Abstract. Cognitive metaphor theory (CMT) has over the past 25 years amply argued that we do not only write and speak but, more importantly, actually think in metaphors. If this tenet of CMT is correct, metaphor should necessarily manifest itself not just in language but also in other modes of communication, such as pictures, music, sounds, and gestures. However, non-verbal and multimodal metaphor have been far less extensively studied than their verbal sisters. The present article provides discussion of work done in this area, focusing on a number of issues that require further research. Some are specific for non-verbal and multimodal metaphor, others invite reconsideration in the realm of linguistic metaphor. These issues include the proposal to distinguish between monomodal and multimodal metaphor; reflections on the distinction between structural and creative metaphor; the question of how verbalizations of non-verbal or conceptual metaphors may affect their possible interpretation; thoughts as to how similarity is created between target and source once the realm of purely verbal discourse is left; and suggestions about the importance of genre for the construal and interpretation of metaphor;

Keywords: Monomodal metaphor; multimodal metaphor; pictorial metaphor; structural versus creative metaphor; similarity in metaphor; genre.

1 Introduction

Andrew Ortony’s edited volume Metaphor and Thought (Ortony 1979) and Lakoff and Johnson’s monograph Metaphors We Live By (Lakoff and Johnson 1980) were milestone publications in the sense that they marked the switch from research into metaphor as a primarily verbal to a predominantly conceptual phenomenon. The “conceptual metaphor theory” (CMT), as the Lakovian-Johnsonian model is habitually referred to, has been a very productive one (e.g.,
Kövecses 1986, 2000, 2002, Lakoff 1987, 1993, Lakoff and Turner 1989, Lakoff and Johnson 1999, 2003, Johnson 1987, 1993, Sweetser 1990, Gibbs 1994, Turner 1996). A key notion in this theory is that “the mind is inherently embodied, reason is shaped by the body” (Lakoff and Johnson 1999: 5; Chapter 3). Very briefly, what this means is the following. Human beings find phenomena they can see, hear, feel, taste and/or smell easier to understand and categorize than phenomena they cannot. It is perceptibility that makes the former phenomena concrete, and the lack of it that makes the latter abstract. In order to master abstract concepts, humans systematically comprehend them in terms of concrete concepts. Thus abstract concepts such as LIFE, TIME, and EMOTIONS are systematically understood in terms of concrete phenomena. Life is understood as a journey (“He’s without direction in his life”; “I’m at a crossroads in my life”) – but also, for instance, as a story (“Tell me the story of your life”; “Life’s … a tale told by an idiot …”). Time is comprehended in terms of spatial motion (“The time for action has arrived”; “Time is flying by”; “He passed the time happily”). Emotions are typically represented by drawing on the domain of forces. (“I was overwhelmed”; “I was swept off my feet”; examples from Lakoff and Johnson 1980, Lakoff 1993, Kövecses 2000). Conceptualisations of many phenomena, CMT proposes, have deeply entrenched metaphorical forms, in which the metaphor’s target (topic, tenor) is abstract and its source (vehicle, base) is concrete. A metaphor’s interpretation boils down to the “mapping” of pertinent features from the source to the target; a mapping that in the case of entrenched metaphors such as the above occurs automatically. Since “concreteness” is apprehended perceptually, metaphorical source domains are strongly rooted in the functioning of the human body. Metaphorical reasoning is thus governed by the “arch” metaphor mind is body (Lakoff and Johnson 1999: 249). A more recent development rooted in CMT is “blending theory” (Fauconnier and Turner 2002). Rather than postulating a target and a source domain, it presents two (or more) “input spaces.” The input spaces have both shared and unique characteristics, and it is this combination that allows for the construal of a so-called “blended space.” Blending theory is a, hitherto mainly descriptive, model claiming to be superior to metaphor theory in being able to account for ad hoc linguistic creativity, metaphorical and otherwise. Hitherto it cannot quite convince (for a critical review of blending theory, see Forceville 2004a), but new work, taking into account pragmatic rhetorical factors, is promising (see Coulson and Pascual, in press).

CMT has inspired conferences (e.g., those organized by the International Cognitive Linguistics Association, and the Researching and Applying Metaphor [RAAM] group), journals (e.g., *Metaphor and Symbol, Cognitive Linguistics*), as well as empirical research (for references see Gibbs 1994, Lakoff and Johnson 1999: 587-88). Its importance is evident: if CMT is basically correct, it provides crucial insights into what, thanks to embodiment, lays claim to being
universal in human cognition, and what is rooted in, or shapes, (sub)cultural differences.

However, CMT is restricted in at least the following very important dimension. Even though Lakoff and Johnson’s characterization of metaphor’s essence as “understanding and experiencing one kind of thing in terms of another” (1980: 5) emphatically avoids the word “verbal” or “linguistic,” the validity of CMT’s claims about the existence of conceptual metaphors depends almost exclusively on the patterns detectable in verbal metaphors. This entails two dangers: in the first place, there is the risk of a vicious circle: “cognitive linguistic research suffers from circular reasoning in that it starts with an analysis of language to infer something about the mind and body which in turn motivates different aspects of linguistic structure and behavior” (Gibbs and Colston 1995: 354; see also Cienki 1998). Clearly, to further validate the idea that metaphors are expressed by language, as opposed to the idea that they are necessarily linguistic in nature, it is necessary to demonstrate that, and how, they can occur non-verbally and multimodally as well as purely verbally. Secondly, an exclusive or predominant concentration on verbal manifestations of metaphor runs the risk of blinding researchers to aspects of metaphor that may typically occur in multimodal representations only. This latter awareness, of course, exemplifies a more general principle. Ever since Marshall McLuhan’s “the medium is the message” (McLuhan, 1964: 24 et passim), it is a truism that as soon as one changes the medium via which a message (including both its factual and emotive aspects) is conveyed, the content of this message is changed as well (see also Bolter and Grusin 1999). Each medium – here defined as a material carrier and transmitter of information – communicates via one or more signalling systems. The medium of non-illustrated books, for instance, exclusively draws on the mode of written language; radio relies on the modes of spoken language, non-verbal sound, and music; advertising billboards on written language and visuals; and post-silent film on visuals, written language, spoken language, non-verbal sound, and music. If, as is argued here, each of these signalling systems (which will henceforward be called “modes”) can cue, independently or in combination, metaphorical targets as well as metaphorical sources, a full-blown theory of metaphor cannot be based on its verbal manifestations alone, since this may result in a biased view of what constitutes metaphor.

In this paper I will sketch how adopting the view that metaphors can assume non-verbal and multimodal appearances can and should guide the research of a new generation of metaphor scholars. I will do so partly by bringing to bear multimodal perspectives on issues already familiar from research by language-oriented metaphor scholars, and partly by discussing issues that have either been neglected by language-oriented metaphor researchers or were simply not pertinent to purely verbal metaphors. The paper
should be seen as a map of mostly uncharted territory, with only a few details inked in, much of it reporting theory-driven analyses and informed speculation awaiting empirical testing. Multimodal metaphor researchers have a vast amount of work to look forward to.

2 Multimodality versus monomodality

In order to distinguish multimodal metaphor from monomodal metaphor, it should first be further clarified what is meant by “mode.” This is no easy task, because what is labelled a mode here is a complex of various factors. As a first approximation, let us say that a mode is a sign system interpretable because of a specific perception process. Acceptance of this approach would link modes one-on-one to the five senses, so that we would arrive at the following list: (1) the pictorial or visual mode; (2) the aural or sonic mode; (3) the olfactory mode; (4) the gustatory mode; and (5) the tactile mode. However, this is too crude a categorization. For instance, the sonic mode under this description lumps together spoken language, music, and non-verbal sound. Similarly, both written language and gestures would have to be part and parcel of the visual, since one cannot hear, smell, taste, or touch either conventionally written language or gestures (although a blind person can feel Braille language and, by touch, perceive certain gestures – for instance those of a statue). If justice is to be done to these distinctions (between images and gestures, between spoken and written language, between spoken language, sounds, and music), other factors need to be taken into account, such as the manner of production (e.g., printed versus Braille letters in relief on paper; signs made with parts of the body versus signs whose use is governed by the grammar and vocabulary rules of a natural language). There are other problematic issues. For instance, what is music and what “mere” sound may differ from one culture, or period, to another. Similarly, it is impossible to assess objectively where music shades off into sound effect. And is “typeface” to be considered an element of writing, of visuals, or of both? In short, it is at this stage impossible to give either a satisfactory definition of “mode,” or compile an exhaustive list of modes. However, this is no obstacle for postulating that there are different modes and that these include, at least, the following: (1) pictorial signs; (2) written signs; (3) spoken signs; (4) gestures; (5) sounds; (6) music (7) smells; (8) tastes; (9) touch.

We can now provisionally define monomodal metaphors as metaphors whose target and source are exclusively or predominantly rendered in one mode. The prototypical monomodal metaphor is the verbal specimen that until recently was identical with “metaphor” tout court, and which has yielded thousands of studies (Shibles 1971, Van Noppen et al. 1985; Van Noppen and Hols 1990; de Knop et al. 2005). A type of monomodal metaphor that has more
recently become the subject of sustained research is pictorial or visual metaphor.¹

An early discussion of metaphor in pictures is Kennedy (1982). The perception psychologist Kennedy takes “metaphor” in the all-encompassing sense of what literary scholars call a “trope” or a “figure of speech” (and which Tversky 2001 calls “figures of depiction”) and identifies some 25 types, including “metonymy,” “hendiadys,” and “litotes.” Kennedy’s attempts to describe an extensive catalogue of “figures of depiction” using the tenor/vehicle distinction that Richards (1965) specifically coined for one of them (i.e., “metaphor”) are sometimes strained. There are other problems. It is a matter for debate whether the names he selects for his examples are always necessarily the best ones, and each trope is illustrated with one or two examples only, making generalizations difficult (the same problem also adheres to Durand 1987). This having been said, Kennedy makes a number of points that are illuminating for a theory of pictorial/visual-metaphor-in-the-narrow-sense. In the first place he argues that for a phenomenon to be labelled a visual metaphor it should be understandable as an intended violation of codes of representation, rather than as being due to carelessness or error. Secondly, Kennedy emphasizes that target and source are, in principle, irreversible, which ensures that what he labels metaphor remains commensurate with a generally accepted criterion in theories of verbal metaphor. Thirdly, Kennedy introduces the helpful notion of “runes”: the kind of non-iconic signs used profusely in comics and cartoons to indicate speed, pain, happiness, anger and many other phenomena by means of straight or squiggly lines, stars, bubbles etc. using surrounding characters or moving objects (see also Smith 1996). In later work, Kennedy elaborates on his theoretical work in various experiments. Kennedy (1993) reports, among other things, how congenitally blind children metaphorically draw a spinning wheel.

Whittock (1990) describes cinematographic metaphor. While his numerous examples are subsumed under ten subtypes, and thus are less wide-ranging than those by Kennedy, they still go beyond metaphor-in-the-narrow-sense, including for instance metonymy and synecdoche. Whittock is criticized by Carroll (1994) for failing to take into account what the latter considers the most typical variety of visual or cinematographic metaphor, the visual hybrid (see also Carroll 1996). Carroll, unlike Kennedy and Whittock, moreover argues that visual metaphors differ from verbal ones in often allowing for reversal of target and source. In Forceville (2002a), expanding on earlier work (Forceville 1988, 1994, 1995, 1996, 2000), I in turn question Carroll’s choice for visual

¹ The topic of metaphor and music also has inspired studies over the past few years, but for lack of expertise in this area I will not dwell on these. For references, see Johnson and Larson 2003; see also Cook 1998, Zbikowski 2002, Thorau 2003, and Spitzer 2004 – the last one rather difficult for laymen.
hybrids as “core filmic metaphor” (Carroll 1996: 218) as well as his proposal for the typical reversibility of target and source in visual metaphors. My argument rests on the claim that Carroll is biased by his exclusive reliance on examples rooted in Surrealist art. My own model, largely developed with respect to advertising representations (but cf. Forceville 1988) and based on Black’s (1979) interaction theory of metaphor (see also Indurkya 1991, 1992; Gineste et al 2000), centres on the answerability of the following three questions: Which are the two terms of the pictorial metaphor?; which is the target and which is the source?; and which is/are the features that is/are mapped from source to target? The last question pertains to the metaphor’s (potential) interpretations: in principle all elements metonymically associated (by a whole community or by a single individual) with the source domain qualify as potential candidates for a mapping. The crucial issue what is actually mapped by a specific addressee in a specific situation is governed by the relevance principle as developed by Sperber and Wilson (1995). (For more discussion of the role of metonymy in metaphor, see various contributions in Dirven and Pöring 2002; for the pertinence of Relevance Theory to the interpretation of pictorial metaphors, see Forceville 1996, chapters 5 and 6.) In this model, Carroll’s examples would rank as one of three (Forceville 2002b) or four (Forceville 2005a, forthcoming a) subtypes of monomodal metaphor. It is to be noted that the distinction between two of these types – metaphor and simile – is also made by Kaplan 1990, 1992; see also Rozik 1994; 1998.)

By contrast to monomodal metaphors, multimodal metaphors are metaphors whose target and source are each represented exclusively or predominantly in different modes. The qualification “exclusively or predominantly” is necessary because non-verbal metaphors often have targets and/or sources that are cued in more than one mode simultaneously. To give a fictive example: imagine somebody wants to cue, for whatever reason, the metaphor CAT IS ELEPHANT pictorially in an animation film. She could do this for instance by depicting the cat with a trunk-like snout and large flapping ears; by showing the cat with a canopy on its back in which a typical Indian elephant rider is seated; by juxtaposing cat and elephant in the same salient pose; or by letting the cat behave (for instance: move) in an elephant-like manner. These variants would constitute monomodal metaphors of the pictorial kind, featuring hybrid, contextual, simile, and integrated subtypes respectively (see Forceville 1996, 2002b, 2005a) – and of course these subtypes could be combined. Now imagine the producer wishes to cue the same metaphor multimodally. She could for instance have the cat make a trumpeting sound or have another cat shout “elephant!” to the first one (note that this is not a case of synaesthesia, since there is no conflation of the two domains). In these cases the source domain ELEPHANT would be triggered in two modes (sound and language, respectively) that are different from the target (visuals). By this token, the metaphor would be
truly multimodal. But, as in the case of the visual mode alone, the producer would of course not have to choose between any of these modes: she could depict the cat with a trunk-like snout and large ears and have it trumpet, and have another cat shout “elephant!” In this case, the source is cued in three modes simultaneously, only one of these (namely: the visual) exemplifying the same mode as the target. In such a case I propose to label the metaphor multimodal. Of course the metonymy cueing the source domain in itself is often chosen for its specific connotations. Both tusks and a trunk trigger ELEPHANT, but the former connotes, among other things, aggressiveness, whiteness, costliness and the latter among other things flexibility, sensitivity, and “instrument-to-spray-water-or-sand-with.” For examples, as well as more discussion of, multimodal metaphors involving (moving) images, see Forceville (1999a, 2003, 2004b, 2005b, forthcoming a, b). There is also a growing literature on multimodal metaphors involving language and gestures (Cienki 1998, Müller 2004, McNeill 1992), in which the gesture-modality cues the source rather than the target domain (McNeill 2005: 45).

3 Structural versus creative metaphor

Lakoff and Turner (1989) have argued that not only metaphors occurring in everyday verbal communication can be traced back to conceptual metaphors, but also those in artistic texts, specifically poetry. Particularly when poems thematize abstract concepts such as life and death, they cannot but draw on the same conceptual metaphors that permeate non-artistic language. Thus Lakoff and Turner cite, and richly illustrate, many passages featuring LIFE and TIME as metaphorical targets (LIFE IS A JOURNEY, LIFE IS A PLAY, LIFE IS BONDAGE, LIFE IS A BURDEN; TIME IS A THIEF, TIME IS A MOVER, TIME IS A DEVOURER) concluding that “although human imagination is strong, empowering us to make and understand even bizarre connections, there are relatively few basic metaphors for life and death that abide as part of our culture” (1989: 26). They acknowledge that the art and craft of good poets resides in finding fresh, original verbal formulations for these conceptual metaphors, and that these formulations resonate both with the rest of the poem and with the extra-textual knowledge of the reader. I take it that Lakoff and Turner allow that this, in turn, may result in temporary readjustments of the basic level conceptual metaphors, and thus that they would agree that the linguistic level of the metaphor is not a mere illustration or exemplification of the pre-existing basic conceptual level. But not all verbal metaphors in poetry, as Lakoff and Turner acknowledge, reflect basic conceptual metaphors. While conventional metaphors can be expressed either in common or in idiosyncratic language, “modes of thought that are not themselves conventional cannot be expressed in conventional language” (1989: 26) and hence
require idiosyncratic language. (If I understand her correctly, Renate Bartsch
would probably object to the label “conceptual metaphor” for the metaphorical
schema that underlies such novel metaphors. She stipulates that a phenomenon
deserves the name of “concept” only if it has a stable interpretation in a
community, and hence must by definition have been “linguistically explicated”
(Bartsch 2003: 50). A stable interpretation requires that the community agrees on
the phenomenon’s characteristic features and these, in turn, reveal themselves in
the true predications that can be used for it. Novel metaphors focus attention on
non-characteristic features and therefore, in Bartsch’ reasoning, cannot (yet) have
the status of being “conceptual”). Often, the border between conventional
metaphors and idiosyncratic ones is difficult to draw, not least because
conventional metaphors may have idiosyncratic extensions. Lakoff and Turner are
thus completely clear about the fact that by no means all poetic metaphors are
conventional ones, but the bulk of their examples and discussions pertain to the
latter. This is unsurprising, since their aim is to show that poetic metaphors
normally tap into conventional ways of thinking (“great poets can speak to us
because they use the modes of thought we all possess,” 1989: xi).

Nonetheless Lakoff and Turner’s account raises some questions. In the
first place, it is not clear how representative their chosen examples are of poetic
metaphors in general. While they convincingly show that structural metaphors
pervade poetry, the relative distribution of metaphors may depend on time and
place: older poetry, or non-Western poetry, may feature more, or less, instances
of a given metaphor. Gevaert (2001), basing herself on corpus-based data,
questions for instance the embodied, “timeless” status of ANGER IS HEAT claimed
for conceptualisations of anger in Lakoff (1987). She demonstrates, among
other things, that in the Old English period SWELLING was a much more
important source domain in ANGER metaphors than HEAT, and speculates that the
latter’s growing popularity in more recent periods may well be due to the
humoral theory that dominated mediaeval times; that is, to cultural no less than
embodied knowledge (see also Gevaert 2005). Moreover, a systematic, corpus-
based analysis might reveal that many poetic metaphors are not so easily
amenable to conventional ones. Numerous poetic metaphors may simply not
have abstract concepts such as LIFE, DEATH, TIME, PURPOSE as their target
domain. As pointed out by Grady, they may more often be “resemblance
metaphors” than, possibly (near)universal, “correlation metaphors” (Grady
1999). This would not invalidate Lakoff and Turner’s impressive findings, but
their one-sided emphasis on correlation and generic-level metaphors in poetry
may inadvertently lead to an uncritical acceptance of the view that most poetic
metaphors are of this kind.

Lakoff and Turner describe generic-level metaphors as metaphors which are minimally specific in two senses:
“they do not have fixed source and target domains, and they do not have fixed lists of entities specified in the
mapping” (1989: 81). They introduce the term using the example of the EVENTS ARE ACTIONS metaphor which
they contrast with LIFE IS A JOURNEY, one of its specific-level instantiations.
One important difference between conventional and idiosyncratic metaphors is that the interpretation of the latter is, by definition, far less governed by entrenched, pre-existing correspondences between the schematic structures in target and source. It is only by downplaying this difference that Lakoff and Turner can say that “the preservation of generic-level structure is, we believe, at the heart of metaphorical imagination, whether poetic or ordinary” (1989: 83; for other critical accounts of this view see Stockwell 1999; Crisp 2003).

Secondly, we should not forget that a metaphor can also conceptualise the concrete in terms of the concrete. Lakoff and Turner, to be fair, are aware of this. They discuss at some length the Elizabethan notion of the “Great Chain of Being” (see e.g., Tillyard 1976), which endorsed the idea of “natural” hierarchies within various types of creatures – angels, humans, birds, mammals, etc. – and state that “THE GREAT CHAIN METAPHOR can apply to a target domain at the same level on the Great Chain as the source domain” (Lakoff and Turner 1989: 179, emphasis in original). Put differently, metaphors may have targets as well as sources that are directly accessible to the senses. But given that CMT puts great emphasis on metaphor’s role in conceptualising the abstract in terms of the concrete, this possibility receives rather scant attention, while CONCRETE IS CONCRETE metaphors are particularly relevant once we leave the realm of the purely verbal. In the case of monomodal metaphors of the pictorial variety, both target and source are depicted. Since, in advertising, metaphorical targets usually coincide with promoted products, the targets are depicted – and hence are necessarily concrete: a beer brand is depicted as a wine; an elegant watch as a butterfly, a close-fitting bathing suit as a dolphin’s tight and supple skin (examples from Forceville 1996). The same holds for metaphors in feature films (Whittock 1990, Forceville 2005b). In short, to what extent monomodal metaphors of the non-verbal variety and multimodal metaphors are amenable to the correlation metaphors that are the centre of attention in CMT is an empirical question. Some of them no doubt do; for instance, the personification of commodities is a very familiar marketing strategy, and ties in with CMT views (Lakoff and Turner 1989: 72). But many pictorial and multimodal metaphors are of the OBJECT A IS OBJECT B TYPE. Traditional CMT has not much to say about these. Even Lakoff and Turner’s (1989) invocation of the Great Chain metaphor is only of limited use here, since it depends on typological hierarchies that may be subverted, or simply irrelevant, in creative metaphors, many of which function in contexts creating highly specific, ad hoc metaphorical resemblances (see Black 1979).

There is a third aspect in which CMT has a somewhat one-sided emphasis. As discussed above, the typical source domain’s concreteness has in CMT been traditionally connected to the notion of “embodiment.” The embodied nature of source domains emphasises their physical nature: it is
human physical interaction with the world that familiarizes humans with it to such an extent that the resulting knowledge structures can in turn be mapped onto abstract concepts. Knowledge about source domains is not simply a matter of embodiment, however, but also of cultural connotations, as Lakoff and Turner acknowledge (1989: 66). More recent studies have demonstrated in a variety of ways how the structure of source domains – and the salient (and hence: easily mappable) elements in it – is influenced by culture (Shore 1996; Gibbs and Steen 1999; Yu 1998; Kövecses 2005). Indeed, the cultural connotations that are metonymically related to a source domain are often more important for potential mappings to a target than its embodied aspects. In a Dutch commercial promoting a Gazelle bicycle in terms of a dressage horse the embodied mapping of “riding a horse” to “riding a bicycle” is less important for the interpretation of the metaphor than the mapping of the cultural connotations from the dressage horse’s owner, champion Anky van Grunsven, to the prospective buyer and user of the bike. Similarly, while advertising a high-tech Senseo coffee machine in terms of a motorbike certainly has embodied aspects, the subcultural connotations associated with motorbike-riding evoked by Steppenwolf’s “Born to be wild,” audible on the commercial’s soundtrack, are at least as important in the mapping (Forceville 2004b, forthcoming b, Forceville et al. in preparation a). The relevant “similarity” that is created between target and source pertains to these connotations more than anything else. The examples bear out Bartsch’ observation that in metaphor “the role of similarity is not restricted to the identity of internal properties of objects and situations, rather similarity also is due to identity of external contiguity relationships between objects, between situations, and it is due to relationships of objects and situations with emotional attitudes, desires, and behavioural dispositions of people” (Bartsch 2003: 52). Indeed, it might be ventured that “a single, embodied correspondence between target and source is enough to trigger a wide range of further ‘cultural’ correspondences between target and source, and hence of inferences about the target” (Forceville et al., forthcoming). The old adage that a picture tells more than a thousand words should not blind us to the fact that pictures and other multimodal representations seldom communicate automatically or self-evidently. As in verbal metaphors, it is connotations rather than denotations of source domains that get mapped in metaphors, and these may substantially differ from one (sub)cultural group to another (see e.g., Maalej 2001). Even when non-verbal metaphors verge toward the conventional, as in comics representations of ANGER (Forceville 2005c; Eerden 2004; see Simons 1995 for multimodal instantiations of structural metaphor in pre-election TV spots promoting political parties), it may well be the case that, as in their verbal counterparts (see Kövecses 1986, 2000, 2005), there is cultural variation (cf. Shinohara and Matsunaka 2003).
The three issues briefly mentioned above (“metaphors frequently have concrete rather than abstract targets”; “many metaphors with idiosyncratic surface manifestations do not reflect correlational metaphors”; “what gets mapped from source to target domains are often cultural no less than embodied features”) are thrown in relief when the study of purely verbal metaphors gives way to that of non-verbal and multimodal metaphors. That is, generalizing observations on metaphor based on the systematic investigation of verbal specimens need to be considered afresh by testing these observations in metaphors occurring in other modes.

4 The verbalization of non-verbal metaphor and the nature of “similarity”

We have seen that within the CMT paradigm, most surface metaphors should be amenable to a pre-existing conceptual A IS B format. Inevitably, in order to discuss the metaphor, this A and B must be named, i.e., rendered in language. It is by no means a foregone conclusion, of course, that the “language of thought” is actually a verbal language. The convention to verbalize the image-schematic structures underlying surface metaphors by using SMALL CAPITALS – useful inasmuch as this facilitates analysing them – should not blind scholars to a number of consequences that seem to me more problematic in the discussion of representations that are not (exclusively) verbal ones than of purely verbal ones. One of these consequences is that it is the analyst’s responsibility to find an adequate or acceptable verbal rendering of the metaphor’s underlying image-schematic level, but such a verbalization, even though used as a convenient shorthand, is never neutral. The design of the Senseo coffee machine suggests the posture of somebody bending over and modestly offering something (i.e., a cup) on a plate. But should this awareness result in the verbalization COFFEE MACHINE IS SERVANT, or is COFFEE MACHINE IS BUTLER more appropriate? Although “servant” and “butler” share many features, they also differ: “butler” is more specific, and may in some people (but not in others) evoke connotations of Britishness and standards of service that “servant” does not. As a result, the mappings suggested by the two verbalizations may differ. Bartsch might conclude that this very inability to agree on a single verbalization of the source domain that is shared within a community shows that the source has no conceptual status, and reflects a “quasi-concept” at best (Bartsch 2003: 50). However, to the extent that there is a community that recognizes the source as cueing a serving person, the source admits predicates understood as “true” in the community (such as “is there to serve the user,” “obeys your requests,” and “is almost always available”). In a visually literate society, a vast number of endlessly repeated and recycled images (such as famous paintings, photographs, film shots, flags, logos, animation characters) evoke specific phenomena and
events in a clichéd, shorthand manner widely shared within a community, and hence arguably aspire to conceptual status. But this speculation leads us far beyond the concerns of the present chapter and deserves in-depth reflection elsewhere.

Another consequence is that verbalization of a non-verbal metaphor is necessarily a conscious action, and a fairly unusual one at that. It is only the scholar writing an academic paper who, to be able to discuss a multimodal metaphor, needs to resort to GAZELLE BICYCLE IS VAN GRUNSVEN DRESSAGE HORSE; the metaphor is not verbalized in this form in the commercial, and it is an open question whether the construal of a non-verbal metaphor requires its verbalization by the audience. The question can be reformulated as follows: does comprehension of a non-verbal or multimodal metaphor imply that recipients “mentally” verbalize the metaphor? It is an important question, but extremely difficult to test empirically.

Since non-verbal modes of communicating by definition do not have the “is” or “is like” in order to signal a metaphorical identity relation between two entities belonging, in the given context, to different categories, one issue that deserves attention is by what stylistic means the similarity is triggered. Of course this holds for verbal metaphors that do not have the paradigmatic A is (like) B format as well (cf. Brooke-Rose 1958; Goatley 1997). But whatever means are chosen in this latter case, the cues are themselves of a verbal nature. In non-verbal and multimodal metaphors, the signals that cue metaphorical similarity between two phenomena are different, and bound to differ depending on the mode(s) in which the metaphorical terms are represented. Here are some possibilities that are deployed in isolation or in combination:

**Physical resemblance.** This can only function as a trigger in the case of monomodal metaphors: only a visual representation can physically resemble another visual representation; only a sound can physically resemble another sound. In the case of visual resemblance, there is a range of choices: two things can resemble one another because they have the same size, colour, position, posture, texture, materiality, etc. Note that the resemblance need not reside in the “things” themselves, but may surface in their manner of representation: they are for instance photographed from the same unusual angle, or filmed with the same unusual camera movement.

**Filling a schematic slot unexpectedly.** Placing a thing in a certain context may strongly, even inescapably, evoke a different kind of thing, namely the thing for which the given context is the natural or conventional place. Put differently, we encounter deviations from typical gestalts or schemas. For example, when in a musical environment a violin case contains a monkey wrench, this suggests the metaphor MONKEY WRENCH IS VIOLIN.

**Simultaneous cueing.** If two things are signalled in different modes, metaphorical identification is achieved by saliently representing target and
source at the same time. For instance a kiss could be accompanied by the sound of a car crash, of a vacuum cleaner, or of the clunking of chains, to cue metaphorical mappings of, say, disaster, dreary domestic routine, and imprisonment, respectively. Alternatively, in a variant on the previous mechanism, two disparate things can be linked because of an unexpected filling of a slot, as when a photograph of a kiss has the caption “imprisonment.”

5 The influence of genre

Human beings in most cases appear to construe a text automatically, very quickly, and probably largely subconsciously as belonging to one genre rather than to another. Anecdotal support for this claim is the experience of channel-surfing: seasoned TV watchers guess in a split second what kind of programme they surf into (and decide at once whether they want to spend time with it. Another illustration for the claim is the very funny, self-reflexive trailer for the film Comedian (http://www.geocities.com/tvtranscripts/misc/comedian.htm). A man, Jack, sits in a booth in a sound studio to record the voice-over for the trailer, but as soon as he has said a few words he is impatiently interrupted by the director:

Jack: “In a world where laughter was king …”
Director: No “in a world,” Jack ….
Jack: What do you mean, “No ‘in a world’”? 
Director: It’s not that kind of movie.
Jack: Oh … OK … “In a land that …”
Director: No “in a land” either …
Jack: “In the time …”
Director: No, I don’t think so.

In a rapid exchange Jack makes one abortive attempt after another – “One man …” “When your life is no longer your own …” “When everything you know is wrong …” “In an outpost …” “On the edge of space …” – only to be cut short by the director straightaway. For present purposes the point to be made is that, to an audience with expertise in the area of film, the few words uttered by Jack suffice to cue an entire genre. Finally, Hayward (1994) offers empirically attested support for the claim that people are able to decide very fast to what genre a text belongs. Hayward found that almost 80% of experimental subjects, given randomly selected passages of history or fiction writing, recognized the genre of the work even on the basis of very short passages (5 to 15 words).

The genre within which a text (in whatever medium) is presented, or the genre to which it is attributed, determines and constrains its possible
interpretations to an extent that is difficult to overestimate (see Altman 1999; Zwaan 1993; Steen 1994; Forceville 1999b, 2005d; Charteris-Black 2004). For this reason, it is important to study how genre has an impact on the production and interpretation of metaphors (monomodal and multimodal alike). In advertising, for instance, the targets of metaphors often coincide with the product promoted (Forceville 1996). This is to be expected: an advertisement or commercial predicates something about a product, brand, or service, and this neatly and naturally fits the metaphor’s TARGET IS SOURCE format. Moreover, the features mapped from source to target are positive ones (unless the metaphor is used to disqualify a competitor’s brand, in which case the mapped features are typically negative). But in feature films, there is no phenomenon that in a similar, “natural,” way qualifies as a metaphorical target. Metaphors in artistic narratives pertain to phenomena that, for whatever reason, are deemed salient by their producers. These can be protagonists, but also objects. The mapped features will often be less clear-cut, and may have a richer “aligned structure” (Gentner and Loewenstein 2002) than those in advertising.

Metaphors in artistic representations may also differ in other respects from those in commercial messages. For instance, while in commercials there will seldom be a question what is target (usually: the product) and what is source in a metaphor, an artistic narrative may give rise to two different construals of a metaphor: both A IS B and B IS A is appropriate. (While Carroll 1994, 1996 calls such metaphors “reversible,” I prefer to say that, in the given context, both the metaphors A IS B and B IS A are pertinent, in order to retain the notion that target and source in a metaphor are, in principle, irreversible.) Commensurate with this, metaphors in artistic contexts presumably allow for greater freedom of interpretation than do metaphors in commercials (cf. also Shen 1995).

Another parameter that deserves further research is whether any of the subtypes of pictorial metaphor or of the manifold varieties of multimodal metaphor can be systematically related to certain text genres. For instance, it seems that commercial advertising seldom makes use of the hybrid variety of pictorial metaphors (in Forceville 1996 these were called MP2s). Again, this makes sense: if metaphorical targets typically coincide with products, advertisers would want their product portrayed in their entirety, and not in a manner that might evoke connotations of incompleteness or mutilation. Hybridizing it with a metaphorical source domain would not fit in this goal. By contrast, in animation films, or science fiction films, no such problem arises.

Finally, it would be worthwhile to investigate whether in the case of multimodal metaphors there are any systematic correlations between textual genres and the modes in which a target and source are represented. In advertisements, the visual mode is typically used for representing the target – and this may well be true for different genres as well. But perhaps alternative
patterns in the choice of mode for the source domain are detectable in different types of texts, while this may also change over time within a genre.

6 Concluding remarks

Researching multimodal metaphor, in short, is a natural next step in the further development of metaphor studies – a development in which theoretical reflection will have to go hand in hand with empirical testing. If creative and conventional metaphor are key factors in human thinking, and if human thinking is reflected in more than verbal manifestations alone, investigating multimodal metaphor is highly worthy of extensive scholarly effort. Given its long disciplinary tradition, the robust insights of metaphor scholarship can in turn fruitfully feed into the budding field of multimodality in general (Kress and Van Leeuwen 1996, 2001; Ventola et al. 2004). Genres to be investigated include advertising (Forceville 1996, 2003, forthcoming a; Phillips 2003; McQuarrie and Mick 2003; Wiggin and Miller 2003.), political cartoons (El Refaie 2003), film (Whittock 1990, Forceville 1999a, 2005b, Rohdin 2003), oral speech accompanied by gestures (Cienki 1998, Müller 2004, McNeill 1992, 2005), design (Cupchik 2003; Van Rompay 2005). And inasmuch as multimodal representations (in the form of advertising, videoclips, games, TV-formats, mainstream films, animation) travel faster and more easily across the world than verbal ones, examining their metaphorical manifestations will help focus on what remains stable and what changes in cross-cultural communication. Furthermore, such work may provide the starting point for how other tropes besides metaphor can assume multimodal appearances (e.g., metonymy, irony, hyperbole, oxymoron, see Kennedy 1982; Gibbs 1993; Teng and Sun 2002). Here the analysis of multimodal metaphor ties in with the study of rhetoric. In a global society in which media are increasingly used, or abused, as mouthpieces for the views of powerful factions (politicians, industrial tycoons, religious leaders), the critical analysis of the tools of persuasive discourse in the broadest sense constitutes an excellent interface between academic research and its possible usefulness in the world beyond its walls.

On virtually all of the issues discussed in this article, Forceville and Urios-Aparisi (in preparation b) will provide new insights.

Acknowledgment I am indebted to Francisco Ruiz-de Mendoza and René Dirven for comments and suggestions on earlier drafts of this paper.
References

Altman, Rick

Bartsch, Renate

Black Max

Bolter, Jay David, and Richard Grusin

Brooke-Rose, Christine

Carroll, Noel

Carroll, Noel

Charteris-Black, Jonathan

Cienki, Alan

Cook, Nicholas

Coulson, Seana, and Esther Pascual
in press For the sake of argument: Mourning the unborn and reviving the dead through conceptual blending. *Annual Review of Cognitive Linguistics* 4.

Crisp, Peter

Cupchik, Gerry D.

Dirven, René, and Ralf Pörings (eds)

Durand, Jacques

Eerden, Bart

El Refaie, Elisabeth

Fauconnier, Gilles and Mark Turner

Forceville, Charles

Forceville, Charles

Forceville, Charles

Forceville, Charles

Forceville, Charles

Forceville, Charles

Forceville, Charles

Forceville, Charles
Forceville, Charles

Forceville, Charles

Forceville, Charles

Forceville, Charles
2004b The role of non-verbal sound and music in multimodal metaphor. In: Henk Aertsen, Mike Hannay, and Rod Lyall (eds), *Words in their Places: A Festschrift for J. Lachlan Mackenzie*, 65-78. Amsterdam: Faculty of Arts, VU Amsterdam.

Forceville, Charles

Forceville, Charles

Forceville, Charles

Forceville, Charles

Forceville

Forceville

Forceville, Charles, Paul Hekkert, and Ed Tan

Forceville, Charles, Gerry Cupchik, and Michelle Hilscher
in prep. a The interpretation of multimodal metaphors in Dutch commercials
by Dutch and Canadian subjects [provisional title].

Forceville, Charles, and Eduardo Urios-Aparisi (eds)

Gentner, Dedre, and Jeffrey Loewenstein

Gevaert, Caroline

Gevaert, Caroline

Gibbs, Raymond W., Jr.

Gibbs, Raymond W., Jr.

Gibbs, Raymond W. Jr. and Gerard J. Steen (eds)

Gibbs, Ray W. jr., and Herbert L. Colston

Gineste, Marie-Dominique, Bipin Indurkhya, and Véronique Scart

Goatley, Andrew

Grady, Joe

Hayward, Malcolm

Indurkhya, Bipin
Indurkhya, Bipin


Johnson, Mark


Johnson, Mark


Johnson, Mark, and Steve Larson


Kaplan, Stuart J.


Kaplan, Stuart J.


Kennedy, John M.


Kennedy, John M.


Knop, Sabine de, René Dirven, Ning Yu, and Birgit Smieja (eds)


Kövecses, Zoltán


Kövecses, Zoltán


Kövecses, Zoltán


Kövecses, Zoltán


Kress, Gunther, and Theo van Leeuwen


Kress, Gunther, and Theo van Leeuwen

Lakoff, George

Lakoff, George

Lakoff, George, and Mark Johnson

Lakoff, George, and Mark Johnson

Lakoff, George, and Mark Johnson

Lakoff, George, and Mark Turner

Maalej, Zouhair
2001 Processing pictorial metaphor in advertising: A cross-cultural perspective. Academic Research, 1: 19-42 [Sfax, Tunisia].

McLuhan, Marshall

McNeill, David

McNeill, David

McQuarrie, Edward F., and David Glen Mick

Müller, Cornelia

Noppen, Jean-Pierre van, Sabine de Knop, and René Jongen (eds)
Philadelphia: John Benjamins.  
Noppen, Jean-Pierre van, and Edith Hols (eds)  
Amsterdam/Philadelphia: John Benjamins.  
Ortony, Andrew (ed.)  
Phillips, Barbara J.  
Richards, I.A.  
Rohdin, Mats  
Rompay, Thomas van  
Rozik, Eli  
Rozik, Eli  
Shen, Yeshayahu  
Shibles, Warren A.  
Shinohara, Kazuko, and Yoshihiro Matsunaka  
Shore, Bradd  
Simons, Jan  
1995 *Film, Language, and Conceptual Structures: Thinking Film in the Age of Cognitivism*. Unpublished PhD dissertation, Department of Film and TV Studies, University of Amsterdam, The Netherlands.  
Smith, Ken

Sperber, Dan and Deirdre Wilson


Spitzer, Michael


Steen, Gerard J.


Stockwell, Peter J.


Sweetser, Eve E.


Teng, Norman Y., and Sewen Sun


Thorau, Christian


Tillyard, E.M.W.


Turner, Mark


Tversky, Barbara


Ventola, Eija, Cassily Charles, and Martin Kaltenbacher


Whittock, Trevor


Wiggin, Amy A., and Christine M. Miller


Yu, Ning


Zbikowski, Lawrence M.

Zwaan, Rolf