

INTEGRATING SYSTEMIC FUNCTIONAL AND COGNITIVE APPROACHES TO MULTIMODAL DISCOURSE ANALYSIS

Dezheng Feng

The Hong Kong Polytechnic University

Elaine Espindola

The Hong Kong Polytechnic University

Abstract

This study explores the complementarities between systemic functional and cognitive metaphorical approaches to multimodal discourse analysis. Their common concern with the construal of human experience and their shared theoretical foundation of viewing language and other semiotic systems as meaning-making resource make it possible to integrate them in analyzing linguistic and multimodal data. Meanwhile, as far as multimodal discourse analysis is concerned, the theoretical strengths of systemic functional visual grammar and multimodal metaphor theory are able to bridge existing gaps. On the one hand, the systemic functional framework provides a comprehensive modeling of the visual realization of metaphor; on the other hand, conceptual metaphor theory provides an epistemological status to the semiotic description of visual images. Therefore,

Ilha do Desterro	Florianópolis	nº 64	p. 085- 110	jan/jun 2013
------------------	---------------	-------	-------------	--------------

from the results here obtained, it is possible to conclude that the integration of these two major theoretical approaches is significant to furthering our understanding of multimodal discourse.

Keywords: Systemic Functional Theory; Conceptual Metaphor Theory; Multimodal Discourse Analysis.

1. Introduction

The current state of the art of multimodal discourse analysis shows that visual images are mainly analyzed based on two theoretical paradigms: Systemic Functional Grammar (SFG henceforth) (Halliday & Matthiessen, 2004) and Cognitive Metaphor Theory (CMT henceforth) (Lakoff & Johnson, 1980). These two theoretical foundations give rise to two approaches to visual analysis: systemic functional visual grammar (e.g. Kress & van Leeuwen, 2006; O’Toole, 1994) and multimodal metaphor theory (e.g. Forceville, 1996; Forceville & Urios-Aparisi, 2009). While these approaches have produced significant insights to the understanding of visual semiosis, they present some issues in what concerns their analytical and explanatory nature that hindered their independent expansion. On the one hand, the issue holding SF visual grammar back is the subjectiveness of assigning semiotic values to visual images (Zhu, 2007). After exploring the compositional meaning of newspaper layout using their framework, Kress and van Leeuwen (1998, p. 218) also admitted that “the major challenge to our approach is the epistemological status of our claim). These authors also explain: “For instance, how can we know that left and right, top and bottom have the values we attribute to them, or more fundamentally, have any value at all?” (Kress & van Leeuwen, 1998, p. 218). On the other

hand, the issue surrounding the study of multimodal metaphor is that researchers have not yet proposed a theoretical framework able to model the several types of visual metaphors (Feng, 2011a). As El Refaie (2003, p. 80) also observes, “there seems to be a whole range of different forms through which metaphorical concepts can be expressed visually”.

Based on the observations above, we intend to provide the epistemological status of systemic functional visual grammar from a cognitive perspective and provide a systematic modeling of the visual mechanisms of metaphor construction based on the systemic functional framework, so as to integrate SFG and CMT, two perspectives that were isolated from one another as presented in in Feng (2011a; 2011b). However, before elaborating how such theoretical integration can be put to action in Section 3 and Section 4, we will briefly discuss the theoretical foundations for integrating these two relevant theoretical paradigms in Section 2. Finally, in Section 5 we draw conclusions in what concerns the integration of systemic functional and cognitive perspectives as a significant approach for exploring and understanding visual semiosis.

2. Systemic Functional and Cognitive Approaches to Multimodal Analysis: Space of Integration

In this section, we will introduce the foundations of SFG and CMT, as well as their applications to multimodal discourse analysis. We then further the discussion to how these approaches may complement each other when analyzing visual images.

SFG models language as sets of inter-related systems of choice that are metafunctionally organized. The “systemic” principle regards grammar as systems of paradigmatic choice that are

modeled as system networks. The “functional” principle implies that language simultaneously provides resources for construing three interdependent metafunctions, which in turn construe three layers of meaning, namely, ideational meaning, interpersonal meaning and textual meaning. Social semioticians argue that these principles are applicable to non-linguistic resources as well, resulting in the development of metafunctional frameworks for semiotic resources such as images, architecture and mathematical symbols (e.g. Kress and van Leeuwen, 2006; O’Toole, 1994; O’Halloran, 2005). According to Kress and van Leeuwen (2006), visual images, like language, fulfill the metafunctions of the representation of the experiential world (representational meaning), the interaction between the participants represented in a visual design and its viewers (interactive meaning), and the compositional arrangements of visual resources (compositional meaning). This framework is illustrated below in Figure 1.

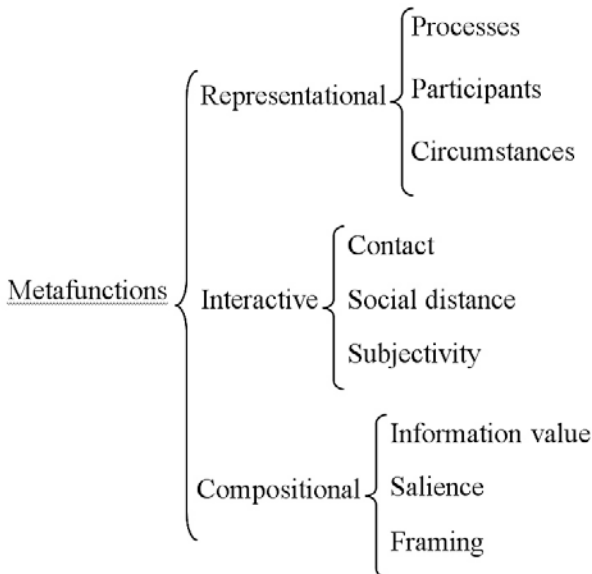


Figure 1: Kress and van Leeuwen’s (2006) metafunctional framework

The framework proposed by Kress and van Leeuwen (2006), especially in what concerns the interactive and compositional meanings, has frequently been challenged as being to a certain extent subjective, that is, the authors assign semiotic values to different camera angles and different locations in the visual space based on observation alone. Taking an SFL perspective to look into the relationship between semantics and the lexico-grammatical configuration of visual images, it is possible to observe that the grammar of visual images is not as conventionalized as that of written and spoken language and therefore needs to be supported by epistemological evidence coming from other areas of enquiry. In this paper, the epistemological status of visual grammar is explored from a cognitive perspective. Specifically, the ‘how do we know’ question quoted in Section 1 is answered by the conceptual metaphor theory proposed by Lakoff and Johnson (1980). They argue that most abstract concepts are metaphorically understood in terms of other (concrete) concepts and this process forms what is known as conceptual metaphor. The formulation of a conceptual metaphor consists of “A IS B” in small capital letters. An abstract concept A is understood in terms of a concrete concept B. A is known as the target domain and B is called the source domain. Understood in this way, the relationship between visual space and semiotic value becomes a metaphorical mapping between the source domain and the target domain, as illustrated in Figure 2 below. This mapping can be considered as the master metaphor, which entails every kind of sub-mappings between elements of the visual space (e.g. up, down, center) and elements of the semiotic meaning (e.g. ideal, real, important). This can be applied to the reading of camera angle as presented in Figure 2 below:

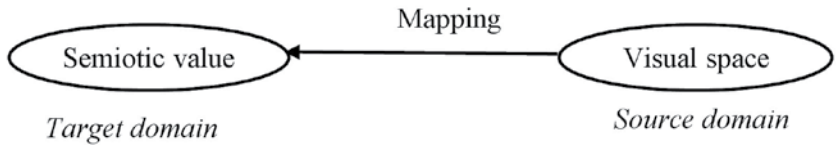


Figure 2: The meaning of visual space as a master metaphor

In this approach, instead of assigning semiotic values to the camera and visual space, we are led to question “how do we understand the abstract concepts through camera positioning and visual space?” Therefore, Kress and van Leeuwen’s (2006) descriptive model that puts forward the idea that “Left is Given” becomes, for example, in the cognitive approach, GIVEN IS LEFT, as tentatively illustrated in Figure 3. This is a significant propositional change because, in systemic functional terms, what was “given information” (Token) in the first proposition – “Left” – becomes “new information” (Value) in the second. That is, in the first proposition we are assigning meaning to the visual space, in the second the abstract concept is being understood through a concrete phenomenon (Feng, 2011b, p. 58).



Figure 3: From description to understanding

The mapping between semiotic value and visual space is not arbitrary, but based on the individual’s embodied experience. The notion of *experiential basis* is fundamental to the analysis of metaphors as “it is only by means of these experiential bases that

the metaphor can serve the purpose of understanding” (Lakoff & Johnson, 1980, p. 20). For the purposes pursued here it is precisely the experiential basis that will aid answering the “how can we know” question: seen as a metaphor, the validity of the association becomes the existence and functioning of the experiential basis. Taking the left/right orientation as an example, in most modern cultures, human beings write and read from left to right (Rogers, 2005), so the left side is taken as Given information and the right side as New information. This is a metaphorical process in that information value is understood in terms of spatial relations, according to our culturally-determined experience of reading. The process can be captured by the metaphor: INFORMATION VALUE IS READING PATH, which entails GIVEN IS LEFT and NEW IS RIGHT. However, the precondition is the experience of writing from left to right, so in cultures where people write from right to left (as in Ancient China, or some Arabic countries), the model is questionable. Moreover, the experiential basis must be functioning with respect to other factors, that is, GIVEN IS LEFT is valid when the reading path is not interrupted or overwhelmed by other factors such as visual salience through size and color. In this way, the epistemological status and the “establishing condition” of the mapping are both in place. The framework for reformulating other dimensions of Kress and van Leeuwen’s (2006) “visual grammar” as systems of metaphors is developed further in Section 3.

Roughly around the same time of the emergence of visual grammar, cognitive linguists also further explored CMT by building it upon the realization of conceptual metaphors for visual images (e.g. Forceville, 1996; Carroll, 1996). However, as Feng (2011a) points out, cognitive theorists have not yet presented a framework that is capable of describing the visual mechanism responsible for realizing metaphors and types of visual metaphors (see also El Refaie,

2003). With this context in view, this paper argues that a thorough modeling of the visual realization of metaphor has to be based upon a systematic description of visual images. In this regard, systemic functional visual grammar (see Figure 1) provides a more thorough and systematic description of visual images than what was available before and, to our view, it is a paramount candidate as an apparatus for modeling the visual realization of metaphors. The application of the systemic functional framework for modeling visual metaphor is more fully elaborated in Section 4.

3. The Cognitive Metaphorical Interpretation of Systemic Functional Visual Grammar

This section aims at reformulating the interactive and compositional dimensions of visual grammar, as described by Kress and van Leeuwen (2006), into systems of metaphors. For the interactive meaning, the focus of attention is given to the dimensions of social distance and subjectivity, which are visually realized by shot distance and camera angle, respectively. Subjectivity can be sub classified into involvement and power relations, and these two are in turn realized by camera angles at the horizontal axis and the vertical axis, respectively. For the compositional meaning, the focus is put on information value, which is realized by the location of the object in the visual space. This system is represented in Figure 4. Based on the approach developed in Section 2, the semiotic systems in Figure 4 are reformulated as inter-related systems of metaphors, as displayed in Figure 5 and Figure 6. By doing so, it is possible to validate the relationship existing between camera positioning/visual space and

semiotic value provided by the experiential basis of the mapping, as we elaborate below (see Feng, 2011a for a detailed discussion).

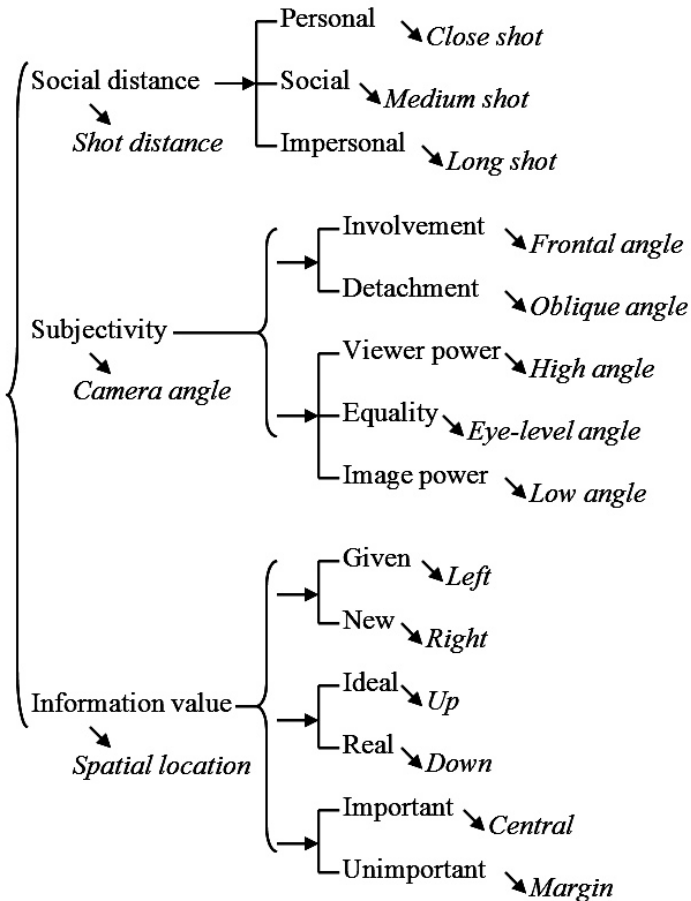


Figure 4 The interactive and compositional dimensions of visual grammar

The metaphorical meaning of camera positioning is adopted based on the iconic nature of image. Therefore, shot distance reproduces the structure features of physical distance in real life

and camera angle reproduces the features of the ways we look at and interact with people. The basis of the mapping between physical distance (hence shot distance) and social distance is well established in the study of proxemics (e.g. Hall, 1969), but such conceptualization is out of the scope of the present study and, thus, will not be elaborated here. The mapping between image-viewer power relation and vertical camera angle is based on the structural features of real-life situations in which we look up to powerful people and look down upon less powerful ones (Messaris, 1994). The mapping of involvement and horizontal camera angle is based on real life situations in which human beings face the person they want to interact with and gaze at him/her, and later turn their faces (gaze) away if there is no longer the desire for interaction. Figure 5 and 6 below illustrate such inter-relations of metaphorical meaning.

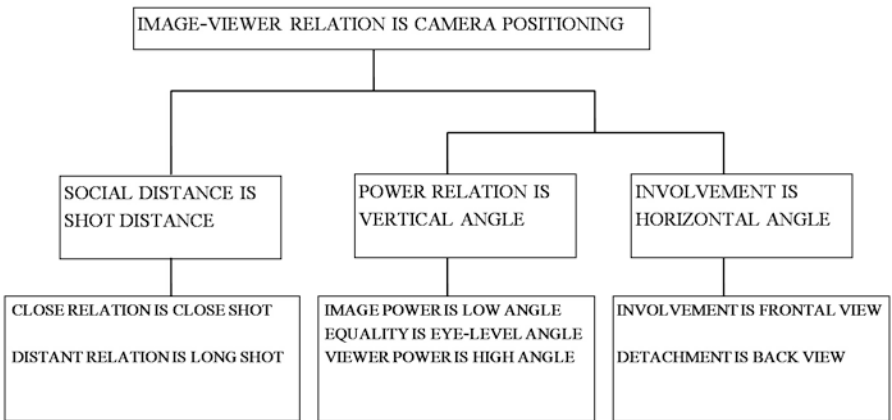


Figure 5 The metaphor of camera positioning (Feng, 2011a, p. 27)

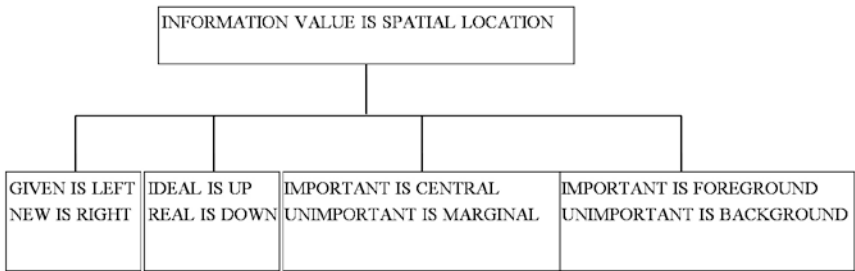


Figure 6 The metaphor of the visual space

The information values of given/new, ideal/real and important/unimportant are realized by the spatial orientations of left/right, top/bottom and central/marginal respectively. Similar to interactive meaning, visual compositional meaning is also derived from individual’s embodied shared experience. GIVEN IS LEFT/NEW IS RIGHT is based on the experience that in most cultures, language users write and read from left to right, so that left is taken as given information and right is taken as new. In IDEAL IS UP, “ideal” has two different but related entailments, they are: desirable and unrealistic (Feng, 2011b). DESIRABLE IS UP is synonymous to the well-established metaphor GOOD IS UP (Lakoff & Johnson, 1980) and need no further explanation. UNREALISTIC IS UP uses a different sense of the common-sense notion of “up”, meaning here: high. This may be explained by human beings natural shared experience that high located materials/items/things (almost as out of reach) are often more difficult or unrealistic to achieve or reach (e.g. stars). Therefore, ideal things, while desirable, may be unrealistic. The metaphor REAL IS DOWN is just the opposite of UNREALISTIC IS UP, that is, things that are located in lower position are more accessible, or “real” to our perception. The association between central and important is so conventionalized that “important” has become a lexical meaning of

“central”. It may arise from our biological feature that the most vital organs, tissue or substances (heart, liver, marrow) are located near the center of our bodies (Goatly, 2007). Finally, the meanings of the concepts of foreground/background can be explained in relation to the notion of “depth”, which is “the distance between the viewers’ eyes and any point in the visual field” of realization (Messaris, 1994, p. 51). Thus, foreground is read as near to the viewer and background as remote to the viewer. The human biological feature of vision results in the different visual impact of far and near objects: we are able to notice what is foregrounded first (most likely for reasons of survival) and take it as the most important or prominent element than that which appears in the background.

Through these experiential bases, it can be argued that these metaphors do exist and are conventionalized in our everyday ordinary conceptual system. However, these conventional or default interpretations of camera positioning and composition may be overridden by other factors in specific contexts. For example, Dick (2005) points out that sometimes film scripts require a high or low angle shot for the sake of consistency rather than for symbolism purposes. From the cognitive perspective, this is because certain semiotic choices (e.g. low angle) are not motivated by the default experiential basis, or it is motivated by both the default experiential basis and other factors (e.g. inter-textual and discursive consistency). In such cases, the context of situation may point to a specific interpretation. In this way, the reformulation of visual grammar as metaphor system not only makes it possible to validate the association between semiotic values and visual resources by providing experiential basis to it, but it also helps in expounding on the conditions where the association does not seem to exist.

4. The Systemic Functional Modeling of Visual Metaphor

Recently, the study of visual realization of metaphor has attracted much attention (Forceville, 1996; Forceville & Urios-Aparisi, 2009; El Refaie, 2003). However, as has been pointed out in Section 2, cognitive theorists have not yet designed a model to systematically account for the visual mechanisms used for representing metaphors. Thus, the aim of this section is to provide a systematic account of the visual realization of metaphor by relating it to Kress and van Leeuwen's (2006) visual grammar framework. The systemic functional model describes visual images in a more thorough and systematic way than cognitive theorists do. This comparison is claimed based on two factors: (i) at the semantic level, visual images are seen as metafunctional constructs (see Figure 1, above); (ii) at the lexico-grammar level, the visual resources for realizing metafunctions are modeled as interrelated systems of choice. These two factors allow a more comprehensive and holistic understanding of the object of study.

Forceville (1996) distinguishes three kinds of pictorial metaphors: MP1 (one metaphorical term is present), MP2 (two metaphorical terms are present and integrated) and pictorial simile (two metaphorical terms are juxtaposed). From a systemic functional perspective, Forceville's (1996) three types of pictorial metaphor are seen to be based on the systemic choices of spatial relations between the "metaphorical subject" (typically the target domain, that is, the primary subject) and the "pictorial context" as illustrated next in Figure 7.

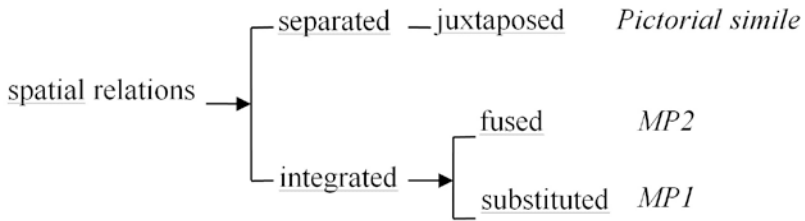


Figure 7 Forceville's (1996) typology of pictorial metaphor

However, in the systemic functional model, visual images involve not just spatial relations, but also representational and interactive resources. In the present framework, the three metafunctions all provide resources for realizing metaphors. We shall, then, briefly discuss how interactive and compositional resources realize metaphors. The visual metaphors discussed in Forceville (1996) are mostly novel metaphors, which are for decorative purposes in advertisements, while the more conventional metaphors are not included. From a systemic functional perspective, it is found that the meaning of interactive and compositional resources, namely, camera positions and visual spaces, are acquired through conventionalized metaphorical mapping, as elaborated in Section 3 above. Therefore, at the same time they explain the meaning of camera positioning and spatial location, the metaphor systems in Figure 5 and Figure 6 are also types of visual metaphors. However, the “targets” of such metaphors are not present in images, but derived from *correlations* of our basic shared experience of the world (Lakoff & Johnson, 1980). The interpretation of such metaphors, therefore, does not depend on immediate context, but on physical and cultural experiences that are common to human beings or specific cultural communities, as has been previously discussed in Section 3.

Aside from the interactive and compositional resources, the representational structure is also fundamental for interpreting the visual mechanism of metaphor construal. In what follows, a systematic functional theorization of Forceville's (1996) categorization of novel visual metaphors (see Figure 7) based on Feng (2011a) will be provided. In such a context, we see Forceville's "metaphorical subject" and "pictorial context" as belonging to one unified grammatical unit in the representational meaning structure. We draw on Kress and van Leeuwen's (2006) structure of representational meaning to model the relation between the "source domain" and the "target domain" in a more precise and systematic way.

Kress and van Leeuwen (2006) identify two types of structure in terms of representation: narrative structure and conceptual structure. The distinction between these two structures refers to the ways through which the participants of an image are related to one another. In other words, the distinction is based on either the "unfolding of actions and events, processes of change", or based on the "generalized, stable and timeless essence" of what is going on in the visual image. There are five types of process within the category of *narrative representation*. The first four types, actional, reactional, verbal and mental processes involve a distinct agent (Actor, Senser, Carrier, Sayer, etc.) and are categorized as agentive processes. An actional process represents the action of an agent. Reactional process is typically formed by the eyeline direction of a represented participant. Verbal and mental processes are constructed by dialogue balloons and thought bubbles respectively. Finally, the non-agentive process type of conversion of the *narrative structure* involves a change of a state of affairs of the represented participant in the order of things within the image.

As for the *conceptual structure*, the participants are related through taxonomic relations, part-whole relation or symbolic relations, termed classificational process, analytical process and symbolic process, respectively. A classificational process relates represented participants to each other in terms of taxonomy, with these participants as the subordinates of another participant, namely the superordinate participants. The taxonomy can be overt or covert, depending on whether the superordinate is represented in the image or not. In analytical process, participants are related based on a part-whole structure. The two types of represented participants involved in an analytical process are Carrier (i.e. the whole), and Possessive Attributes (i.e. the parts that constitute the whole). A symbolic process defines the meaning or identity of a represented participant. The process types discussed above can be visualized and summarized in the following figure.

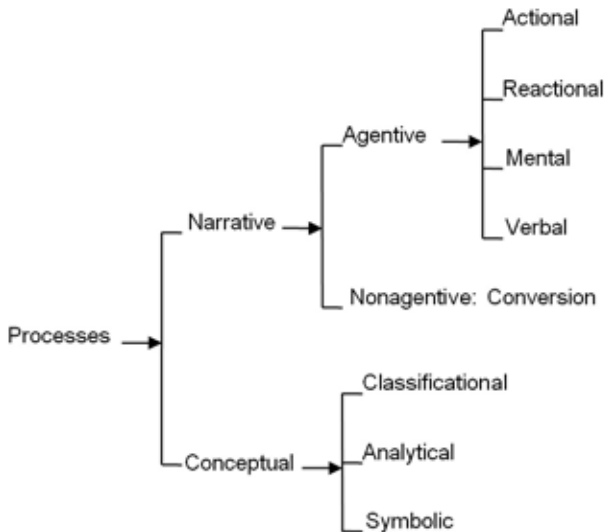


Figure 8 Process types in visual images (Kress & van Leeuwen, 2006)

Bearing this classification in mind, we suggest that novel visual metaphors are mainly constructed by anomaly, or unconventionality, of visual elements in the representational structure, in a similar manner to the colligational interpretation of metaphor in language (Goatly, 1997). Anomaly has different meanings in different process structures since metaphorical terms are related to it in different ways. In narrative structure, the target domain relates to other elements through actional process, verbal process, mental process, and so on; in conceptual structure, the target domain relates to other elements through relational processes in the form of taxonomic relations (classification process), part-whole relations (analytical process) or identifying relations (symbolic process). In the following exemplifications, we shall mainly examine the realization of metaphor in the actional process, the classificational process and the analytical process in advertisement campaigns to show how the novel visual metaphor is construed by anomaly in the representational structure.

In actional process, the conventional participant (i.e. actor or goal) or circumstantial element (e.g. medium) associated with an action (termed B) is substituted by an unconventional one (termed A). As a result, the metaphor A IS B is formed. In Forceville's (1994, p. 10) example, a person is killing himself by pointing a gas nozzle to his head. The metaphor GAS NOZZLE IS GUN is constructed because the gas nozzle adopts the role of a gun. For verbal and mental processes, since the agent of these two process types has to be endowed with consciousness, that is, human being, therefore, if non-human agents are performing these process types, then, visual personification is formed.

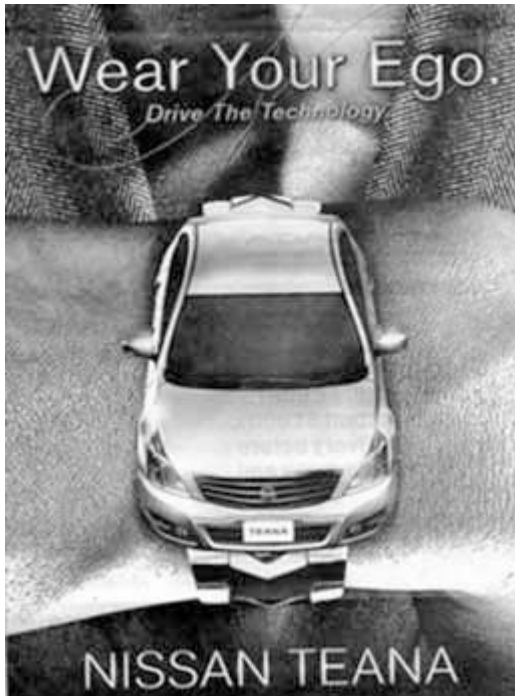


Plate 1 Nissan Teana, from *The Straits Times*, 4th December, 2008, C17

In the car advertisement in Plate 1, the car is worn on the wrist like a watch. Apparently, the car takes the place of a watch, which results in colligational anomaly. By taking the place of a watch, the car adopts the attributes of watch. That is, the attributes of the watch are projected onto the car, constituting the metaphor CAR IS WATCH. The process of substitution within the narrative structure is illustrated in Table 1 as follows:

Grammatical role	Actor	Process	Goal	Circumstance
Unconventional element (current text)	man	wears	car	on wrist
Conventional element (background knowledge)	man	wears	watch	on wrist

Table 1 Participant substitution in actional process

Abstract *processes* can also be visualized as concrete processes. The image in Plate 2 below can be described as “ocean water is being poured into drinking glass” in which the ocean is recognized by the blue water (in the original picture) together with the fish and seaweed immersed in the ocean water. The target, that is the linguistic text superimposed on the image, consists of the complex process of desalination. Therefore, it is possible to read the metaphor DESALINATION IS POURING OCEAN WATER *into drinking glass*, through which desalination is understood.



Plate 2 From the front cover of *The Economist*, June, 7th, 2008

The classificational process constructs metaphor mainly in two ways. First, entity A is an unconventional member of a category whose conventional member is entity B. As a result, A borrows the salient features of B and the metaphor A IS B is formed. In Teng's (2009, p. 198) example, where an American newspaper is put among horror books on the bookshelf labeled 'horror', the resultant metaphor AMERICAN NEWS IS LIKE HORROR NOVELS is an example of this kind of case construction. Second, two entities may be put together unconventionally to form a covert category (Kress & van Leeuwen, 2006). The formation of covert category requires a crucial visual feature — that is, symmetry in composition, such as equality in size, framing and arrangement (Kress & van Leeuwen, 2006). This process is similar to the visual simile in Forceville's (1996) categorization. However, the source and target cannot be structurally determined in this case because they are represented on the premise that these two elements have to be equal. In this sense, we have to draw upon discursive purposes. The advertisement in Plate 3 (presented next) is a good case to illustrate such point. The minivans are juxtaposed with weight-lifting champions. They form a covert category by being identical in number and arrangement. Since it is an advertisement for the minivan, the minivan is the target and the metaphor thus formed is MINIVANS ARE WEIGHT-LIFTING CHAMPIONS. The salient feature of the athletes, that is, strength, is mapped onto the minivans.



Plate 3 Wuling minivan, from *Qilu Evening Paper*, July, 19th, 2008

Anomaly in analytical process occurs when there is an unconventional part in the whole composition. This is typically realized as the unconventional part A taking the place of the conventional part B. As a result, A inherits the salient features of B and forms the metaphor A IS B. The well-known example from Forceville (1996, p. 110) in Plate 4 below, which shows a man's torso with a suit but the tie is substituted by a shoe, is a good case in exemplifying this anomaly. This image is commonly viewed as a formal attire that typically includes a suit with a tie. By taking the place of the tie, the shoe borrows the salient features of the tie and the metaphor SHOE IS TIE is formed.

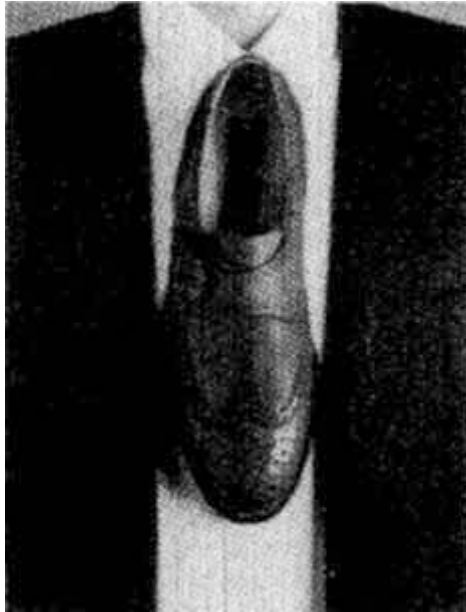


Plate 4 Shoe advertisement from Forceville (1996, p. 110)

To summarize, we have looked at the visual mechanisms for the construction of metaphor in the representational structure. It is observed that the source and target of novel visual metaphors can mostly be identified through examining the way they are represented (i.e., anomaly in different process structures). However, since metaphors are not constructed by decontextualized visual components, representational resources alone may not be able to specify the source and target focal points. Moreover, the cues from process of construction may contradict with the more explicit contextual cues of interpretation. Such awareness of context is in accordance with the systemic functional approach, whereby representational anomalies are seen as *resources* for metaphor realization, rather than rigid semiotic *codes*. In this regard, the systemic

functional modeling of context (e.g. Halliday and Matthiessen, 2004; Martin, 1992) also provides a more systematic framework for the contextual interpretation of visual metaphor than Forceville (1996)'s proposal. This comparison needs further exploration in future attempts to integrate systemic functional and cognitive approaches to the reading and interpretation of metaphor in images.

Meanwhile, as explained in Section 3, the relation between camera positioning and interactive meaning, as well as that between spatial orientation and information value, is seen as metaphorical mapping between the source domain and the target domain. From the perspective of conceptual metaphor theory, these metaphors are visual realizations of “conventional metaphors” which are based on *correlations* in our bodily experience (Lakoff & Johnson, 1980, p. 139). The systematic categorization of metaphors in interactive and compositional resources complements cognitive studies of visual metaphor mostly focused on *what* is in the image, instead of *how* the image is represented (e.g. Forceville, 1996; El Refaie, 2003). Therefore, our framework of visual metaphor includes both “novel” metaphors constructed by visual anomaly and conventional metaphors that are implicit in camera positioning and spatial orientation. That is, in systemic functional terms, we explored metaphors realized in representational resources, interactive resources and compositional resources.

5. Conclusion

The present paper provides a synthesized framework for integrating the systemic functional and cognitive metaphorical approaches to the analysis of visual images. It argues that the epistemological status of Kress and van Leeuwen's (2006) association

between camera angle/spatial orientation and semiotic value can be established by viewing it as a metaphorical mapping of image realization. Meanwhile, Kress and van Leeuwen's (2006) systemic functional visual grammar provides a comprehensive modeling of the visual realization of both novel and conventional metaphors. The integration of these two analytical paradigms is significant for further understanding and explaining visual semiosis. On the one hand, the cognitive support of the descriptive visual grammar provides more solid theoretical basis for the analysis of visual images; on the other, the systematic account of how different types of metaphor are realized in images sheds further light on the nature and working mechanism of visual metaphors.

Note

1. The dimension of 'modality' which refers mainly to the realness of the image is out of the scope of the present study and is, therefore, not included in the discussion here.
2. For the purpose of the present study, we shall call the cross-domain mapping a "conceptual metaphor" in accordance with CMT, but we are aware and stand by Halliday and Matthiessen's (1999) position that metaphor is a linguistic (semantic) phenomenon.

References

- Carroll, N. (1996). A note on film metaphor. *Journal of Pragmatics*, 26(6), 809-22.
- Dick, B. F. (2005). *Anatomy of film*. Boston, Mass: Bedford/St. Martins.
- El Refaie, E. (2003). Understanding visual metaphor. *Visual Communication*, 2(1), 75-95.
- Feng, D. (2011a). The construction and categorization of multimodal metaphor: A systemic functional approach. *Foreign Languages Research*, 1, 24-29.

- Feng, D. (2011b). Visual space and ideology: A critical cognitive analysis of spatial orientations in advertising. In K. L. O'Halloran & B. Smith. (Eds.), *Multimodal studies: Exploring issues and domains* (pp. 55-75). London: Routledge.
- Forceville, C. (1994). Pictorial metaphor in advertisements. *Metaphor and Symbolic Activity*, 9(1), 1-29.
- Forceville, C. (1996). *Pictorial metaphor in advertising*. London: Routledge.
- Forceville, C., & Urios-Aparisi, E. (Eds.). (2009). *Multimodal metaphor*. Berlin: Mouton de Gruyter.
- Goatly, A. (1997). *The language of metaphors*. London: Routledge.
- Goatly, A. (2007). *Washing the brain: Metaphor and hidden ideology*. Amsterdam: John Benjamin.
- Hall, E. (1969). *The hidden dimension*. London: Bodley Head.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (1999). *Construing experience through meaning: A language based approach to cognition*. London: Cassell.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2004). *An introduction to functional grammar* (2nd ed.). London: Arnold.
- Kress, G., & T. van Leeuwen. (1998). Front pages: (The critical) analysis of newspaper layout. In A. Bell & P. Garrett (Eds.), *Approaches to media discourse* (pp. 186-219). Oxford: Blackwell.
- Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). London. Routledge.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Martin, J. R. (1992). *English text: System and structure*. Amsterdam: John Benjamins.
- Messaris, P. (1994). *Visual literacy: Image, mind and reality*. Boulder: Westview Press.
- O'Halloran, K. L. (2005). *Mathematical discourse: Language, symbolism and visual Images*. London: Continuum.

- O'Toole, M. (1994). *The language of displayed art*. London: Leicester University Press.
- Rogers, Henry. (2005). *Writing systems: A linguistic approach*. Oxford: Blackwell.
- Teng, N. Y. Image alignment in multimodal metaphor. In C. Forceville & E. Urios-Aparisi (Eds.). (2009). *Multimodal metaphor*. Berlin: Mouton de Gruyter.
- Zhu, Y. (2007). The theoretical foundation and methodology of multimodal discourse analysis. *Foreign Language Research*, 5, 82-86.

[Received in 20/05/2013. Approved in 27/06/2013]