Metonymy and Visual Representation: Towards A Social Semiotic Framework of Visual Metonymy

Abstract: This study proposes that metonymy is fundamental to visual meaning making and develops a social semiotic framework to elucidate how conceptual metonymies are realized in both static and moving images. While we all accept that visual images are iconic, this study demonstrates systematically that they are also indexical (i.e. metonymic), in terms of their representation of both objects/events and abstract concepts. Based on the social semiotic visual grammar of Kress and van Leeuwen (2006), systems of metonymy in actional, reactional, classificational and analytical processes are developed to map out the types of metonymies in visual representation. The metonymy systems bring a wide array of resources under a coherent framework for analysts to scrutinize the choices of representation in visual media such as comics, film and TV commercial. This study develops current theories of multimodal metaphor and metonymy on the one hand, and provides new insights into the process of visual meaning making on the other.

Key words: visual metonymy; conceptual metaphor theory; social semiotics; visual grammar; multimodality

INTRODUCTION

The central argument of conceptual metaphor theory (Lakoff and Johnson, 1980) is that metaphor is a conceptual phenomenon, realized not simply in language but in other communication modes as well, such as visual image and gesture (e.g. El Refaie, 2003; Forceville, 1996; Goatly, 2007; Kövecses, 2010). Recently, the study of the non-linguistic realizations of metaphor has attracted much attention, notably with the publication of *Multimodal Metaphor* (Forceville and Urios-Aparisi, 2009) and the special issue on "multimodality and cognitive linguistics" in the journal *Review of Cognitive Linguistics* (Pinar, 2013). While cognitive linguists have been interested primarily in conceptual metaphors, in the past two decades conceptual metonymy has attracted an increasing attention as a fundamental device of human cognition (e.g. Barcelona, 2000; Panther and Radden, 1999; Panther and Thornburg, 2004). In this context, investigating nonverbal metonymy is a logical next step for (cognitive) multimodal

researchers (Forceville, 2009: 56). Some cognitive linguists have started to explore the nonverbal manifestations of metonymy. For example, Forceville (2009) studied multimodal metonymy in static and moving images through the analysis of several of advertising texts and feature films; Moya (2013) analyzed part-whole metonymies in children's picture books; Yu (2009) and Urios-Aparisi (2009) studied the multimodal interaction of metaphor and metonymy in educational TV advertisements and TV commercials, respectively. However, compared with the relatively well-developed theoretical models of visual metaphor (e.g. El Refaie, 2003; Forceville, 1996; Feng and O'Halloran, 2013a), so far there is no systematic description of the visual mechanisms that realize metonymies.

From the perspective of Peircean semiotics, metonymic relations are seen as derived from the semiotic principle of indexicality[1], including relations such as cause and effect, part and whole, container and contained, and so on (Norrick, 1981). Relating to the Peirce's typology of signs, the images discussed in this paper are primarily iconic (i.e. they resemble reality in a straightforward way), which include print advertisements, moving images, comics and so on, but not symbolic images such as mathematical symbols, diagrams, etc. However, besides iconicity, images are also indexical in at least two senses: first, visual images are never exact replications of reality, but can only be partial representations of it (Feng and O'Halloran, 2012); second, visual representations of abstract concepts which are invisible (e.g. emotions) can only be through depictions of visible objects (e.g. symptoms of emotions) related to them (see for example the discussion of emotion representation in Forceville, 2005; Feng and O'Halloran, 2012). The concern of this paper is precisely on the theorization of these two types of indexicality.

Feng and O'Halloran (2013a) propose a framework to elucidate the visual realization of metaphors based on Kress and van Leeuwen's (2006) visual grammar. They argue that the theorization of visual metaphors has to be based on a systematic description of meaning construction mechanisms in visual images. Adopting the same approach, this paper provides a systematic account of the visual realization of conceptual metonymy. In the following section, I will first introduce the cognitive theory of metonymy and the social semiotic model of visual images, and then put forward a framework to integrate these two approaches. After that, the

---

[1]In Peircean semiotics, indexical signs make meaning through causal-continuity relations with the objects, iconic signs through resemblance and symbolic signs through conventions.

social semiotic framework of visual metonymy is elaborated in terms of the metonymic nature of representational meaning and attitudinal meaning. The framework is then applied to the analysis of meaning construction in a TV commercial. Finally, the study concludes that the social semiotic framework can provide a comprehensive account of the visual construction of metonymy, and that visual metonymy provides a new perspective for understanding the nature of visual semiosis.

CONCEPTUAL METONYMY AND SOCIAL SEMIOTIC VISUAL GRAMMAR

In cognitive linguistics, metonymy is regarded as a cognitive process that evokes a conceptual frame, rather than merely a matter of the substitution of linguistic expressions (Panther and Radden, 1999: 9). Like metaphor, metonymy involves the understanding of one thing in terms of the other. Different from metaphors which involve the mappings between two conceptual domains, in metonymies, "the mapping or connection between two things is within the same domain" (Gibbs, 1994: 322). Simply put, metonymy is "using one entity to refer to another that is related to it" (Lakoff and Johnson, 1980: 36)[2]. Common metonymic relations include part-for-whole (e.g. we need more *hands* here), place-for-institution (e.g. the *White House* declared war), artist-for-artwork (e.g. she likes reading *Shakespeare*), and so on.

Many theoretical models have been proposed to explain conceptual metonymy (e.g. Gibbs, 1994; Radden and Kövecses, 1999; Lakoff, 1987; Panther and Thornburg; 2004; Warren, 2002). Among them, Radden and Kövecses (1999) have proposed a comprehensive framework to elucidate the ontological realms of metonymic relationships and the types of mappings[3]. In terms of ontological realms, they distinguish the world of "concepts", the world of "forms", in particular, the form of language, and the world of "things" and "events", corresponding to Ogden and Richards' (1923) trichotomy of thought, symbol and referent. Based on this trichotomy, Radden and Kövecses (1999) distinguish three types of metonymies: sign metonymy, reference metonymy, and concept metonymy. In sign metonymy, a linguistic (or nonverbal) form is used to stand for a concept. They argue that the very nature of language is based on this metonymic

---

[2]The boundary between metaphor and metonymy is fuzzy and there is no agreement among cognitive linguists. This paper follows the definition of Gibbs (1994) and adopts the categorization of Radden and Kövecses (1999). That is, I consider a mapping a metonymy if it is based on part-whole relations or part-part relations. The ambiguity will be explained where relevant in the following discussions.

[3]Following cognitive linguists such as Lakoff and Johnson (1980) and Radden and Kövecses (1999), "mapping" is used to refer to the connection or link between the source domain and the target domain.

principle, which Lakoff and Turner (1989: 108) describe as WORDS FOR THE CONCEPTS THEY EXPRESS. In reference metonymy, the typical case is that a sign (i.e., form-concept unit) stands for a thing or event it refers to (e.g. the word "cow" for a real cow). Concept metonymies involve a shift from Concept A to Concept B, which are related to each other in some specific way within the same conceptual domain (or Idealized Cognitive Model, ICM) (e.g. bus for driver, White House for US government). Concept metonymy is the type that is usually taken as metonymy and is the focus of most cognitive studies. Radden and Kövecses (1999) continue to elucidate the conceptual relationships within an ICM which may give rise to metonymy, that is, metonymy-producing relationships. They propose two general types of relations: those between the whole ICM and its part(s), and those between parts of an ICM. The typical case of the former is when part of an object/event stands for the whole object/event, or vice versa (e.g. America for The United States of America). In an example provided by Gibbs (1999: 66), the answer of "I waved down a taxi" to the question "How did you get to the airport?" stands for the scenario that "I got to the airport by hailing a taxi, having it stop and pick me up, and then having it take me to the airport." In part-part configurations, one part of an ICM stands for the other part. Typical relations include effect for cause (e.g. slow road for bad road condition), agent for action (e.g. to author a book), producer for product (e.g. I bought a Ford), and so on (see Radden and Kövecses, 1999).

Although cognitive theorists agree that metonymy is a conceptual phenomenon and can be realized in both language and other semiotic resources, so far there has been no systematic discussion of the visual manifestation of metonymy, like Radden and Kövecses (1999) have provided for linguistic metonymy or Forceville (1996) for visual metaphors. In this paper, I propose to model the visual realization of metonymies drawing on the social semiotic theory of Halliday (e.g. Halliday 1978; Halliday and Matthiessen, 2004) and the visual grammar inspired by it (Kress and van Leeuwen, 2006). According to Kress and van Leeuwen (2006), visual images, like language, fulfill three metafunctions, namely, 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). Similar to Lakoff and Turner's (1989: 108) metonymy of WORDS FOR THE CONCEPTS THEY EXPRESS, we can extend the resource of words to visual resources (e.g. characters' appearances, facial expressions, costumes, etc.) as well and get the

general *sign metonymy* of VISUAL RESOURCES FOR MEANING (i.e. metafunctions in social semiotic terms). This agrees with the proposal that "since we have no other means of expressing and communicating our concepts than by using forms, language as well as other communication systems are of necessity metonymic" (Radden and Kövecses, 1999: 24). However, it should be noted that for some of the visual metonymies identified in this paper, the mappings between the source and target are not as conventionalized and strong as the lexicalized ones in language, due to the nature of visual semiosis (i.e. visual images lack a clear vocabulary system and many interpretations rely on context). It follows that metonymy identification in visual images operates at the level of discourse semantics and is context dependent (cf. Forceville, 1996, 2009). From this perspective, "a metonym consists of a source concept/structure, which via a cue in a communicative mode (language, visuals, music, sound, gesture, etc.) allows the metonym's addressee to infer the target concept/structure" (Forceville, 2009: 58). The "addressee" makes inference from the visual cues drawing upon his/her cultural knowledge as well as the immediate context of communication.

In terms of the relation between visual images and the reality, in social semiotic terms, human experience is construed as different process types (cf. Halliday and Matthiessen, 2004). Kress and van Leeuwen (2006) identify two types of process in terms of visual representation: narrative processes and conceptual processes. The distinction between them lies in the ways in which the image participants are related to each other, that is, whether it is based on the "unfolding of actions and events, processes of change", or based on their "generalized, stable and timeless essence" (Kress and van Leeuwen, 2006: 59). There are four main types of processes within the category of narrative representation: actional, reactional, verbal, and mental processes. An actional process depicts the action of a participant (e.g. running, hugging, and punching). A reactional process depicts a participant's reactions, typically formed by facial expressions (e.g. smiling, crying, and frowning). Verbal and mental processes are constructed by dialogue balloons and thought bubbles respectively (e.g. in comics). In conceptual representation, the participants are related through taxonomic relations, part-whole relations, or symbolic relations, termed classificational process, analytical process and symbolic process, respectively. A classificational process relates the represented participants to each other in terms of taxonomy, with these participants as the subordinates of another participant, which is their superordinate (similar to hyponymy relation among linguistic concepts). In analytical processes, participants

are related based on a part-whole structure (e.g. appearance or clothing as part of a person). 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). Symbolic processes define the meaning or identity of a represented participant through certain cultural associations (e.g. cross stands for church or Christianity).

These processes may construct two types of meanings. First, they record and reconstruct reality, which is the representational meaning. Unlike words which are symbolic, visual resources are primarily iconic and necessarily partial, as it is impossible to reproduce reality (Feng and O'Halloran, 2012; Kress, 2010). More importantly, visual sign makers always choose from a range of available choices motivated by their interests (Kress, 2010: 67). Therefore, the process of visual meaning construction is metonymic, similar to reference metonymy in the model of Radden and Kövecses (1999). This type of metonymy is elaborated in Section 3.

Second, they can indirectly construct attitudinal/evaluative meanings (Martin and White, 2005; Feng and O'Halloran, 2013b). According to the Appraisal system of Martin and White (2005), attitude includes three subcategories, namely, emotional responses (Affect), values by which human behaviors are socially assessed (Judgment), and values which address the aesthetic qualities of objects and entities (Appreciation). Attitudinal meanings can be explicitly constructed through attitudinal lexis (e.g. happy, kind, valuable), but can also be indirectly constructed through ideational events (i.e. things/events which elicit the attitude) (cf. Feng and O'Halloran, 2013b). For example, "happiness" can be inferred from "I got the job", and "kindness" can be inferred from the utterance "she donated all her money to the poor". As attitude is an abstract concept, which cannot be inscribed in visual images, visual representation of attitude is of necessity metonymic (cf. Forceville, 2009). If we consider the attitude schema in Figure1 as an ICM (cf. Feng, 2012: 87), then attitude (e.g. happiness) can be metonymically represented through the visual depiction of eliciting condition (e.g. meeting an old friend) or reaction (e.g. smiling and hugging). This type of metonymy will be elaborated in Section 4.

Insert Figure 1 here

The three types of metonymies based on Ogden and Richards' (1923) trichotomy, which is illustrated in Figure 2, provide the general framework of this paper. The second type (marked as ② in Figure 2), sign metonymy, is mentioned previously as a general principle of meaning

making. Focusing on representational meaning, in the next two sections, I will elucidate the first type (marked ① in Figure 2), reference metonymy, in terms of the partial representation of reality, and the third type (marked ③ in Figure 2), concept metonymy, in terms of how representational meaning invokes attitudinal meaning. Then in Section 5, I will demonstrate the application of the theory through analyzing meaning construction in a TV commercial. In this way, this paper elucidates the metonymic relations that are fundamental in visual meaning making, that is, how visual images, as the source domain of the mappings, construe reality and abstract concepts.

Insert Figure 2 here

METONYMY AND THE PARTIALITY OF VISUAL REPRESENTATION

In this section, the partial nature of visual meaning making is elucidated. Practically, it is impossible to reproduce every aspect of the three-dimensional reality in two-dimensional static or moving images. For events which unfold in time, static images can only capture snapshots, and moving images are often only fragments of the whole process (Painter, Martin and Unsworth, 2013: 58). This process of reduction and abstraction is explained as metonymy from the perspective of reception and production in this section. After that, a framework is proposed to model different types of partiality in the process of visual meaning making.

In terms of reception (i.e. how viewers make sense of visual images), cognitive theorists explain that our knowledge is stored in memory in the form of scripts or schemata (e.g. Bartlett, 1932; Schank and Abelson, 1977). Scripts consist of well-learned scenarios describing structured situations in everyday life. The knowledge in long-term memory of a coherent, everyday series of events can be metonymically referred to by the mere mention of one salient subpart of these events (Gibbs, 1999: 69). The ability to infer a whole script from the mere mention of a part makes it possible for us to make sense of seemingly anomalous and disconnected statements in multimodal texts. For complex multimodal film texts, the cognitive film theory that is currently dominant in the field is based on this very metonymic principle. Cognitivists' concern is with how viewers make sense of the inherently incomplete form of discourse (i.e. combinations of shots in film) by using their capacity for inference generation. A wide range of studies have been published which describe the functions and stylistic conventions of filmic devices cuing spectators to various dimensions of film comprehension (e.g. Bordwell, 1989; Carroll, 1996).

From this perspective, Bordwell (1985) has developed a theory which includes the two-tier construct of *fabula* and *syuzhet* as formulated by the Russian formalists: "The imaginary construct we create, progressively and retrospectively, was termed by formalists the fabula (sometimes translated as 'story') …The syuzhet (usually translated as 'plot') is the actual arrangement and presentation of the fabula in the film. It is a more abstract construct, that is, the patterning of a story as a blow-by-blow recounting of the film could render it" (Bordwell, 1985: xii). From the perspective of metonymy, this proposal can be reformulated as SYUZHET FOR FABULA, in which we infer the coherent fabula based on the incomplete syuzhet.

In terms of production (i.e. how visual images are created), Norrick (1981) attends to the part-whole principle in nonverbal sign systems over three decades ago. Norrick (1981: 53-54) argues that "central to traditional painting and sculpture stood the idea that an entire event could be evoked by portraying a single representative moment of the whole. Nor is even the total reality of the single moment portrayed, but only pertinent aspects of it, and from some particular perspective, which necessarily concentrates the attention on a single part of the whole scene". In more recent social semiotic terms, Kress (2010: 70) explains that visual representation is always partial, motivated by the sign maker's interest[4]. As it is impossible to represent everything, we constantly make choices according to our interest in the process of sign making. Therefore, in social semiotics, "arbitrariness is replaced by motivation in all instances of meaning making, for any kind of sign" (Kress 2010: 67). For example, in analyzing a young child's representation of car using two circles, Kress (2010: 70) proposes that the process involves two steps: circles are like wheels and wheels (not other parts) stand for a car[5]. Motivation makes the partiality more than just a choice of form, but also a choice of meaning. As Kress (2010: 71) points out, "partiality of interest shapes the signified at the moment of the making of the sign". This is consistent with cognitive linguists' proposal that "the essence of metonymy is highlighting" (Warren, 2002: 123). Developing the principles set out by Norrick (1981) and Kress (2010), in the rest of this section, I will provide a framework to elucidate the types of partiality in visual sign making.

---

[4]Different from Painter, Martin and Unsworth's (2013: 61) distinction between "complete" and "metonymic" representations in picture books, in the present definition, all representations are of necessity metonymic.

[5]Kress (2010) considers this as a "double metaphorical process", focusing on the analogy in the process of meaning making. As the process involves part-whole relation, it is considered as metonymic in this paper.

Partiality is realized in different ways in different processes in the representational structure of visual images. For narrative processes, the focus is on actional processes and reactional processes. Verbal and mental processes are not so relevant here for the reason that are realized by speech/thought bubbles which usually projects language. By definition, actional and reactional processes involve actions which unfold in time (actional process shall be used as a cover terms in this section). Continuous videotaping reproduces the most details of an event (see Figure 4). In film, the representation of the whole process in one shot is rare; instead, it involves different configurations of shots, called "syntagmas" (Metz, 1974). For example, in the representation of a beheading, as it is impossible to represent the whole process of cutting off one's head, we can only see the raising of the broadsword, and then a fake head on the ground. In *Gladiator* (Scott, 2000), there is a scene in which Commodus is smothering his father to death. Instead of reproducing the event in one shot, the director uses rapidly alternating shots to feature Commodus' facial expression and his father's struggling hands, as illustrated in Figure 3. Aside from selecting salient moments from the event, the representation also involves framing (e.g. close-up shots or panoramic shots), which will be discussed in the following.

Insert Figure 3 here

In static images, visual representations provide a snapshot of movement to represent the whole re/action. This can be illustrated using the representation of gesture, touch, and facial expression. Kendon (2004) explains that a prototypical gesture passes through three phases, namely, the preparation, the stroke, and the retraction. Similar components can be identified in touch and facial expression. For example, a slapping behavior includes raising the hand, the slapping, and taking the hand back. Images can only use a snapshot of actual contact or the preparatory stage of the behavior. For example, in emoticons in online chatting, the kiss is signified by pouted mouth and the hug is signified by outstretched arms, both of which are preparatory to the actual behavior (cf. Feng and O'Halloran, 2012). Therefore, in moving or static images, we can get the metonymy of MOMENTS OF AN ACTION FOR THE WHOLE ACTION. As illustrated in Figure 4, the metonymy is manifested as a degree of partiality from continuous videotaping to snapshots, depending on the amount of moments selected (the slanted square bracket signifies the cline of change).

In conceptual structures, metonymic relations are based on selections of members from a category (classificational process) or parts from an entity (analytical process)[6]. Thus, we get the two general metonymies of MEMBER OF A CATEGORY FOR THE CATEGORY and PART OF AN ENTITY FOR THE ENTITY. In a classificational process, one or more members of a category are used to stand for the category as a whole, which is termed "generic representation" in van Leeuwen (2008). In talking about racist pictures, van Leeuwen (2008: 143) argues that "at first sight, it might seem that images can only show specific people. Yet, there is a difference between concentrating the depiction on what makes a person unique and concentrating the depiction on what makes a person into a certain social type". From the perspective of metonymy, the characters represented (e.g. one Muslim woman) are used to stand for the whole category (e.g. all Muslim women). As it is impossible to represent all members of a category in visual images, all visual representations of categories are necessarily metonymic. In textbooks, for example, in explaining the species of koala, an image of a koala is usually provided. In a car advertisement, the image does not refer to that very car, but metonymically stands for the brand or the model. This is discussed under the indexical principle of specific-generic relation in Norrick (1981:35), which states that "any specific instantiation of a class calls forth the whole class, and consequently serves as a motivated sign of it". Similarly, the human figures in advertisements are often generic representations. This point is aptly summarized by van Leeuwen (2008: 143): "when people are photographed as desirable models of current styles of beauty and attractiveness, their individuality can seem to disappear behind what categorizes them—behind the hairdo, the makeup, the dress, the status accessories". More importantly, the representations carry stereotypical judgments of the character, for example, the characters in Calvin Klein advertisements may stand for "cool and sexy" people. Such metonymic reference to more abstract stereotypical meaning will be discussed in the next section.

In analytical processes, two types of partiality can be identified: framing and abstraction. In terms of framing[7], in visual images there are a limited number of standard framings to represent something less than totally: extreme close-up, close-up, medium close-up, medium shot, medium long shot (Forceville, 2009: 63; Kress and van Leeuwen, 2006). Meanwhile, the camera can only

---

[6]Symbolic processes are also essentially metonymic, but they are Concept metonymies which involve complex conventionalized associations and are relevant here (see Forceville, 2013 for detailed discussion).

[7]Note that the term 'framing' is used here to refer to camera use, different from Kress and van Leeuwen's (2006) use which refers to the degree to which dividing lines are created to disconnect elements.

take one perspective: frontal, back, profile or oblique angles, any of which only provides a partial view of the object. Therefore, different choices of camera positioning result in different partial representations and often imply a change in salience or perspective. As Forceville (2009: 58) affirms, "the choice of metonymic source makes salient one or more aspects of the target that otherwise would not, or not as clearly, have been noticeable, and thereby makes accessible the target under a specific perspective". Moya (2013: 339) also notes that metonymies are frequently used to highlight some aspect of the message and attract the reader's attention to relevant parts of a multimodal ensemble. In Figure 3, the close-up shot of Commodus' facial expression is a partial representation of the character and highlights his anguish in killing his father (see Feng, 2012 for a detailed analysis).

The second aspect of partiality in analytical process is abstraction, which refers to the reduction of analytical features. Kress and van Leeuwen (2006) discuss this phenomenon under the notion of "modality", which, in a naturalistic coding orientation, refers to how close the representation is to reality. Borrowing the terms of Painter, Martin and Unsworth (2013), I distinguish between "naturalistic" representations and "minimalist" representations, depending on how much details are included (see Figure 4). For example, if we take a photograph of a smiling face as an ensemble of analytical features, the smiles in comics only make use of some of the essential features (e.g. lip corners drawing back and upwards, cheeks raising and so on). In the emoticon "☺", the representation is more abstract as the facial features are reduced to stylized lips. In this way, we get a scale of abstraction according to the features that are represented, the choice of which is dependent upon the sign maker's interest in the context of sign making (e.g. genre conventions) (Painter, Martin and Unsworth, 2013).

To summarize, in this section, it is proposed that visual representation is inherently partial. The partiality is explained both in terms of the interest of the sign maker and in terms of its cognitive processing. It is shown how this partiality is manifested in the representational structure of visual images, creating a systematic framework to model the different types of metonymic relations, as illustrated in Figure 4. Such a typology provides an explicit modeling of the visual manifestation of part-whole metonymies, and at the same time offers a new perspective in understanding the nature of visual meaning making and Kress and van Leeuwen's (2006) visual grammar. For example, camera use and modality (abstraction) are discussed by them separately from analytical processes, but it is shown here that they are intricately related. It

should be pointed out that all these types of partiality may occur simultaneously in an image (e.g. a snapshot of a person running with a close shot which stands for the category of athletes), so curly bracket is used in Figure 4.

Insert Figure 4 here

METONYMY AND ATTITUDINAL MEANING IN VISUAL IMAGES

As explained in Section 2, representational resources in visual images can invoke attitudinal meanings based on the attitude schema. The main focus here will be on human beings in images, so that resources such as actions and reactions can be discussed. In accordance with social semiotic terminology, they will be called "participants". There are two types of attitudinal meanings metonymically constructed through visual processes: (1) we infer the participant's emotions based on his/her reactions to the eliciting conditions; (2) we infer the participant's attributes (e.g. capacity, morality) based on his/her actions, analytical features (e.g. dress, accessories), and social identity (e.g. doctor, student). These correspond to the two categories of Affect and Judgment respectively in the Appraisal system of Martin and White (2005). The framework therefore also extends appraisal resources to nonlinguistic modes and provides a theoretical basis for understanding how visual resources signify attitude.

The emotions and attributes are metonymic as they are inferences made by viewers based on visual cues (as well as other contextual factors). The interpretation conforms to the definitions of both Forceville (2009) and Radden and Kövecses (1999) since the emotions/attributes and the visual cues belong to the same ICMs. In what follows, I elucidate how the emotions and attributes of visual participants are metonymically constructed through actional, reactional, classificational and analytical processes[8]. As previously noted, metonymy identification in visual images is context dependent and the audience makes inference from the visual cues by drawing upon his/her cultural knowledge as well as the immediate context of communication. Therefore, the attitudinal meanings discussed in this section are "invoked" from the audience, rather than unambiguously "inscribed" in the text (cf. Martin and White, 2005). While there is a high degree of sharedness in inferring attitude (e.g. we can generally recognize other people's emotions accurately; there are accepted social standards on what is right and wrong), I am not arguing that

---

[8]Same as above, verbal and mental processes are not included, as in visual images they are represented by speech and thought bubbles in the framework of Kress and van Leeuwen (2006) which usually projects language.

different actions invoke exactly the same attitude from all viewers. Instead, the aim of this section is to elucidate how the representational semiotic resources invoke attitudinal meanings, which further enables us to understand why certain choices of actions, clothing, and so on are made by image designers to elicit intended attitudes from viewers.

Constructing Participant Emotion through Representational Meaning

The main theoretical basis for elucidating the representation of emotion is cognitive appraisal theories which argue that emotion antecedents drive response patterning in terms of physiological reactions, motor expression, and action preparation (Frijda, 1986; Lazarus, 1991; Oatley and Johnson-Laird, 1987; Scherer and Ellgring, 2007). Therefore, cognitive theorists agree, with slight differences, that all emotions include antecedents, the interpretation and evaluation of antecedents, subjective feelings, physiological changes and behavioral reactions. In this study, a three-stage scenario involving Eliciting Condition, Feeling State, and Expression/Reaction is adopted (see Figure 1). This schematic representation significantly facilitates our recognition of emotion because one or a number of the components are able to activate our knowledge of a specific emotion (cf. Gibbs, 1999). As emotion is an abstract concept, in visual images it can only be represented metonymically by depicting the eliciting condition (the cause) or the emoter's behavioral reaction (the effect) (Feng and O'Halloran, 2012, 2013b). In the following discussion, I will elaborate the metonymic representations of emotive meanings through eliciting conditions and behavioral reactions.

The first metonymy, REACTION FOR EMOTION is common in both static and moving images. Forceville (2005) and Feng and O'Halloran (2012) both explain emotive behavior as indexes of emotion (causal relation). The signifying power of emotive behavior is premised on psychological research which reveals that the nonverbal expression of emotion can be recognized with a high degree of accuracy (e.g. Ekman and Friesen, 1975; Izard, 1971). In visual media, such as film, advertisements, and comics, behavioral reactions, especially facial expressions, are the primary resources for construing emotive meaning (see for example, the study of facial expression in film by Plantinga, 1999; in comics by Forceville, 2005 and Feng and O'Halloran, 2012).

The second metonymy is ELICITING CONDITION FOR EMOTION. As Ortony et al. (1988: 3) note, "it is apparent that writers can reliably produce in readers an awareness of a character's affective

states by characterizing a situation whose construal is assumed to give rise to them." The reason for this is that the appraisal of an eliciting condition is generally shared amongst members and groups of a society (Bless et al., 2004). In multimodal forms of representation such as film, filmmakers are able to predict, most of the time correctly, viewers' emotional reactions based on cultural knowledge. It is thus possible for filmmakers to "design" emotions to optimize engagement with viewers. Eliciting condition can be visually represented by various processes, for example, a person may be terrified by a skeleton (analytical features), disgusted by a despicable behavior (actional process), saddened by another person's cry (reactional process), or delighted by the kiss he/she receives (actional process). Feng and O'Halloran (2013b) have developed a social semiotic framework to elucidate this very phenomenon. In TV commercials, advertisers usually present a problem to elicit negative emotions from characters and viewers, and then present the product as a solution to elicit positive emotion (Feng and Wignell, 2011).

Constructing Participant Attribute through Actional Process

In the following three sections, the focus is shifted from viewers' recognition of visual participants' emotions to viewers' evaluation of their attributes through actional, analytical and classificational processes. In terms of actions, the standards according to which they are evaluated are shared among the members of a social group. According to van Dijk (1976: 291), action involves a conscious being bringing about some change (in his body, in an object, in a situation) with a given purpose, under a certain circumstance. That is, actions cannot be defined in pure behaviorist terms, but need to include the actor's *intention*. It is only in this way that the actor is held responsible for the action, and it is precisely the intention that is subject to value judgment. Therefore, our knowledge of an action is stored as a schema which includes intention and visible action. The intention or subjectivity may or may not be verbally inscribed, but it can normally be recognized as part of the "action schema".

  In visual representations such as film and comics, a character's actions are the main resource to construct his/her attributes, e.g. as hero or villain, relying on the metonymic relation ACTION FOR ATTRIBUTE. As Smith (1995: 121) points out, actions are an essential source of the traits which we assign to characters, as it is unusual (and uninteresting) to explicitly inscribe characters' attributes through narration. Certain actions, aside from contributing to the plot, are specifically designed to invoke viewers' judgments, for example, to reinforce viewers' allegiance

to the hero and antipathy to the villain (Smith, 1995). For example, in *Gladiator* (Scott, 2000), Commodus presses his father Marcus' head very hard against his chest while his father struggles painfully (see Figure 3). We recognize instantly that Commodus' intention is to smother his father to death. This intention is automatically judged as immoral in that context according to our shared social standards. The metonymic relation is further supported by Norrick's (1981: 65) indexical principle that a conventional property may be attributed to an individual based on present observations of its experiential manifestations (e.g. actions). Therefore, as with the elicitation of emotions, character actions automatically invoke judgments from viewers if certain social standards are activated. That is why various visual art forms can assign intended attributes to characters through designing their actions and predict viewers' reactions for persuasive or recreational purposes.

Constructing Participant Attribute through Analytical Process

Aside from actions, another important resource for constructing a character's attributes is through the analytical features. As explained in Section 2, analytical processes relate part-whole relations of participants. They involve two kinds of participants: the Carrier (the whole) and any number of Possessive Attributes (the parts – see Kress and van Leeuwen, 2006: 87). In Section 3 above, I discussed the notion of metonymic mapping when the part is used to stand for the whole. At a higher level of abstraction, the outer physical attributes (analytical features) of a person (or an entity) can *index* inner conceptual attributes based on the shared cultural knowledge or folk psychology (Feng, 2012: 181). This process borders on symbolic process discussed by Kress and van Leeuwen (2006), but I shall consider characters as Carrier, clothing as Possessive Attributes, and the characters' inner attributes as "Possessive Attributes that acquire symbolic value" (Kress and van Leeuwen, 2006: 97) based on causal-continuity relations.

   Norrick's (1981: 62-63) discussion of the indexical relation between a costume (as analytical features) and its wearer's attributes is particularly relevant here:

> It appears historically accurate to claim that costumes arise as a result of the efforts of certain groups to set themselves off from others, i.e. to be recognized as such and, consequently, as different from other groups…Theater and cinema trade on our productive application of this principle to identify their characters as members of certain national, historic, professional and other groups…Contemporary examples proliferate in cultures where one class is distinguished from another, one age group from another, one sex from the other on the basis of modes of

dress…Various items of clothing, jewelry etc. will be accepted as evidence of the presence of a member of a particular ethnic, religious or age group.

Norrick (1981: 67) further suggests that "any physical object recognizable as the property of someone will serve as a sign of that person: we intend our homes, cars, clothing and accessories to stand for ourselves to all observers". This relation thus produces a metonymy ANALYTICAL FEATURE FOR ATTRIBUTE. The relation between physical appearance and people's perceived quality/identity has been extensively studied in theories of nonverbal communication (e.g. Andersen, 2008; Richmond et al., 2008). Right or wrong, receivers of this physical information make attributions about our attractiveness, competence, moral character, social status and friendliness (Andersen, 2008: 32). Good-looking people are perceived as more talented, kind, more credible and more honest (Richmond et al., 2008). Taller people are perceived as more capable, more confident, more powerful, and are more likely to succeed (Henley, 1977). Painter, Martin and Unsworth (2013: 56) also point out that "while the only experiential aspect of a character that can be explicitly inscribed into an image is his or her physical appearance, the viewer can, as in life, infer other categories, such as age, class, ethnicity, role and place in the family".

Aside from being manifested through their own visual features, the abstract attributes of participants can also be constructed by the visual features of other entities/persons which are affected by the attributes. This strategy is frequently used in advertisements where the effectiveness of the product is metonymically constructed through analytical features of the characters. For example, the effectiveness of slimming pills is constructed through the character's slim body; the effectiveness of shampoo through the character's smooth and shiny hair; the effectiveness of toothpaste through white teeth, and so on. These analytical features are assumed to be caused by the quality of the product, and they often co-construct the effectiveness of the product together with characters' reactions, as we shall see in the analysis in Section 5. Here I sum up by stating that the evaluative attributes of a person or entity are often metonymically constructed through his/her/its visual analytical features, or through the features they bring about in other things.

Constructing Social Stereotypes in Generic Representation

As explained in Section 3, in generic representations, the member does not only stand for the category, but may also invoke more abstract stereotypical knowledge of the category. As Lakoff (1987: 79) points out, "a member or subcategory can stand metonymically for the whole category for the purpose of making inferences or judgments". In the two sub-sections above, I discussed the metonymic construction of evaluative attributes through actions and analytical features. In many cases, an attribute is not just about the specific participant, but represents that of the category it belongs to, that is, it constructs the generic identity of the participant. In relation to attitudinal meaning, generic identities are often judged according to social stereotypical knowledge. For example, the stereotypical second-hand car salesman in some cultures is eloquent but dishonest and the stereotypical doctor is trustworthy. Therefore, we get the metonymy IDENTITY FOR STEREOTYPICAL ATTRIBUTE (OF THE CATEGORY). As Lakoff (1987: 79) argues, social stereotypes are cases of metonymy—where a member or subcategory has a socially recognized status as standing for the category as a whole, usually for the purpose of making quick judgments about people. Lakoff (1987: 85) continues that "since social stereotypes are commonly used to characterize cultural expectations, they tend to be exploited in advertising and in most forms of popular entertainment".

To use the film *Pretty Woman* as an example (Marshall, 1990), the protagonist Vivian is depicted using the shots reproduced in Figure 5. After a scene in which prostitutes are soliciting customers, the film cuts to a close-up of a woman in bed with black lace underwear. The woman gets up, slips on her pull-up stockings and zip-up boots and slides into her trademark "hooker" outfit: a pink halter that is attached to a black mini skirt by way of a big, silver ring. The choice of such visual representation, also with the hints given by the appearance of the previous prostitutes, suggests that she is also a prostitute. Negative judgments of her social status and morality naturally follow from our stereotypical knowledge of this profession. Her identity is constructed by the analytical features, but our judgment of her is based on the knowledge of the category (i.e. her identity), even before we know anything about her as an individual.

Insert Figure 5 here

Another genre that constantly invokes such stereotypical knowledge is TV commercials. Because TV commercials are short, and the space for the elaboration of characters' attributes is limited, they often rely on a character's identity to invite viewers to jump to conclusions about

his/her attributes. For example, a doctor is perceived to be an expert and honest and his/her comment on a health product is therefore more reliable, so advertisements for health products often use representational resources to assign the fictional identity of doctor to characters, in particular, through actional (e.g. medical checkup) and analytical processes (e.g. costume).

The attitudinal meanings based on metonymic mappings in reactional processes, actional processes, analytical processes and classificational processes are summarized in Figure 6. In visual images such as films and advertisements, the eliciting conditions, reactions, clothing, identities are not authentic as in real life, but are semiotic discursive constructs designed by the artists. The framework enables us to map out the visual choices available in representing attitudinal meaning and what choices are made for intended communicative effects. In Section 5, I will apply the theoretical framework proposed in Section 3 and this section in the analysis of a TV commercial.

Insert Figure 6 here

CASE STUDY: ANALYSIS OF A TV COMMERCIAL

In this section, I will elucidate the metonymic mappings in the visual sign making of a Colgate toothpaste TV commercial, as illustrated in Table 1. I will explain the partiality of the representation and the metonymic representation of the characters' attributes by analyzing actional, reactional, analytical and classificational processes, and explain how they contribute to the discursive purpose of persuasion.

Insert Table 1 here

First, in terms of actional process, the advertisement has three scenes which represent three social activities respectively: reporting, dental check, and explaining. There are four main characters, the reporter, the dentist, the patient and the Colgate scientist. The representation of the actional processes is partial in the sense that only brief moments of the whole activity are represented (MOMENTS OF AN ACTION FOR THE WHOLE ACTION). In particular, a dental check is a complex process, but here it is represented with a four-second shot (image 2a). The brief moments are used to activate our schematic knowledge of these activities, which serve to contextualize the product and dilute the activity of promoting.

Second, in terms of analytical processes, the characters are represented with medium and close-up shots (PART OF AN ENTITY FOR THE ENTITY). Such framings exclude non-essential features and serve to highlight the characters' actions and reactions. In particular, in image 2c, the close-up shot is used to highlight the woman's facial expression and her white teeth. Aside from highlighting, the framing also serves to construct differing degrees of symbolic social distance between viewers and the characters (cf. Kress and van Leeuwen, 2006). The reporter, dentist and Colgate scientist are all represented by medium shot and are perceived as "socially close" to viewers, while the patient is represented by close up shot which is perceived as "intimate" to viewers. The representation is highly naturalistic, with maximum details depicted in the background, which aims to make viewers perceive the scenes as real.

Actional and analytical processes are the primary resources for constructing character identities. That is, their identities are not explicitly labeled (as apparently they are just actors/actresses), but are metonymically represented by what they do and what they wear (ACTION FOR ATTRIBUTE, ANALYTICAL FEATURE FOR ATTRIBUTE). The identity of the reporter is represented by his action of reporting and the microphone he is using; the identity of the doctor is visually represented by the actional process of checking up the patient and the analytical feature of his uniform; the identity of the patient is visually represented as the "goal" of the check-up in image 2a, and image 2c refers back to it. Unlike the first three characters, the identity of the Colgate scientist is represented as real, which is articulated through verbal labeling in the caption, and co-constructed by the analytical feature of the uniform and the action of explaining.

Third, in terms of classificational processes, the characters are not represented as specific individuals, but stand for the category they belong to (MEMBER OF A CATEGORY FOR THE CATEGORY). Therefore, their identities and our stereotypical knowledge of the identities are crucial in this multimodal text. The characters' identities activate our stereotypical knowledge and attitudes toward the group they belong to (IDENTITY FOR STEREOTYPICAL ATTRIBUTE). In this example, the identities of the reporter, dentist and Colgate scientist invoke viewers' positive judgment about them and hence the reliability of their utterance. The reliability comes from the "fact" that everyone is doing his own job, rather than directly selling the product. The advertised information then invokes our knowledge and attitude towards the social practice. Specifically, the popularity of the product is metonymically represented as the news report of the result that

"everyone here uses Colgate toothpaste". Viewers' stereotypical judgments about the objectivity of reporters and news reports lend credibility to this claim. Next, the effect of the product is metonymically represented through the dentist's diagnosis. The dentist recounts that the last time the patient had had three buccal problems, but after using Colgate, "it is much better now". The reliability of the dentist and his report makes the effect of the product trustworthy. It can be seen that both the reporter and the dentist are just "doing their jobs", instead of promoting the product, and it is this illusion that makes their identities effective. In contrast, the identity of the Colgate scientist is real and his job is to provide scientific evidence for the effectiveness of the product. Therefore, he is in the position to talk about the effect of the product. But as a scientist, he is not offering subjective opinion, but reporting the objective results of "clinical experiments".

Finally, in terms of reactional process, only the female patient shows a clear facial expression. Her broad smile metonymically constructs happiness (REACTION FOR EMOTION) and invokes viewers' positive emotion and the desire to identify with her. Her smile and white teeth further metonymically construct the effect of the product, which is the eliciting condition (REACTION FOR ELICITING CONDITION). Although the emotion is simple in this example, many TV commercials depict contrasting facial expressions to show the effect of the product. First, a real life problem is created which is followed by a close-up shot of a sad face, and then the product comes as solution, which is followed by another close-up shot of a smiling face. The representation of contrasting reactional processes not only metonymically constructs characters' emotions and the effect of the product, but also engages viewers' emotions through empathy and the desire to identify (Feng, 2012: 277).

The case study demonstrates the fundamental role of metonymy in visual meaning making on the one hand, and elucidates the implicit mechanism of persuasion on the other. Advertisers attribute their claims to characters with different identities whose stereotypical attributes lend credibility to the advertised information. The fictional identities are metonymically constructed through actional and analytical processes. The identities and actions of the characters are carefully designed to elicit desirable rational judgment and emotional engagement from viewers within the short duration of the advertisement. The effect of the product is not directly promoted, but metonymically represented through news report, diagnosis, and patient reaction.

The framework is therefore also useful for the analysis of representational choices in complex multimodal discourse such as advertisement and film. As Bateman and Schmidt (2012:

24) point out in analyzing film, "the analysis demands powerful theoretical and technical tools whose principal focus is *signification* itself (emphasis added). Without this, there is little guidance of what lower-level patterns to focus on and why, and accounts proposed at higher levels of abstraction remain overly subject to intuitive and impressionistic descriptions". The framework of visual metonymy brings semiotic tools which are discussed separately under representational meaning (Kress and van Leeuwen, 2006), interactive meaning and modality (Kress and van Leeuwen, 2006), and attitudinal meaning (Martin and White, 2005) into a coherent, interrelated whole for analysts to scrutinize the choices in the signification process. Analysts are *guided* to focus on essential representational features for their purposes in a principled manner and to make interpretations based on the systematic semiotic descriptions. Two general guiding questions based on the framework (Figure 4 and Figure 6) are: (1) what are the metonymic targets that are constructed? (2) what are the metonymic sources that are selected? Specifically, an analyst may ask the following questions to guide his/her analysis of representational meaning. The analyst should then address why such choices are made in specific contexts.

a. Actional Process. What social actions are represented? What salient moments are selected? What emotion and attribute can we infer from the actions as eliciting condition?

b. Reactional Process. What are the emotions represented? How are they represented? What reactions are depicted?

c. Classificational Process. What characters are represented? What categories do they stand for (e.g. profession, age, and ethnicity)? What stereotypical attitude do they invoke from viewers?

d. Analytical Process. How are the characters depicted? What are the camera angles and shot distances? What is the level of abstraction? What are their analytical features (e.g. appearance, clothing, accessories)? What are the attributes metonymically constructed by the analytical features?

CONCLUSION

This study proposes that visual meaning making is to a large extent metonymic and provides a framework to theorize the types of visual metonymies. Arguing that a systematic account of visual metonymy has to be based on a comprehensive theory of visual images, this study

provides a social semiotic modeling of the visual manifestations of metonymy in the representational structure of visual grammar (Kress and van Leeuwen, 2006). Drawing upon Radden and Kövecses (1999), I propose two types of visual metonymic mappings, namely, the partial representation of reality, and the invocation of attitudinal meanings (Figure 2). Systems of metonymy in actional, reactional, classificational and analytical processes (Figure 4 and Figure 6) are developed to map out the types of metonymies in visual representation. A wide range of theories are draw upon to support the metonymic mappings, such as cognitive psychology, Norrick's (1981) semantic theory, film studies, and nonverbal communication.

The study provides an alternative understanding of the nature of visual semiosis by suggesting that the familiar analytical tools we work with (e.g. camera angle, modality, and attitude) are the result of metonymic processes. While we all accept that visual images are iconic, this study demonstrates systematically that they are also indexical (i.e. metonymic), in terms of their representation of both objects/events and abstract concepts. In doing so, it also brings a wide array of resources under a coherent framework for analysts to scrutinize the representational choices. The study demonstrates that the integration of social semiotics and cognitive linguistics is beneficial to both. On the one hand, the social semiotic visual grammar enables us to describe the visual realizations of metonymy in a systematic way; on the other hand, the conceptual metonymy theory provides us with new insights into the process of visual meaning making. It is argued that such integration should continue to further develop the theory and to deal with the complexity of multimodal discourse. Meanwhile, the theory also needs to be evaluated against data to locate the inadequacies, and to "establish the degree to which they can cover and explain uses of multimodality more generally" (Bateman, 2014: 238).

REFERENCES

Andersen, P. (2008) *Nonverbal Communication: Forms and Functions*. Long Grove: Waveland Press.

Bartlett, F. (1932) *Remembering*. Cambridge: Cambridge University Press.

Bateman, J. (2014) 'Using multimodal corpora for empirical research', in C. Jewitt (ed.) *The Routledge Handbook of Multimodal Analysis* (2ndedn.), pp. 238-252. London: Routledge.

Bateman, J. and Schmidt (2012) *Multimodal Film Analysis: How Films Mean*. London: Routledge.

Bless, H., Fiedler, K. and Strack, F. (2004). *Social Cognition: How Individuals Construct Social Reality*. Hove: Psychology Press.

Bordwell, D. (1985) *Narration in the Fiction Film*. Madison: University of Wisconsin Press.

Bordwell, D. (1989) *Making Meaning: Inference and Rhetoric in the Interpretation of Cinema*. Cambridge: Harvard University Press.

Carroll, N. (1996) *Theorizing the Moving Image*. Cambridge: Cambridge University Press.

Ekman, P. and Friesen, W. (1975) *Unmasking the Face*. Englewood Cliffs, NJ: Prentice Hall.

El Refaie, E. (2003) 'Understanding visual metaphor: the example of newspaper cartoons', *Visual Communication* 2(1): 75-95.

Feng, D. (2012) *Modelling Appraisal in Film: A Social Semiotic Approach*. Unpublished PhD Thesis. Singapore: National University of Singapore.

Feng, D. and O'Halloran, K. (2012) 'Representing emotion in visual images: A social semiotic approach', *Journal of Pragmatics* 44 (14): 2067-2084.

Feng, D. and O'Halloran, K. (2013a) 'The visual representation of metaphor: A social semiotic perspective'. *Review of Cognitive Linguistics* 11(2): 320-335.

Feng, D. and O'Halloran, K. (2013b). 'The multimodal representation of emotion in film: Integrating cognitive and semiotic approaches', *Semiotica* 197: 79-100.

Feng, D. and Wignell, P. (2011) 'Intertextual Voices and Engagement in TV Advertisements', *Visual Communication* 10(4): 565-588.

Forceville, C. (1996) *Pictorial Metaphor in Advertising*. Routledge, London.

Forceville, C. (2005) 'Visual representations of the idealized cognitive model of anger in the Asterix album La Zizanie', *Journal of Pragmatics* 37(1): 69-88.

Forceville, C. (2009) 'Metonymy in visual and audiovisual discourse', in E. Ventola and A. Moya (eds.) *The World told and the World Shown: Multisemiotic Issues*, pp. 56–74. Basingstoke/New York: Palgrave Macmillan.

Forceville, C. (2013) 'Metaphor and symbol: SEARCHING FOR ONE'S IDENTITY IS LOOKING FOR A HOME in animation film', *Review of Cognitive Linguistics* 11(2): 250-268.

Forceville, C. and Urios-Aparisi, E. (2009) *Multimodal Metaphor*. Berlin: Mouton de Gruyter.

Frijda, N. (1986). *The Emotions*. Cambridge: Cambridge University Press.

Gibbs, R. (1994) *The Poetics of Mind: Figurative Thought, Language and Understanding*. Cambridge: Cambridge University Press.

Gibbs, R. (1999) 'Speaking and thinking with metonymy', in K. Panther and G. Radden (eds.) *Metonymy in Language and Thought*, pp. 61-76. Amsterdam: John Benjamins.

Goatly, A. (2007) *Washing the Brain: Metaphor and Hidden Ideology*. Amsterdam: John Benjamin.

Halliday, M. A. K. (1978) *Language as Social Semiotic: The Social Interpretation of Language and Meaning*. London: Arnold.

Halliday, M. A. K. and Matthiessen, C. M. I. M. (2004) *An Introduction to Functional Grammar* (3rdedn.) London: Arnold.

Henley, N. (1977) *Body Politics: Power, Sex, and Nonverbal Communication*. Englewood Cliffs, NJ: Prentice Hall.

Izard, C. (1971) *The Face of Emotion*. New York: Appleton Century Crofts.

Kendon, A. (2004) *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.

Kövecses, Z. (2010) *Metaphor: A Practical Introduction*. Amsterdam: John Benjamins.

Kress, G. and van Leeuwen, T. (2006) *Reading Images: The Grammar of Visual Design*. London: Routledge.

Kress, G. (2010) *Multimodality: A Social Semiotic Approach to Contemporary Communication*. London: Routledge.

Lakoff, G. (1987) *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.

Lakoff, G. and Johnson, M. (1980) *Metaphors We Live By*. Chicago: University Chicago Press.

Lakoff, G. and Turner, M. (1989) *More than Cool Reason: A Field Guide to Poetic Metaphor*. Chicago: University of Chicago Press.

Lazarus, R. (1991) *Emotion and Adaptation*. New York: Oxford University Press.

Marshall, G. (1990) *Pretty Woman.* Buena Vista Pictures.

Martin, J. R. and White, P. (2005) *The Language of Evaluation*. New York: Palgrave Macmillan.

Metz, C. (1974) *Film Language: A Semiotics of the Cinema* (M. Taylor, Trans.). Oxford: Oxford University Press.

Moya, A. (2013) 'Visual metonymy in children's picture books', *Review of Cognitive Linguistics* 11 (2): 336-352.

Norrick, N. R. (1981) *Semiotic Principles in Semantic Theory*. Amsterdam: Benjamins.

Oatley, K. and Johnson-Laird, P. (1987) 'Towards a cognitive theory of emotions', *Cognition and Emotion* 1(1): 29-50.

Ogden, C. and Richards, I. (1923) *The Meaning of Meaning*. New York: Harcourt, Brace and World.

Ortony, A., Clore, G. and Collins, A. (1988) *The Cognitive Structure of Emotions*. Cambridge University Press, New York.

Painter, C., Martin, J. R. and Unsworth, L. (2013) *Reading Visual Narrative*. London: Equinox.

Panther, K. and Radden, G. (1999) *Metonymy in Language and Thought*. Amsterdam: Benjamins.

Panther, K. and Thornburg, L. (2004) 'The role of conceptual metonymy in meaning construction', *Metaphorik* 6: 91-116.

Pinar, M. (2013) (ed.) Multimodality and cognitive linguistics. *Review of Cognitive Linguistics*, 11(2): 227-235.

Plantinga, C. (1999) 'The scene of empathy and the human face on film', in C. Plantinga and G. Smith (eds.) *Passionate Views: Film, Cognition, and Emotion*, pp. 239-255. Baltimore: Johns Hopkins University Press.

Radden, G. and Kövecses, Z. (1999) 'Towards a theory of metonymy', in K. Panther and G. Radden (eds.) *Metonymy in Language and Thought*, pp. 17-60. Amsterdam: John Benjamins.

Richmond, V., McCroskey, J. and Hichson, M. (2008). *Nonverbal Behavior in Interpersonal Relations*. Boston: Pearson Education.

Schank, R. and Abelson, R. (1977) *Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures*. Hillsdale, NJ: Lawrence Erlbaum.

Scherer, K. and Ellgring, H. (2007) 'Multimodal expression of emotion: Affect programs or componential appraisal patterns?', *Emotion* 7(1): 158-171.

Scott, R. (2000) *Gladiator*. Universal Pictures.

Smith, M. (1995) *Engaging Characters: Fiction, Emotion, and the Cinema*. Oxford: Oxford University Press.

Urios-Aparisi, E. (2009) 'Interaction of multimodal metaphor and metonymy in TV commercials: Four case studies', in C. Forceville and E. Urios-Aparisi (eds.) *Multimodal Metaphor*, pp. 95–117. Berlin/New York: Mouton de Gruyter.

Van Dijk, T. (1976) 'Philosophy of action and theory of narrative', *Poetics* 5(4): 287-338.

Van Leeuwen, T. (2008) *Discourse and Practice*. Oxford: Oxford University Press.

Warren, B. (2002) 'An alternative account of the interpretation of referential metonymy and metaphor', in R. Dirven and R. Pörings (eds.) *Metaphor and Metonymy in Comparison and Contrast*, pp. 113-132. Berlin: Mouton de Gruyter.

Yu, N. (2009) 'Nonverbal and multimodal manifestations of metaphor and metonymies: A case study', in C. Forceville and E. Urios-Aparisi (eds.) *Multimodal Metaphor*, pp. 119-143. Berlin/New York: Mouton de Gruyter.

Figure Captions

Figure 1 The schematic components of attitude

Figure 2 Types of visual metonymy based on the ontology of sign

Figure 3 The partial representation of action in film (fair use of screenshots)

Figure 4 Types of partiality in visual representation

Figure 5 The invocation of character identity in film (fair use of screenshots)

Figure 6 Metonymic mapping between representational and attitudinal meanings

Eliciting Condition ⟹ Attitude ⟹ Reaction

Figure 1

Attitudinal meaning

participant emotions
  reactional process REACTION FOR EMOTION
  various processes ELICITING CONDITION FOR EMOTION

representational meaning

participant attributes
  actional process ACTION FOR ATTRIBUTE
  analytical process ANALYTICAL FEATURE FOR ATTRIBUTE
  classificational process IDENTITY FOR STEREOTYPICAL ATTRIBUTE
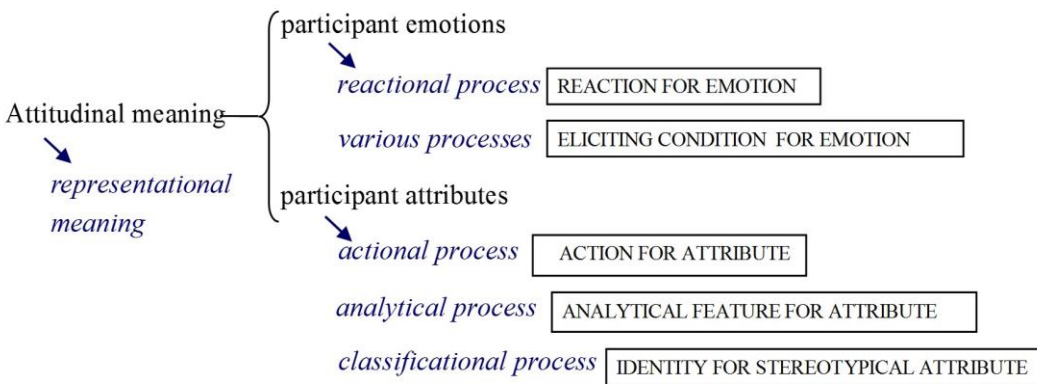
Figure 2

Figure 3a

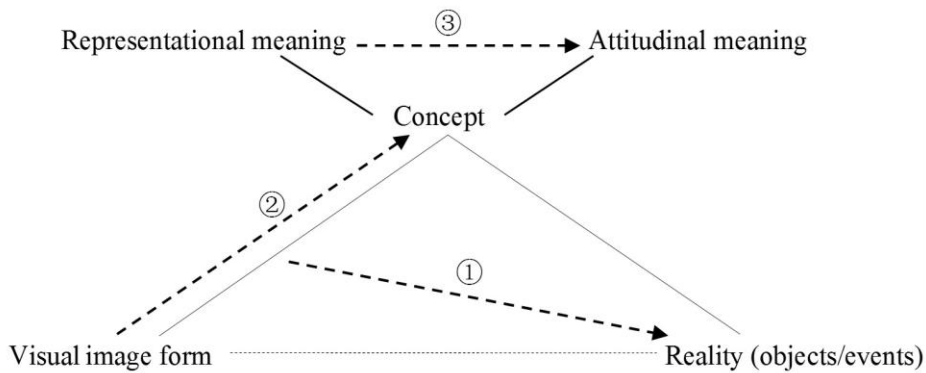Figure 3b

Figure 3c
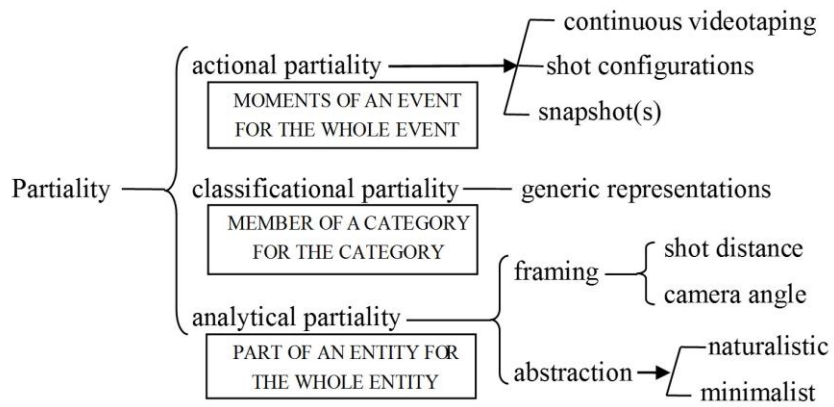
Figure 4
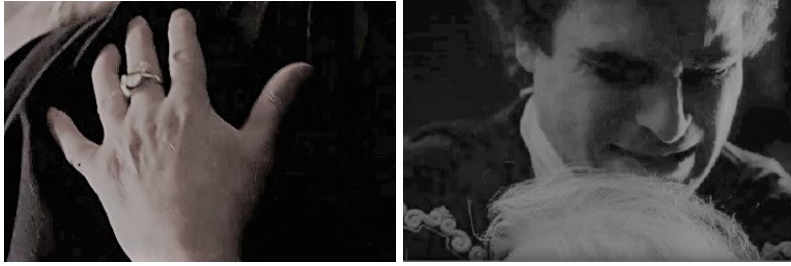
Figure 5

Figure 6

Table 1

Colgate toothpaste TV commercial (fair use of screenshots).

	Visual image	**Image description**	**Sound transcription**
1		Reporter speaking with smiling face.	We are back to sports center. Everyone here uses Colgate toothpaste.
2a		The doctor is re-checking a patient who has been using Colgate.	Last time you had three buccal problems. Let's see how it looks now.

2b		The doctor smiles, with the screen displaying the problems Colgate solves.	Great! It is much better now.
2c		The patient smiling, displaying white, beautiful teeth.	No utterance
3		Colgate scientist speaking.	Clinic experiments proved that Colgate can improve dental health.