A study in cinematic subjectivity
Metaphors of perception in film

Maarten Coëgnarts and Peter Kravanja
University of Antwerp / Katholieke Universiteit Leuven

This article offers a metaphorical and embodied examination of the representation of perception in narrative cinema. Using insights from Conceptual Metaphor Theory we argue that the perceptual states of characters can be represented cinematically via audio-visual expressions of metaphors related to the physical functioning of human bodies. More specifically, we show how a predominant pair of conceptual mappings, namely the metonymy PERCEPTUAL ORGAN STANDS FOR PERCEPTION and the metaphor PERCEPTION IS CONTACT BETWEEN PERCEIVER AND PERCEIVED, plays a crucial role in the non-verbal representation of the characters’ perceptual experience.

Keywords: Cinematic subjectivity, conceptual metaphor, embodied cognition, perception

1. Introduction

During the last two decades, cognitive science has rigorously challenged the old literal view of meaning according to which concepts are inherently disembodied, that is, characterised by abstract symbols that get their meaning solely by virtue of their capacity to correspond to things and states of affairs in the world. Research in cognitive linguistics and cognitive neuroscience increasingly shows that reasoning about abstract concepts is largely shaped by aspects of the human body (Barsalou, 1999, 2008; Gallese and Lakoff, 2005; Lakoff and Johnson, 1980, 1999). For instance, Sweetser (1990) has argued that there exists an extensive subsystem of metaphors, belonging to the generic metaphor MIND IS BODY, according to which our mental life is conceptualised in terms of different kinds of physical functioning (pp. 28–32) (see also Lakoff & Johnson, 1999, p. 235). When considering the mental activity of perception, many scholars have argued that our concepts related to seeing and hearing are conceptualised in terms of corresponding concepts
related to our possession of sensory-motor knowledge (Lakoff 1995; Noë, 2004; Sweetser, 1990; Yu, 2003, 2004). With regard to the latter, Sweetser (1990), for example, has pointed to the importance of physical manipulation and touching as concrete source domains for words meaning ‘sight’ (i.e., visually picking out a stimulus) (p. 38). Similarly, Lakoff (1995) has distinguished two special cases or metaphorical duals, both involving the physical domain of movement: perceiving is touching and perception is reception (p. 139; see also Yu, 2004, p. 676).

In the first metaphor perception occurs “when the perceiver moves his organs of perception to the thing perceived and touches it” (p. 139). Examples include such expressions as “I can’t take my eyes off her” or “They made eye contact” (Yu, 2004, p. 676). By contrast, in the second metaphor, perception occurs “when the thing perceived moves to the perceiver’s organs of perception” (p. 139). This metaphor underlies such expressions as “A comet came into my sight” or “The noise came through the walls” (p. 139). Both metaphors are special cases of a more general metaphor, which Lakoff (1995) refers to as the perception is contact between perceiver and perceived metaphor (p. 139). Following Lakoff, Yu (2004) in turn has suggested that both English and Chinese share the same conceptual metaphors for perception. More specifically, he has shown that although both languages differ on the linguistic level (i.e., the way the conceptual metaphor is manifested in each language), they both conceptualise perception in terms of spatial movement (from the eyes to the target or vice versa) and physical contact, i.e., the contact between the perceptual organ (eyes) and the perceived target. He also stressed the importance of metonymy for perception metaphors, arguing that perceptual organs usually stand for the whole domain of perception (p. 665). For instance, eyes and brows are such important features that they can be metonymically linked to the whole look of a person. Consider, for example, the following linguistic expression: “They keep an eye on the youngsters.” In this example, which involves visual perception, the phrase keeps an eye on is mapped onto the metonymical target domain of seeing. Hilpert (2006) in similar terms refers to this metonymical extension of eye as the instrument for activity metonymy eye for watching (p. 130). In sum, then, one can schematize the underlying metaphorical and metonymical structure of perception as in Figure 1 (after Yu, 2004, p. 680).

The conceptual lower level consists of two mappings onto the same target domain of perceptual experience. The first mapping is metonymical and expresses the relationship between the perceptual experience and the eyes as perceptual organ. The second mapping is metaphorical and involves the projection of elements from the source domain of physical contact between perceiver and perceived onto the target domain of perception. Furthermore, this source domain can take two forms: reception and touching. As Yu (2004) points out, these metonymical and metaphorical mappings “show how lower bodily experiences work their
way up to help conceptualize higher mental experiences, or how the more abstract is understood in terms of the more concrete” (p. 679). In a similar way Shapiro (2011) refers to this upward movement as the “trickle up effect” (p. 88).

1.1 Conceptual metaphors and cinema?

The existence of the above metaphorical mappings has been almost solely proven on the basis of linguistic expressions. They are found in words, sentences and phrases. As such, these verbal metaphors are subjected to the rules and conventions of linguistic meaning. When considering these metaphors in the light of the film medium, however, one is challenged by the fact that cinema, unlike language, does not operate on the basis of arbitrary conventions and linguistic symbols. The following question therefore imposes itself: Can these metaphors be discerned also at the non-verbal level of film? There is at least one good reason for thinking that the answer to this question is “yes.”

According to Conceptual Metaphor Theory (CMT) metaphorical mappings are conceptual rather than linguistic. Metaphor is only derivatively a linguistic phenomenon. Hence, if linguistic metaphors are only the expression of “metaphors
in a person’s conceptual system”, as Lakoff and Johnson (1980, p. 6) claim, then it is plausible to assume that other (non-verbal) modes of expression reflect this activation as well (e.g., Forceville, 2009; Pecher, Boot, & Van Dantzig, 2011). Indeed, by placing meaning on a higher psychological level (i.e., above the arbitrary linguistic rules of syntactic and semantic categories), it instigates a distinction between a conceptual level, on the one hand, and a formal expression or manifestation level (linguistic or otherwise), on the other hand. At the formal level, the specific expressions that manifest the underlying conceptual metaphors can differ. However, at the conceptual level, there exists a commonality that is cross-modal and cross-linguistic.

Considering this argument, one can hypothesize, then, that the study of conceptual metaphors (and meaning in general) does not have to be limited to language alone, for the processes of embodied meaning in non-verbal discourses are probably the very same ones that constitute linguistic meaning (see also Johnson, 2007, p. 208).

Only recently, the plausibility of this claim has been confirmed by various non-linguistic studies (Coëgnarts & Kravanja, 2012a, 2012b; Fahlenbrach, 2007, 2008; Forceville, 2011; Forceville & Jeulink, 2011; Forceville & Renckens, 2013; Ortiz, 2011) that indicate that non-verbal media and visual art forms such as film are making use of the very same conceptual metaphors as experiential elements in meaning that operate in language.

However, with regard to the non-verbal and multimodal manifestations of the metaphors for perception, the academic output has been rather scarce, not to say non-existent. The goal of this article, then, is twofold in its interdisciplinarity, combining both the domain of CMT and the subfield of cinematic subjectivity. The first goal reflects CMT and aims to offer some non-linguistic and filmic evidence of the conceptual metaphor perception is contact between perceiver and perceived. The second goal is an exercise in film scholarship and seeks to provide insight into the question of how the perceptual experience of characters is grounded in film through embodied cognition. Our article is structured as follows.

The first part aims to situate the fundamental concept of our paper, perception (as a specific abstract category of mental activity), within the ongoing film-scholarly discussions of character subjectivity and such related notions as focalization and point of view. Indeed, if we tend to analyse the perception of characters from the perspective of CMT, then it is necessary to clarify the former concept first, before we can discuss its relationship to the latter. More specifically, we will offer a description of characters’ perception based on its difference from other mental faculties (e.g., dreams, hallucinations, beliefs).
The second part, then, offers an inquiry into the metaphorical representation of characters’ perception in film. First, we will discuss some of the theoretical implications of applying CMT to cinema. More specifically, we will centre our attention on the answerability of the following crucial question: How can conceptual metaphors be identified in film? Next we will describe a model of four cinematic ways in which the conceptual metaphor PERCEPTION IS CONTACT BETWEEN PERCEIVER AND PERCEIVED can be established in film. We conclude our article with a concise film analysis of a scene from Alfred Hitchcock’s “Notorious” (1946).

2. The paradox of cinematic subjectivity: Perception vs. higher mental activity

Traditionally speaking, the study of characters’ perception falls into the much debated narratological domain of ‘focalization’ and similar terms such as ‘point of view’ (henceforth POV) and ‘mental perspective’. Originally developed by the French narrative theorist Gérard Genette (1980), the notion of focalization refers to a typology of three different narrative situations (pp. 189–190). The first type is labelled ‘zero focalization’ or ‘nonfocalized narrative’ and refers to the so called omniscient narrator. The second type is called ‘internal focalization’ and corresponds to the situation where the narrative is focused through the consciousness of a character. The third type is defined as ‘external focalization’ and describes the situation where the narrative is focused on a character, not through him. The discussion of perception from a character’s perspective, then, can be placed under the character-based level of focalization: the presentation of the story through the perceptual, conceptual, and other mental equipment of a character. Because characters are usually perceived as humans — they are in the words of Eder (2010) “fictional beings” (p. 23) — both entities are believed to share the same human features. It is within this category of character-focalization that many scholars have pointed to the importance of distinguishing the perceptual experience of a character from more private states of mental experience such as cognition and imagination. Let us consider some of these accounts in more detail.

On the level of the subjective shot, Wilson (2006), for example, has distinguished between two kinds of POV shot: the veridical POV shot and the POV shot that is subjectively inflected. The first kind of POV shots “represent (at least approximately) the visual perspective, anchored in an implicit visual vantage point, of a designed character at a given moment in time” (p. 84). In this case the visual contents are to be imagined as representing the private field of vision of the perceiving character. This type of POV shot is contrasted with the second type, in which the characters’ visual experience is affected by the phenomenological
qualities or contents of his or her psychological condition (p. 85). Consider, for example, a POV shot in which certain special effects of focus and filtering reflect the fuzzy state of mind of a drunken or hallucinating character. In this case the visual content is not primarily determined by his or her perceptual state, but by his or her present state of hallucination. It differs from the veridical POV shot in that a range of the visual properties of the shot are supposed to represent subjective enhancements and distortions of the characters’ field of vision at the time. This subjective infliction of the visual perspective is complete, or in the words of Wilson, “subjectively saturated” (p. 85), if the visual contents are wholly determined by those private visual contents (e.g., dreams or inner visualisations). For instance, when a character memorizes the past, the spectator is instructed not to imagine that he is being provided with information outside the character’s mind. The spectator is mandated to seeing only the content of the characters’ private and psychological state of mind (i.e., the content of the memory), not the content of the characters’ perceptual state.

This contrastive relationship between perception and private forms of subjectivity has also been highlighted by Torban Grodal. In his book “Embodied Visions” (2009), Grodal offers a typology of six key ways in which subjectivity can be elicited in film by way of perceptual distortion (p. 236). One important mode, for example, states that the feeling of subjectivity can be intensified by “a represented space that impedes perceptual access” (pp. 237–239). As Grodal argues, people usually don’t notice the mental processes that mediate our perception of the world. When our view is unrestricted, spaces and objects appear as objective and delineated phenomena in a well-lit three dimensional world (p. 237). However, when our vision or hearing is blocked by such blurring techniques as rain, darkness or fog, we become aware of these unconscious mediating processes. Our perception is challenged, thus giving rise to heightened feelings of subjectivity. Similar to Wilson, Grodal then defines subjectivity primarily by relating it to the concept of perception. More specifically, the former is delineated in function of its differences from the latter.

The same distinction also reflects two of Branigan’s (1992) character-based levels of narration, namely the distinction between “internal focalization, surface” and “internal focalization, depth” (p. 103). According to Branigan internal focalization can range from the field of (a) perception (surface), and (b) impressions, to what he labels (c) “deeper thoughts” (depth) (p. 103). The first mirrors Wilson’s notion of the veridical POV shot and includes the distribution of narrative information through the characters’ unrestricted perceptual experience. The second is similar to Wilson’s category of subjectively inflected POV shots and consists of the out-of-focus POV depicting a character who is drunk or dizzy. The last group co-
incides with Wilson’s third type of subjectively saturated POV shots and includes higher-level forms of subjectivity such as dreams, hallucinations, and memories.

This differentiation of perception from other subjective faculties of a character also mirrors Berys Gaut’s (2010) notion of “imaginative identification” (pp. 255–260). According to Gaut the viewer “imagines” herself to be the character with whom she identifies (p. 255). This identification can take many forms. It can be categorized in terms of imagining what a character is seeing (perceptual identification), imagining feeling what a character is feeling (affective identification), imagining believing what a character believes (epistemic identification), and so on.

Falling back on results from the scientific study of human beings, Eder (2010) in turn proposes a system of anthropological categories for the analysis of fictional characters. More specifically, he carries over three property domains of humans to the domain of what he calls “fictional beings”: ’corporeality’, ‘sociality’, and ‘mind’ (p. 24). It is within the latter that he further distinguishes perception from other mental faculties such as cognition, evolution, motivation and emotion (p. 24). Similar distinctions can also be found in the work of Kawin (2006, pp. 6–12) and Kuhn (2011, pp. 122–133) (for a good discussion of these distinctions, see also Reinerth, 2010/2011).

What all these different accounts suggest, then, is a clear distinction between, on the one hand, the simple visual or acoustic perception of a character, and, on the other hand, such complex forms of introspective subjectivity as memories, internal visualizations and beliefs. Furthermore and more important in the light of the current paper, this distinction is perceived by these authors as being highly hierarchical and asymmetrical, i.e., the character’s degree of subjectivity is defined by the degree of detachment from his or her real sensory perception. There is, however, a paradox to this logic of cinematic subjectivity. On the one hand, perception is necessary to show and to visualize the character’s internal mind set. One can only see or hear the internal content of a character’s mind by externalizing it. This in turn always implies some kind of perceptual experience. On the other hand, however, subjectivity takes its value from what it is not, that is, the absence of perception (partially by distortion or wholly by substitution) is a determinant of the cinematic representation of subjectivity.

Given this dual and paradoxical nature of cinematic subjectivity, one can argue then that this characterization mirrors the same unidirectional and constitutive relationship between perception and subjectivity which is underlying CMT, and in particular such metaphors as understanding is seeing in which the concepts related to knowing and other mental activities are conceptualised in terms of corresponding concepts related to the physical activity of seeing (Johnson, 2007, p. 165; Lakoff & Johnson, 1999, p. 238; Sweetser, 1990, p. 38; Yu, 2003, p. 149; Yu, 2004, p. 669).
However, a cinematic analysis of the semantic extensions of perception first and foremost requires an inquiry into the concept of perception itself. To adopt Branigan’s terminology, we first need to know how “inner focalisation, surface” can be achieved in cinema, before we can address the question of “internal focalization, depth”. This immediately brings us to the central challenge of this article. Perception like any other private mental activity is a subjective matter. As such, it needs to be externalized and objectified. But how then can perceptual experience be clarified, if not by perception itself? As already suggested in the introduction, the answer to this question is again metaphorically and bodily motivated. In what follows, we will argue that the perceptual experience of characters can be represented through the conceptual metonymy PERCEPTUAL ORGAN STANDS FOR PERCEPTION and the conceptual metaphor PERCEPTION IS CONTACT BETWEEN PERCEIVER AND PERCEIVED. This solution, however, implies the application of conceptual metaphor theory to cinema, which in turn instigates the following fundamental question: how can conceptual metaphors be articulated in non-verbal manifestations or formal expressions? Finding an answer to this question will be the first task of the following section.

3. Identifying conceptual metaphors in film

According to Coëgnarts and Kravanja (2012b) there are at least six features or tools which can be helpful to identify conceptual metaphors in film: ‘type’, ‘quality’, ‘modality’, ‘direction’, ‘homospatiality’, and ‘reality’ (pp. 98–104). Let us briefly consider each feature in turn.

Type. In addition to structural-conceptual metaphors, which are mainly discussed by cognitive linguistics, there exist also more short-lived metaphors which involve not the mapping of conceptual domains, but rather the mapping of images. These metaphors which map conventional mental images onto other conventional mental images by virtue of their internal structure are referred to as ‘image metaphors’ (Lakoff, 1987; Lakoff & Turner, 1989) or ‘creative metaphors’ (Forceville, 2009) (e.g., “My wife…whose waist is an hourglass.”). According to Deignan (2007) image metaphors differ from conventional conceptual metaphors such as understanding is seeing or life is journey in two fundamental ways (pp. 175–176). Firstly, they are rich in imagistic detail (rather than in knowledge and inferential structure). Secondly, they are restricted in their application. They only map onto a single topic, rather than onto a domain or class of topics. With regard to film, Coëgnarts and Kravanja (2012c) have shown how various image metaphors of the type CONCRETE IS CONCRETE are predominant in the silent short films of Buster Keaton. In “The Paleface” (1922), for example, there is a scene in
which Buster, a friend of the Native Americans, has confiscated the black suit of an opposing landowner, causing the Indians in turn to react with a rain of arrows. When the first arrow strikes and lands in the grass next to him, the oblivious Buster stretches out his hand to verify whether it is raining or not, thus triggering the image metaphor ARROW IS RAINDROP (p. 136).

Quality. This feature involves the question whether or not the source and target domains of the metaphor (structural-conceptual or image) are directly accessible to the senses (concrete) or not (abstract). While image metaphors always conceptualise the concrete in terms of the concrete, structural-conceptual metaphors usually map the concrete onto the abstract. Common abstract target domains include such diffuse issues as mental faculties, emotions, morality or time. Domains which, due to their lack of clear demarcation, are difficult to grasp and consequently cry out for metaphorical conceptualization. Common concrete source domains include the human body, machines and tools, health and illness or light and darkness.

Modality. The modality deals with the question of how metaphorical thought (structural-conceptual or image) manifests itself to our senses. With no claim to exhaustiveness, Forceville (2009) categorises nine different modes of depiction: ‘pictorial signs’, ‘written signs’, ‘spoken signs’, ‘gestures’, ‘sounds’, ‘music’, ‘smells’, ‘tastes’ and ‘touch’ (pp. 23–24). Using this as a starting point, he further makes the distinction between two kinds of metaphor: ‘monomodal’ and ‘multimodal’ metaphors. The first are metaphors “whose target and source are exclusively or predominantly rendered in one mode” (p. 23). Examples are the prototypical verbal metaphor (which used to be labelled metaphor tout court) and the pictorial or visual metaphor. In both cases source and target domain are rendered in one mode, respectively written signs and pictorial signs. By contrast, multimodal metaphors are “metaphors whose terms are each represented exclusively or predominantly in different modes” (p. 24). In the example from “The Paleface”, for example, the image metaphor ARROW IS RAINDROP is rendered in more than one mode. The target domain is depicted visually: as spectators we see an arrow flying through the air and falling into the ground. By contrast the source domain — the raindrop which is visually absent from the screen — is cued by a simple gesture of the body: Buster who is simultaneously reaching his hand out in the air, checking whether or not it is raining. By a simple sign of the hand, the arrow is thus metaphorically converted into a drop of rain. In the same sense Noël Carroll (1991) refers to “mimed metaphors” (pp. 30–33). However, because the target domain is still depicted visually, Coëgnarts and Kravanja (2012c) prefer to use the term “visual-mimed metaphors” instead (p. 101). Because the metaphoricity is not in the modality, but in some conceptual mapping, the same image metaphor of the type CONCRETE IS CONCRETE could have been cued monomodally as well by using a visual metaphor. Keaton
could for instance have triggered both domains visually by juxtaposing or superimposing an arrow and a raindrop.

As for structural-conceptual metaphors, there can be only one mode for depicting the abstract target domain directly (i.e., the verbal mode). For example, the target domain of time can be represented by the word “time”. Language (spoken or written signs) is, by virtue of its symbolic and arbitrary nature, the only mode capable of rendering the abstract and generic quality of the target domain. Consequently, the target domain of a structural-conceptual metaphor, if present, is usually depicted indirectly or connotatively by means of a metonymy. The abstract target domain death can for example be triggered by the concrete image of a grave (as a metonymy for death), which, contrary to the abstract target domain, can in turn be rendered directly by a visual representation of a grave or the word grave (see also Lakoff & Turner, 1989, p. 8). Contrary to metaphor, metonymy only involves one conceptual domain. One entity in a schema (i.e., grave) is taken as standing for one other entity in the same schema, or for the schema as a whole (i.e., the schema of death as a whole) (Lakoff & Turner, 1989, p. 103).

Direction. A fourth feature concerns the reversibility, or symmetry, of source and target domain. Can we flip the order of both terms? While linguists usually attribute non-reversibility to structural-conceptual metaphors, there seems to be some confusion on the part of image metaphors. Carroll (1996), for example, has claimed that pictorial metaphors are more liable to symmetry than verbal ones. To support his view, he cites examples from mainly surrealistic paintings, like Magritte’s “Le Viol”, which presents the spectator with a unification of a female face and torso. Forceville (2002) on the other hand, backed by his own findings in the field of advertising, denounced this dictum by saying that these examples of creative metaphors are rather atypical when considering prototypical examples of what in general can be labelled pictorial metaphor. Similarly, research in film seems to back the latter view (e.g., Coëgnarts & Kravanja, 2012c).

Homospatiality. This term, which rules out language, has been introduced by Carroll (1996) to designate the appearance of both source and target simultaneously occupying the same spatial bounded entity. For instance, in Vertov’s “Man with a Movie Camera” (1929), there is a scene in which an eye is superimposed over the camera lens, thus provoking the visual image metaphor eye is camera. In this case both elements, the eye and the camera, are co-present or ‘homospatial’ in the same frame. By contrast, if the identification is based on montage or decoupage, and thus breaking the spatial unity between source and target domain, the term ‘non-homospatiality’ is appropriate.

Reality. Finally, where film is concerned, it is useful to make an analytic distinction between two levels of reality where the metaphor can be located. First, there is the level of ‘ante-filmic (or primary) reality’. It encompasses everything
that took place before the camera during the process of filming. According to the French film scholar Étienne Souriau (1951), this level of reality is divided into two sublevels of existence: the ‘afilmic reality’ and the ‘profilmic reality’. The first level can be defined as “the external reality, the real world, which exists outside the filmic realm but functions as a frame of reference for the filmic universe” (Thanouli, 2014, p. 133). The second level is “the part of the real world which is placed in front of the camera and acquires a physical and organic relation to the film” (Thanouli, 2014, p. 133). Both levels of existence are exterior to the filmic level of reality, which is the ante-filmic reality as transformed by the exclusive capacities of the cinematic medium. This includes such specific cinematic stylistic devices as montage or editing, superimposition and cinematography (see also Rohdin, 2009). The main question is then the following: is the filmic level of substantial importance in order to identify the metaphor? In the formerly cited example from “The Paleface”, the filmic reality barely influences the identification of the metaphor. After all, this metaphor could equally well be identified in the primary reality. The film does not provide any information that an unmediated or direct experience of the metaphor would not have provided. The cinematic medium makes possible a vivid representation of the metaphor, but the metaphor is not specifically cinematic in a deep sense. By contrast, in the example from “Man with a Movie Camera” superimposition is essential to provoke the metaphor. In this case a specific cinematic device actively interferes in the ante-filmic reality, in order to communicate the metaphorical mapping to the spectator. Because the filmic level is essential to identify the metaphor, Coëgnarts and Kravanja (2012b) have labelled this a “filmic metaphor” (p. 103), which, however, is not the same as a visual metaphor. As was pointed out earlier, a visual metaphor is a metaphor whose target and source are both represented visually. Thus, a metaphor can be visual without being filmic (and vice versa). The terms are not mutually interchangeable. The example from “The Paleface” is partly visual (i.e., the arrow is rendered visually), not filmic. By contrast, the example from “Man with a Movie Camera” is totally visual (both the camera and the eye are depicted visually) as well as filmic (superimposition is an essential condition for the realisation of the metaphor).

4. Towards a cinematic model of perception metaphors

When considering the conceptual metaphor PERCEPTION IS CONTACT BETWEEN PERCEIVER AND PERCEIVED in the light of these theoretical concepts, one can make the following analysis. The metaphor PERCEPTION IS CONTACT BETWEEN PERCEIVER AND PERCEIVED involves a structural-conceptual metaphor of the type ABSTRACT IS CONCRETE, in which physical closeness is the concrete source domain
that is mapped onto the abstract target domain of perception. Furthermore, the relationship between target and source can be described as non-reversible. It is common to use terms from physical contact to describe aspects of perception, but it is not common to use perception to clarify grasping or manipulating. Less obvious, however, is the question of reality. How can this conceptual metaphor of perception be instigated on the filmic level, that is without words and sentences? Given the dual constitution of conceptual metaphors, this question breaks down into two sub-questions:

1. How can the abstract target domain of perception be represented non-verbally?
2. How can the contact between perceiver (PR) and the object being perceived (OP) be rendered non-verbally?

As we have already seen in the previous section, the answer to the first question is metonymically motivated. From a non-verbal perspective, the abstract target domain of perception can only be triggered indirectly by showing the perceptual organs of the character (e.g., eyes, ears). This condition, however, is rather vague and could suggest mistakenly that every visualisation of a perceptual organ is metonymically intended. Hence, in order to avoid being too generic, some additional conditions are necessary. A first important condition concerns the distance between the camera and the character’s perceptual organ. The shorter the distance between the camera and the character’s facial features, the likelier it is that the shot will be conceived of as standing for perception as a whole. By contrast, if a person, for example, is rendered in a very long shot, it is much less likely that the target domain of perception will be triggered to the viewer, as the perceptual organ is less visible due to the larger distance between the camera and the organ. The underlying perceptual image schema at work here is that of CENTER-PERIPHERY (Grady, 2005, p. 37; Johnson, 1987, pp. 124–125; Lakoff, 1987, p. 267) (see Figure 2).

![Figure 2. The CENTER-PERIPHERY schema as applied to shot sizes](image-url)
Shot size, albeit crucial, is, however, not the only condition to be fulfilled before the metonymical connection to the target domain of perception can be made. The perceptual organ represented also has to be directed towards something (i.e., the perceived object). Perceiving, like any other type of consciousness, entails intentionality: seeing is seeing something, hearing is hearing something, and so on. The character’s perceptual organ needs to be directed towards the perceived object. In film, this object can be located either on-screen or off-screen. Often the direction of the perceptual organ is helpful in triggering this intentional act. For instance, a character who sees something which is off-screen for the viewer will most likely gaze towards the side of the frame where the object is situated. This bodily gesture in turn allows the viewer to deduce the location of the perceived object (and by extension the idea that the character is likely engaged in the act of perceiving). In addition, the act of perceiving can also be further accentuated by attributes belonging to the visual content of the shot such as a window or binoculars (i.e., looking through something).

The answer to the second question is of a more complex nature. Strictly speaking, the contact between PR and OP can be elicited cinematically by four main strategies divided into two categories: the category of the single shot and the category of two shots (see Table 1).

<table>
<thead>
<tr>
<th>Table 1. Categorization of the cinematic ways in which the contact between PR and OP can be elicited</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Single shot + homospatiality: By showing PR and OP together in one shot (e.g., framing or mise en scène);</td>
</tr>
<tr>
<td>2. Single shot + non-homospatiality: By moving from PR to OP (e.g., tracking, panning, tilting, zoom-function);</td>
</tr>
<tr>
<td>3. Two shots + homospatiality: By presenting PR and OP, each occupying a different shot, as co-present or homospatial entities in the same frame (e.g., superimposition, split screen);</td>
</tr>
<tr>
<td>4. Two shots + non-homospatiality: By cutting from PR to OP (e.g., editing).</td>
</tr>
</tbody>
</table>

Within each category both participating elements can be shown simultaneously on screen or not. As in the previous section, the former case will be labelled ‘homospatiality’. However, we use the concept in a slightly different manner here, to refer to the co-presence of PR and OP in the same larger frame (rather than to the co-presence of source and target). Thus, if each entity occupies a different space in a different shot, albeit in the same larger frame (for example, in the case of split-screen or superimposition), homospatiality is applicable.

We now proceed to a more detailed discussion of each strategy. The various devices have been differentiated for analytical purposes, but they often work
together or are intertwined. We commence our discussion with the category of the single shot.

4.1 Framing and *mise-en-scène*

In the first strategy, perceiver and the object perceived appear simultaneously on-screen in one single uninterrupted shot. This category can take two forms. Either the relationship is established in a POV shot or it is not. In the former case, the visual content of the shot is wholly determined by the perceiver’s perceptual field of vision (i.e., the perceived objects). The physical presence of the perceiver, then, is suggested metonymically by elements in front of the camera, such as hands or smoke, or by the use of a reflective surface in the *mise-en-scène* (e.g., a mirror). In the latter case, we can make a further distinction between two types of non-POV shot: the ‘personal or semi-subjective non-POV shot’ and the ‘impersonal or objective non-POV shot’. In both cases, the perceiver’s body as the locus of perception is present as a distinguishable physical entity on screen. The first type recalls that what Jean Mitry (1965) refers to as the ‘semi-subjective shot’ (pp. 77–78). This type usually has the camera placed somewhere behind a character in the foreground, while the object of his or her gaze is also included in the background of the frame, as in, for example, “Reflections in a Golden Eye” (John Huston, 1967, see Figure 3). In this case, the link is created by using depth of field. In practice, this tension is often further enhanced by such deep focus techniques as the use of racking focus or the split-field diopter lens, the latter allowing an approximation of depth of field by bringing both near and far planes into focus simultaneously (for a good discussion of this technique see Ramaeker, 2007). The second type, the impersonal or objective non-POV shot, is usually centred around the horizontal axis, and has the camera commonly placed next to the character, showing his or her face in profile on one side of the frame (left or right), while the object of his or her gaze occupies the opposite side. As Deleyto has pointed out, this kind of linking between PR and OP depends at least on two important elements, namely framing and the perceiver’s gaze (p. 173). Although this is the most common way to depict an objective and homospatial link between PR and OP, the objective type can also manifest itself in the form of a single shot, in which the perceived object is reflected in (1) the eye of the beholder (see Figure 4) or (2) the reflective surface of a window or other piece of glass through which the character sees something. In the latter case, the camera sits on the opposite side, so that the viewer sees what the perceiver is seeing (i.e., the reflection), as well as the expression on his or her face as he or she sees it.
4.2 Movement

In the second strategy, the relationship between perceiver and perceived is insti-
gated by moving the camera’s point of view from the perceiver’s standpoint to-
wards the perceived entity. This movement can be elicited from a fixed position
(i.e., the zoom function) or not (i.e., the mobile camera). In both cases, the un-
derlying schema at work is that of the source-path-goal image schema (e.g.,
Forceville & Jeulink, 2011; Johnson, 2007, p. 183) (see Figure 5).

Both stylistic solutions involve movement from a starting point (the perceiv-
er), over a pathway, towards an end point (the perceived). Successful perception,
then, takes place when the character’s point of view ‘reaches’ the target: that is, by
the time the mobile camera, or the zoom function, has continued to decrease the

Figure 5. The source-path-goal image schema
distance between the camera’s, or the character’s, perspective and the perceived object. Like the previous strategy, this device can be viewed as a way to connect two entities without interference of editing. However, it differs from the first in that PR and OP are not simultaneously present during the whole length of the shot. As the camera moves towards the perceived object, the perceiver’s presence (depicted either totally or partially) is abandoned to the off-screen space. Furthermore, like the first strategy, the starting point can be represented in three ways, that is the relationship between the camera and the perceiver can be motivated objectively, semi-subjectively or subjectively.

So far we have been dealing with the level of the single shot. The relationship between perceiver and the perceived object is established within the spatial unity of one shot. However, this link can also be realized on the level of two shots: shot A, that establishes the origin of the perceiver, and shot B, that shows the perceived object. This relationship between A and B can take two forms. Either shot A and shot B are shown simultaneously in the larger frame of the screen (i.e., homospatially), usually via such cinematic devices as superimposition or split-screen imagery, or they are shown successively, which is a defining characteristic of editing. We will consider each strategy in turn.

4.3 Superimposition and split-screen or multiple-frame imagery

In the strategies of superimposition and multiple-screen imagery, or split-screen imagery, two images, one containing the perceiver (shot A), the other the perceived object of his vision (shot B), appear simultaneously within the larger frame. Like the first strategy, this process enables the viewer to see both the observer’s gaze and the object of his gaze at exactly the same moment (i.e., homospatially). It differs from it, however, in that both entities are separated by their own frame dimensions and shape. The difference between superimposition and split-screen lies in the way both shots are connected within the larger frame. In the first, the image of the perceived object is laid over the image of the perceiver, as in, for example, “La Signora Di Tutti” (Max Ophuls, 1934, see Figure 6). Notice that the same effect can also be achieved without the need for a second shot (i.e., through the objective (reflective) type of the first strategy). In the second, both entities are placed next to each other with the perceiver situated on one side of the frame, and the perceived object is located on the other side of the frame, as in, for example, “Femme Fatale” (Brian De Palma, 2002, see Figure 7).
4.4 Editing

We conclude our list of cinematic devices with editing. This technique differs from the previous one, in that both entities are not shown simultaneously within the same larger frame, but successively. A shot of the perceived object is pasted after a shot of the perceiver. Usually this takes the form of the shot/reverse shot or the eye line match. Shot A shows an objective and external shot of a character looking off-screen. Shot B shows what the character is looking at, from a position which is either occupied by the character (subjective shot) or not (eyeline match). Because
both entities are not co-present in the same frame, homospatiality is not appropriate. Theoretically speaking, shot A can be presented semi-subjectively, or subjectively. In practice, however, the perceiver in shot A is usually depicted in an objective way. That is, in order to elicit the idea in the viewer that the visual content of shot B coincides with the perceiver’s perceptual perspective, it seems that an objective representation of the perceiver is desired. As Deleyto (1991) rightly points out: “It is as if film narratives required a constant return to objective presentation for a better understanding of the internal gazes that occur in the text” (p. 170). In order to accentuate the subjective quality of the visual content of shot B, this shot is often given movement from the perceiver’s point of view (which falls together with the camera’s point of view) towards the object of his or her perception (cf. the subjective variant of movement). In this case, the viewer initially might think that the suggested movement, either elicited by the mobile camera or by the zoom-in function, coincides with the physical movement of the character’s face and body. Often, this presumption is, however, subverted in a subsequent shot showing the character still stationary from the same position. As with the second strategy, this kind of movement has to be interpreted metaphorically, i.e., as a concrete means to convey and imitate the character’s abstract perceptual experience.

To sum up, then, one can schematize the relationship between the underlying conceptual metaphorical structure and the different cinematic devices as in Figure 8 (below).

In order to test the empirical value of this model, we will now analyse a specific case-study taken from Alfred Hitchcock’s (1946) film “Notorious”. We have chosen this particular film because it is usually considered as a prototypical and canonical example of cinematic subjectivity.¹

5. Case study: “Notorious” (1946)

The scene in question involves the moment in which Alicia (Ingrid Bergman) tries to steal Sebastian’s (Claude Rains) keys to the cellar. At first, the film shows a medium shot of Alicia’s face as she is standing still at the cellar.

¹. All film stills in this contribution are treated as visual citations, in accordance with the established guideline for fair use of film stills from DVDs in scholarly writings. Figures 3 & 4 are captures from “Reflections in a Golden Eye” (John Huston, USA 1967; DVD edition: Warner Home Video, 2006, NTSC, 108min). Figure 6 is a capture from “La Signora di Tutti” (Max Ophüls, Italy 1934; DVD edition: Masters of Cinema, 2010, PAL, 86min). Figure 7 is a capture from “Femme Fatale” (Brian de Palma, France/Switzerland 2002; DVD edition: Paradiso Home Entertainment, 2003, PAL, 110min). Figures 9 & 10 are captures from “Notorious” (Alfred Hitchcock, USA 1946; DVD edition: Dutch Filmworks, 2003, PAL, 101min).
A study in cinematic subjectivity 167

room (see Figure 9). This shot can be interpreted in terms of the conceptual metonymy PERCEPTUAL ORGAN STANDS FOR PERCEPTION. The target domain (i.e.,

<table>
<thead>
<tr>
<th>Conceptual level</th>
<th>Audiovisual expression level</th>
<th>Conceptual level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homospatiality</td>
<td>Subjective</td>
<td>Object</td>
</tr>
<tr>
<td>Non-homospatiality</td>
<td>Perceptual experience</td>
<td>Subjective</td>
</tr>
<tr>
<td>Camera movement</td>
<td>Hand-held camera</td>
<td>Non-subjective</td>
</tr>
<tr>
<td>Superimposition</td>
<td>Split screen</td>
<td>Objective</td>
</tr>
<tr>
<td>Level of two shots</td>
<td>Splitscreen</td>
<td>Subjective</td>
</tr>
<tr>
<td>Physical contact</td>
<td>Zoom function</td>
<td>Objective</td>
</tr>
<tr>
<td>Touching</td>
<td>PROP</td>
<td>Objective</td>
</tr>
<tr>
<td>Reception</td>
<td>PROP</td>
<td>Objective</td>
</tr>
</tbody>
</table>

Figure 8. A cinematic model for analyzing a character's perception in film
the human activity of watching) is metonymically expressed by making visual reference to the body part that is most crucially involved in the activity of seeing. In this way, her eyes stand for watching, which can be subsumed under the more general basic conceptual metonymy body part for its typical functions and for the attributes connected with them (Barcelona, 2003, pp. 265–6). Consequently, framing and the direction of her gaze are constitutive formal factors in communicating this metonymical relationship, and the idea that Alicia is looking into Sebastian’s room. Then the film cuts to Alicia’s point of view, by means of a static shot showing the content of her eyesight: at the back, the reflection of Sebastian’s shadow on his bathroom door, and in the front, his keys which are lying on the desk. In this way the non-homospatial cinematic solution of editing is used to elicit the source domain connection between the perceiver and the

Figures 9 & 10. “Notorious”
OBJECT PERCEIVED. Next the film becomes dynamic, as the camera moves closer from Alicia’s initial position (i.e., the source) to the desk (i.e., the goal), thereby connecting Alicia’s point of view with the keys (see Figure 10). Editing is now substituted by a second filmic solution (viz, camera movement). Initially, one might think that this tracking shot of the camera coincides with the movement of Alicia’s face and body. This presumption, however, is immediately subverted in the subsequent shot showing Alice still stationary at the threshold of the room. She has not moved. Her eyes have made contact with the target, although her body and hands have remained still. The function of the tracking shot is to instigate a physical connection between Alice and the keys as the object of her gaze, thus giving metaphorical expression to her perceptual experience (i.e., the target domain). Only when, in the following shot, she finally decides to physically approach the desk, does her bodily movement take over the place of the schematic spatial movement of the camera, and the contact between perceiver and perceived is made real again: she actually grasps the keys.2

6. Conclusion

In this study we have presented a formal model for analyzing characters’ perception in narrative cinema. Using insights from CMT we have argued that the perceptual experience of a character (as part of the larger domain of character subjectivity) can be represented cinematically via audio-visual expressions of metaphors related to the physical functioning of our human bodies. More specifically, we have demonstrated how a pair of conceptual mappings, namely, the metonymy perceptual organ stands for perception and the metaphor perception is contact between perceiver and perceived, can manifest itself non-linguistically by means exclusive to the cinematic medium (e.g., camera movement, editing, etc.). By providing a non-verbal account of conceptual metaphor in film, our analysis helps to validate CMT’s dictum that metaphor is primarily a matter of thought, and only derivatively a matter of form. CMT also has some important merit for film studies. As a theory concerned with the bodily underpinnings of meaning-making, CMT can provide some insight into the question of how abstract meaning

---

2. The same scene has been analysed by Guerra and Gallese (2012, pp. 200–202). Although they do not explicitly refer to the relationship between metaphor and perception, they similarly acknowledge the embodied nature of the scene. Using Embodied Simulation (ES) theory, they argue that the tracking shot not only mirrors Alicia’s potential approach, but also the viewer’s own potential approach, which, as they state, “turns into a grasping simulation the more the keys are made ready-to-hand, thus evoking the activation of the viewer’s canonical neurons” (p. 201).
is constructed in film. The study presented in this article only reflects one such case. However, systematic studies of visual data may reveal how other abstract concepts are grounded in sensory-motor knowledge as well. For instance, new studies should be able to shed light on how perception, once itself clarified metaphorically, can in turn be extended to other ‘higher’ mental faculties. Indeed, if perception, according to CMT, is used in language as a source domain for such mental functions as thinking and knowing, then it is plausible to assume that the same processes of embodied meaning are underlying the filmic representations of these abstract target domains as well.

References


**Corresponding author’s address**

Maarten Coëgnarts
Department of Communication Sciences, Faculty of Political and Social Sciences
Research group Visual Studies and Media Culture
University of Antwerp
City Campus — Building M
St-Jacobstraat 2
2000 Antwerp
Belgium
maartencoegnarts@gmail.com
Biographical notes

Maarten Coëgnarts has an MA in Film Studies and Visual Culture (University of Antwerp) and an MA in Sociology (University of Antwerp). His research primarily focuses on metaphor in film and embodied visual meaning. He also has a special interest in film analysis and in the relation between film and philosophy. He is currently preparing a PhD in film studies at the University of Antwerp.

Peter Kravanja is a research fellow at KU Leuven (Faculty of Arts, research unit Literature and Culture). He holds an MS and a PhD in Mathematical Engineering and Computer Science (KU Leuven), an MA in Cinema Studies (Université de la Sorbonne Nouvelle — Paris 3) and a BA in Philosophy (KU Leuven). His research interests include analytic philosophy of art applied to cinema, questions concerning analysis, interpretation and form, as well as the relation between film and the other arts.