

Cinematography techniques and their interpretative power

Francesca Proni

University of Siena

francescam.proni@gmail.com

Abstract Schlenker’s proposal of “Super Semantics”, which is aimed at the extension of the application field of formal semantics to non-linguistics objects, based on the assumption that language represents one of the vast modalities of human representation, paved the way for many studies and approaches in the field of depiction. This extension is allowed by iconicity, which is a property that can be retrieved not only in language (gestures) but also in pictures. In this paper I will try to present the most relevant studies in the field of depiction, such as the ones by Abusch and Greenberg, to underline those concepts which are suitable to the development of my own proposal. In fact, with this work I am willing to investigate the mutable relation between natural language and cinema throughout some examples of cinematography techniques that may replicate the way in which verbal discourse produces meaning. More specifically I will focus on those modalities and characteristics that establish a coherent cinematographic discourse.

Keywords: cinematography techniques, iconicity, theory of depiction, multimodality, discourse coherence.

Received 30/11/2021; accepted 03/06/2022.

0. Introduction

Recently, a new branch of studies emerged from the vast variety of subjects belonging to the field of linguistics. The so-called “Super Linguistics”, as defined by Philippe Schlenker, is the application of formal linguistics methodology to non-standard linguistic objects. These methodologies mainly pertain to formal pragmatics and, over the years, many researchers ventured into this newborn discipline applying these linguistic methodologies to a broad range of non-standard linguistic objects such as gestures (Tieu, Pasternak, Schlenker, & Chemla 2018), music (Lewis 1970; Schlenker 2019; Reda 2020), dance (Patel-Grosz, Grosz, Kelkar, & Jensenius 2018), pictures (Greenberg 2011, 2018; Abusch 2013, 2015; Abusch & Root 2017) and animal calls (Fitch & Hauser 2004; Schlenker *et al.* 2016). In few words, Super Semantics is the discipline that aims to integrate formal semantics and semiotics. The theoretical framework proposed by Schlenker could be interpreted as a new contribution to the decades-long debate in the field of semiotics of cinema: “le cinema, langue ou language?” (Metz 1964). This question gave rise to a flourishing literature involving famous critics, directors and philosophers such as Pasolini, Barthes, Eco and Bettetini.

Without going into further details on the whole debate, which has been very faceted and complex, I would like to investigate Metz's position and the relation that this position holds with Schlenker's proposal. Metz and Schlenker's proposals share the same (weak) thrust: both make an appeal to the necessity of expanding the field of application of a discipline to another subject. On the one hand, Metz shared the need of developing a semiotic account for cinema, on the other hand, Schlenker wants to expand the study of formal linguistics to different fields, which were not included in the study so far. However, the underlying frameworks seem to be very different. In fact, Metz's approach is based on the structuralist conception and method of language, while Schlenker's approach seems to be rooted in cognitive linguistics. In Metz's view the concept of language is strictly linked to the act of communication, while in the Schlenker's framework language is an innate cognitive capacity which is shared by every human being. If language is somehow a way of organizing and producing verbal contents could be other human activities such as dance, art, cinema and music be based on a similar processing or system?

Schlenker (2019) argues that non-standard linguistic objects, like music or pictures, display a characteristically linguistic behavior in terms of their inferential properties, providing the ground to extend formal semantics to other media. The most important features that non-standard linguistic objects share with natural language are iconicity and multimodality. We refer to iconicity as the capacity of a medium to represent something through icons. In natural language, iconic modulations can be found in cases where the form of the word resembles its meaning (i.e., "*looong*", which results longer than the conventional word "*long*"). On the other hand, multimodality is the capacity of a medium to extend the meaning of an utterance by interactions with other media. For example, natural language is often enriched by gestures, as well as musical content is enriched by language or filmic content is enriched by language or music.

Iconicity and multimodality contribute to the creation of new word-like elements that can be understood without being known to the speaker through the inferential process that is typical of natural language. It appears that iconic enrichments affect truth-conditions in systematic ways and can replicate the inferential typologies of language. We can conclude that language is multi-modal and non-vocal expressions participate in the same semantic typologies as words and replicate all their inferential typologies (implicatures, presuppositions, supplements and inferences). These findings provide the scope for extending the analysis of the interaction between formal semantics and semiotics beyond standard linguistic objects, such as pictures and films. Acknowledging the existence of previous studies focused on music, gestures, pictures, in this paper I will extend the formal analysis of language to films.

However, Super Semantics represents one of the possible approaches that have been developed in the attempt of investigating the derivation of meaning in other fields. In fact, in this direction another interesting approach is represented by the Conceptual Metaphor Theory (CMT) within the embodied dimension of meaning. CMT claims that «what we call abstract concepts are defined by systematic mappings from bodily-based, sensory-motor source domains onto abstract target domains» (Kravanja and Coëgnarts 2015: 1). In other words, abstract meaning is related to the sensory-motor experience. The CMT theory has also been extended to study of the medium of film, mainly focusing on how abstract content could be communicated through non-verbal aspects of film, therefore without resorting language.

The crucial concept that allows Schlenker's application of semantics to non-standard linguistic objects is iconicity, the property of being able to enrich an "utterance" in ways that do not belong to traditional semantics. In other words, the capacity of "representing something". However, CMT's focus relies on the conceptual metaphor,

defined as the possibility to understand a certain domain of experience in terms of another. George Lakoff describes metaphor as strictly related to the concept of inference:

The heart of metaphor is inference. Conceptual metaphor allows inferences in sensory-motor domains (e.g., domains of space and objects) to be used to draw inferences about other domains (e.g., domains of subjective judgment, with concepts like intimacy, emotions, justice, and so on). (Lakoff and Johnson 2003: 170).

Equally, Schlenker affirms that non-standard linguistic objects may replicate the same inferential typology of language. In other words, inferences allow non-standard linguistic objects to be studied in the same way of language, since they share the same inferential properties.

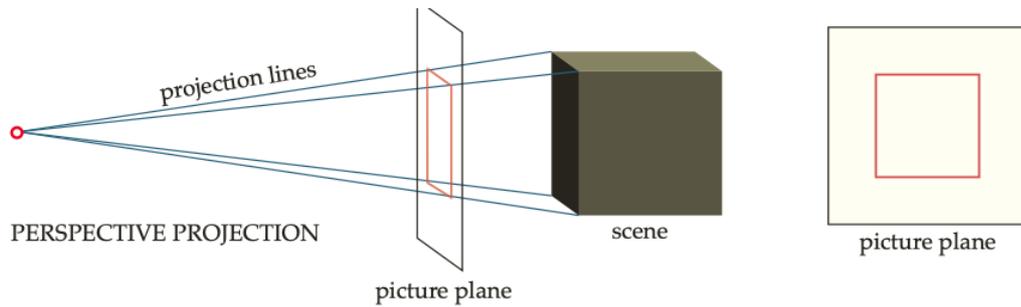
As we can see, even if the two approaches may be similar for some aspects, they completely differ based on their framework. CMT sustains an idea of cognition strictly related to physical states of the body, where metaphor «is primarily a matter of thought, and only derivatively a matter of form (linguistic or otherwise)» (Coëgnarts and Kravanja 2014: 1). Whilst Schlenker's position is mostly eradicated in linguistics, where language is a form of cognition that may be retrieved in other representation modalities.

1. Modern theory of depiction

Before delving into cinematic techniques and their relation to the inferential typologies of language, it is useful to revise the theory of depiction, which is a long philosophical tradition that concerns the study of a form of representation that exceeds the linguistic domain and that can be easily re-conducted to the way of signification of images (Terrone 2015). Many scholars put a great effort in formalizing a picture semantics that could reproduce the tenets of the standard linguistic semantics and a whole literature about "depiction" formed through the last decades. Some proposals are centered on the most common articulations of depiction: pictures. Pictures are considered «vehicles of truth-values» (Eaton 1980) and their signification spawns either from their geometric properties (Greenberg 2011) or from the context in which they captured (Greenberg forthcoming). Other authors, though, depart from the standard declination of the concept of depiction go well beyond the semantics of the single picture. If individual pictures can convey a semantic value on their own, a sequence of pictures may well do the same under certain conditions. Visual narratives, such as comics and films, can also be approached with a semantic eye. Kraft *et al.* (1991) are among the first to establish that the distinction between content and structure in a visual narrative is more or less isomorphic to the distinction between semantics and syntax in standard linguistic objects. Abusch (2013) applies discourse semantics to comics, while Bateman (2012) develops a framework for multimodal discourse analysis for movies. All these attempts are utterly relevant for the project of Super Semantics because they provide a theoretical and empirical ground for extending the field to another class of non-standard linguistic objects: motion pictures.

1.1 From pictures to pictorial narratives: Greenberg, Abusch

The most recent developments of the theory of depiction are based on the Geometrical Projection Theory (GPT), posited by Greenberg (2011). The author develops a model based on “perspective projection”, which is an algorithm that maps scenes into pictures and allows to match the image to its content. The two fundamental elements of this algorithm are the viewpoint and the depiction system. The viewpoint is the place from which we perceive a scene and the depiction system is the set of projection rules that transpose that three-dimensional scene into a two-dimensional picture. The whole process is well exemplified by Figure (1):



(Fig.1)

(Greenberg, 2011)

The accuracy of the picture P , that depicts a certain scene σ , depends on the correct application of a rule of geometric projection G to a certain viewpoint v . In formula:

$$[P] = \{ \langle v, \sigma \rangle \mid G(\sigma) = P \} \quad (1)$$

that reads like “Picture P accurately depicts a scene σ , viewed from a certain viewpoint v , by correctly applying the rules of the geometric projection G to the given scene”.

The force of this model is that it could be extended to other forms of depiction, such as comics and other visual narratives (Abusch 2013). Whereas Greenberg’s GPT treats accuracy just as a product of geometrical rules, Abusch’s extension introduces the communicative aim of the artist as a further element in the system of depiction.

$$[P] = \{ \langle v, \sigma \rangle \mid A(G(\sigma)) = P \} \quad (2)$$

that reads like “Picture P accurately depicts a scene σ , viewed from a certain viewpoint v , by assuming a communicative aim A applied to geometric projection G of the given scene”.

In the expression (2), the artistic transformation A is superordinate to the geometric transformation G but does not modify the indexing relation between the picture P and the scene σ . With this extension, Abusch (2013) widens the theory of depiction by including pragmatics. Pictures, just like common verbal sentences, can be interpreted with an interplay of the semantic approach and the pragmatic approach. When interpreting the semantic value of an image, or a sentence, a set of possible worlds is unlocked. One of them is picked up with the pragmatic interpretation. What is crucial

about the “possible worlds” approach is that dealing with it means checking possibilities. In other words, when a picture’s content picks out a certain possible world, its semantic value is true in that world. At the same time, the picture’s content rules out all the other possibilities.

According to Abusch (2013), two interpretative issues arise with this approach: coreference and temporal succession. Coreference can be found in cases when we deal with a sequence of pictures depicting an action performed or endured by an object or an individual, animal or human. In (2), which is a sequence taken from the popular manga¹ *Death Note* written by Tsugumi Oba and illustrated by Takeshi Obata, where we are observing the fall of an object from the sky into a school’s courtyard.



(Fig. 2)

The semantic model cannot account for the information that leads us to assume that the notebook we see in the top-left picture (let’s call it P_1) is the same notebook appearing in the bottom-right picture (P_2). However, the reader understands that the notebook is the same due to the fact that the existential quantifier of P_1 and P_2 are the same. This interpretation is possible thanks to what Abusch describes as “pragmatic enrichment”, that is an «information that is added conjunctively to literal meaning in constructing the discourse representation for a passage» (Abusch 2013: 17). The same can also be seen in natural language, in the case of a sentence like “He took the money and payed the cab”, where we tend to imply that the man payed the cab with the same money he has taken from his wallet. Therefore, we can conclude that co-reference is a pragmatic process that enriches the meaning of a picture sequence.

Temporal succession is another interpretative issue that is implied by pictorial narratives. They are inevitably structured as sequences, where a scene follows another in chronological order. Consequently, the correct interpretation of the full sequence is only possible when we take in consideration the temporal succession of the pictures. This can be clearly noticed in the following example (3), where we are shown in order: the back of the head of a student that stares out of the window, something falling from the

¹ Manga are traditionally read from left to right. In this dissertation the illustrations follow the traditional disposition.

sky, that same objects landing on the grass and the student that becomes aware of the object.



(Fig. 3)

The correct interpretation of the information of these frames is possible because of the prior knowledge that comics, such as manga, are organized in a fixed structure that itself constitutes a pragmatic enrichment. Thanks to this prior knowledge, in fact, the reader is able to interpret the information given in a certain chronological order. Also in this case, it is possible to draw a comparison between visual narratives and verbal sentences. When someone says “Julia took off her shoes and jumped on the bed”, the listener may infer that the second action chronologically followed the first one.

In conclusion, pragmatic enrichments contribute to the interpretation of a visual narrative in two ways: by picking one of the possible worlds unlocked by the semantic interpretation of the picture and by interpreting visual narratives in a chronological order. Thus, according to Abusch (2015), a visual narrative can be recursively interpreted through an invariant temporal progression model, which states that a picture sequence \sum is true in a world w and a time t' if and only if there is a t such that the following conditions are satisfied:

- i) \sum is true in w and t ;
- ii) P is true as a picture in w and t' ;
- iii) $t < t'$.

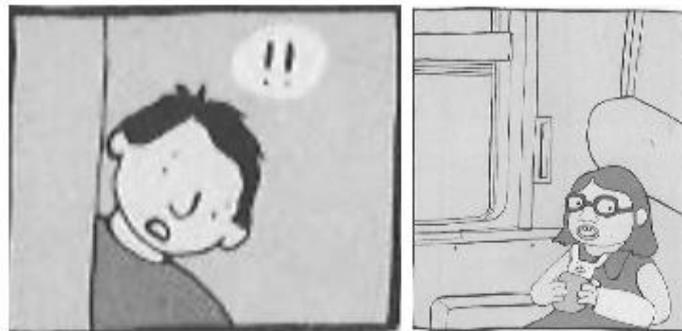
This definition posits the temporal order's importance in the interpretation of visual narratives such as comics. In the next sections, it will be clear that this concept may well be applied to movies too.

2. Discourse relations in movies

As Terrone (2015) highlights, films do not only exemplify colors and forms, but also spatial and temporal relations. These elements are conveyed to the viewer through the use of some discourse relations that are common to natural language and other visual narratives.

2.1 Free perception

We talk about free perception when we deal with pictorial narratives, which show in the starting depiction an individual who looks at something that will be shown in the following depiction. Free perception is a technique that owes its interpretative power to the assumption that an implicit index binds the two frames so that the second one could be interpreted based on the first one. For this reason, free perception is assumed implicitly anaphoric. Free perception can be found in pictorial narratives such as in the case of comics and movies. Figure (4) shows a scene from *Fluffy* by Simon Lia.



(Fig. 4)

(Abusch and Root 2017)



(Fig. 5)

(Abusch and Root 2017)

In (4) we see Michael, the main character of the graphic novel, who is looking for his lost rabbit. He enters a compartment of the train he is traveling on and sees a woman eating a rabbit like a sandwich. The reader is coerced to interpret the sequence as follows: the woman eating a rabbit shown in the second frame is what Michael is looking at in the first frame. Figure (5) is a scene from *The Third man* by Carol Reed (1950) and it is an example of free perception in cinema. In (5) we see two characters. The one on the left is looking at something while, in the next frame, we are shown what he was looking at: a woman.

Free perception, though, is not a phenomenon exclusively related to pictorial narratives. In fact, it may also be found in language, as in the case of eventive sentences, where someone is described to be looking at something and, immediately after, it is told what the person is looking at, like in the following examples:

1. I have heard a noise, I have looked back to see if someone was following me, and it was just a stray dog!

2. When I looked up, I have noticed beautiful jasmines growing all over a balcony. The interpretative power of free perception does not straightforwardly lead to a unique interpretation. In fact, as Abusch and Root (2017) point out, there are two types of scenes that can follow the anaphoric frame: non-veridical and veridical scenes. In the case of non-veridical scenes, they are crucially characterized by the fact that “the base world timeline does not satisfy the content of free perception panel” (Abusch and Root, 2017: p.3). The basic semantics is such that for any world w and viewpoint v that satisfies the narrative, the world w looks like the second panel q from the viewpoint v . In letters:

$$\pi(w, v, l, M) = q \quad (3)$$

Where l and M are projective parameters². When we look at Figure (4) and we notice that the world w that satisfies the narrative does not look like q , we understand that Michael is hallucinating. On the contrary, when we analyze a veridical free perception sequence of panels, q coincides with the world w that satisfies the narrative. In Figure (5), the third panel represents a q that is in line with the world w that satisfies the narrative for the particular viewpoint v of the man looking in the second panel.

The presence of a discourse referent in the set-up picture p plays a key role in indexing the interpretation of the subsequent picture q to the world w and viewpoint v of the looking subject depicted in picture p . This idea is consistent with what Braningan (1984) calls “structuralist poetics”, where the point of view is assumed to be a property of a language system within a general theory of representation. In this light, the camera is not an object but a construct of the viewer that allows him/her to make coherent spatial sense of the discontinuous stimuli of rapidly changing angles (Tomasulo 1987).

In section 4, I show that this kind of compositional interpretation of visual narratives can be extensively applied to films, also when the content of some scenes is non-veridical or when a sequence of scenes happen in the same time t .

2.2 Discourse coherence

The discourse coherence theory relates to studies conducted in linguistics by Hobbs (1985), Mann and Thompson (1988), Kehler (2002), Asher and Lascarides (2003), Lascarides & Stone (2009), and accounts for the fact that sentences are connected in a discourse by coherence relations. Many scholars, such as Kraft (1991), Bateman (2010) and Cumming *et al.* (2017) argue that a similar framework can be applied also to films. In fact, since films represent a world, being the actual or a possible world, what allows the fruition of a movie is coherence.

Kraft (1991) empirically demonstrates that there are some cinematography techniques, namely the establishing shot and the directional continuity, which are employed to coherently present a sequence of non-overlapping shots. The establishing shot shows an

² l represents the projection lines in terms of v , M is a marking rule that qualitatively distinguish the projected elements.

overview of the spatial location of the depicted events, and it is usually the first frame of a sequence. Directional continuity, on the other hand, is often used as a convention for filmmakers and imposes to draw an invisible line along the screen which should not be crossed by the camera (180° rule), so that the viewer is not confused by the different angles and perspectives of the shots. The results of some experiments by Kraft (1991) show that the use of these techniques facilitate the correct interpretation of a filmic sequence and their omission drastically hinders the comprehension of the sequence.

Also, Cumming *et al.* (2017) consider the consistency of the viewpoint as an interpretative convention that allows the fruition of a film. In addition, they define two typologies of viewpoint constraints that act as semantic conventions: the X-Constraint and the T-Constraint. While the X-Constraint corresponds to the 180° rule cited above, the T-Constraint posits that the viewpoints of two consecutive scenes must be related by a translation parallel to the action line, without rotation. Crucially, these conventions cannot be compared to lexical conventions in language, but they could be considered more like “inter-sentential semantic relations which seem to organize all form of linguistic discourse” (Cumming *et al.* 2017: 21). Unlike natural language, filmic materials generate expectations, are conventionalized, motivated and violable.

Even though viewpoints constraints can be understood in terms of discourse coherence relations (providing the film with the same structure of verbal discourse), the voluntary violations of filmic conventions by author can trigger a possible interpretation of the filmic content (Bateman 2016). In this regard, Branigan (1984) warns that the “author” should be seen not as a person, but as a function constructed by the reader to make texts intelligible. In the same fashion as the communicative aim of the comics’ author *A* in Abusch (2013), we could imagine that an artist may play with the discrepancies that could be displayed by violating conventions, hence the expectations deriving from the viewers experience and world knowledge, to create a different effect, perhaps to surprise the viewer or to scare her.

3. Cinematography techniques and their interpretative power

The most recent developments in the literature provide the basis to investigate how cinema, intended as a non-standard linguistic object, can replicate the coherence rules of the verbal discourse (Kraft, 1991) and the inferential typologies of natural language (Abusch and Root, 2017). Nonetheless, it has also been demonstrated (Bateman, 2016) that, unlike language, the communicative aim of the artist behind the movie may act in violation of some semantic or pragmatic rules to trigger a further interpretation of the filmic content. In this section, the mutable relation between natural language and cinema is explored throughout some examples of cinematography techniques that replicate (or not) the way in which a verbal discourse produces meaning.

3.1 Spatial continuity in movies

Bordwell argues that, despite significant stylistic changes have been introduced in modern filmmaking (i.e., special effects, rapid editing), today’s films are far from rejecting traditional continuity. Conversely, the new filmic styles «amount to an intensification of established techniques» that he refers to as «intensified continuity» (Bordwell 2002: 16). A possible way to film a spatially coherent scene is what we called before the “establishing shot” by Kraft (1991), which may also be known as analytical editing. This technique provides a wide shot that allows the viewer to see all relations involved in the picture, and successively single shots of subparts of the wider shot. In this case, the viewer is already oriented when being presented with the subparts of the

establishing shot. Another possible way is represented by constructing editing, where the scene is constructed piece by piece with non-overlapping shots. This last process is possible because the way in which we perceive a pictorial narrative is more similar to the way we usually perceive the world surrounding us than we may tend to intuitively expect (Berliner and Cohen, 2011). In other words, our brain fills up the missing information between filmic frames just like it generally processes visual information in the real world. For this reason, we can conjunct different shots in a coherent location. In sequences like (6), constructive editing helps our unconscious mind to fill the gap and to build up a perfectly coherent scene.

a)



b)



c)



(Fig. 6)

Call me by your name, L. Guadagnino (2017)

Frame (6.a) displays a man with a white shirt coming down the stairs of a house he does not know well, trying to find his way to reach the people that are waiting for him in the terrace. The viewer “follows” the man through a door between frame (6.a) and (6.b) even though the passage from a room to another is not shown. The viewer is also coerced to interpret that the man arrives at the terrace (6.c) after moving out from the second room (6.b). This inferential mechanism applies to spatial continuity just like the inference ensuing from the sentence “Julia took off her shoes and jumped on the bed” applies to temporal succession. The two actions are in continuity with each other.

3.2 Non-veridical frames and stative sequences in movies

Another discourse relation that can be reproduced by cinematography techniques is free perception. Whereas the case of veridical frames in movies has already been investigated by the literature (Abusch and Root 2017), the case of non-veridical frames and stative sequences have received little attention. Branigan (1984) calls them “mental process narration” but, in his theory, they are just a product of the rational subjectivity of the viewer, who passively receives «the level of narration where the telling is attributed to a character» (Branigan 1984: 310). I depart from this literary interpretation, by justifying the interpretation of non-veridical frames using discourse relations. Non-veridical frames are frames q where the world w that satisfies the narrative does not look like q (see equation (3)). These are intentionally produced to trigger the interpretation that the character depicted in the anaphoric frame is hallucinating (or dreaming or similar). Their functioning is similar to the eventive sentences’ one and can be found in comics as well as in movies (Abusch and Root 2017). For example, in (7) we see that the main character (Lester, performed by K. Spacey) is seated at the basketball game with his wife, while attending to their daughter’s cheerleader performance. The frame (7.a) illustrates the show of the dancers, among which there is a beautiful young blond girl, who Lester is attracted to.

a)



b)



c)



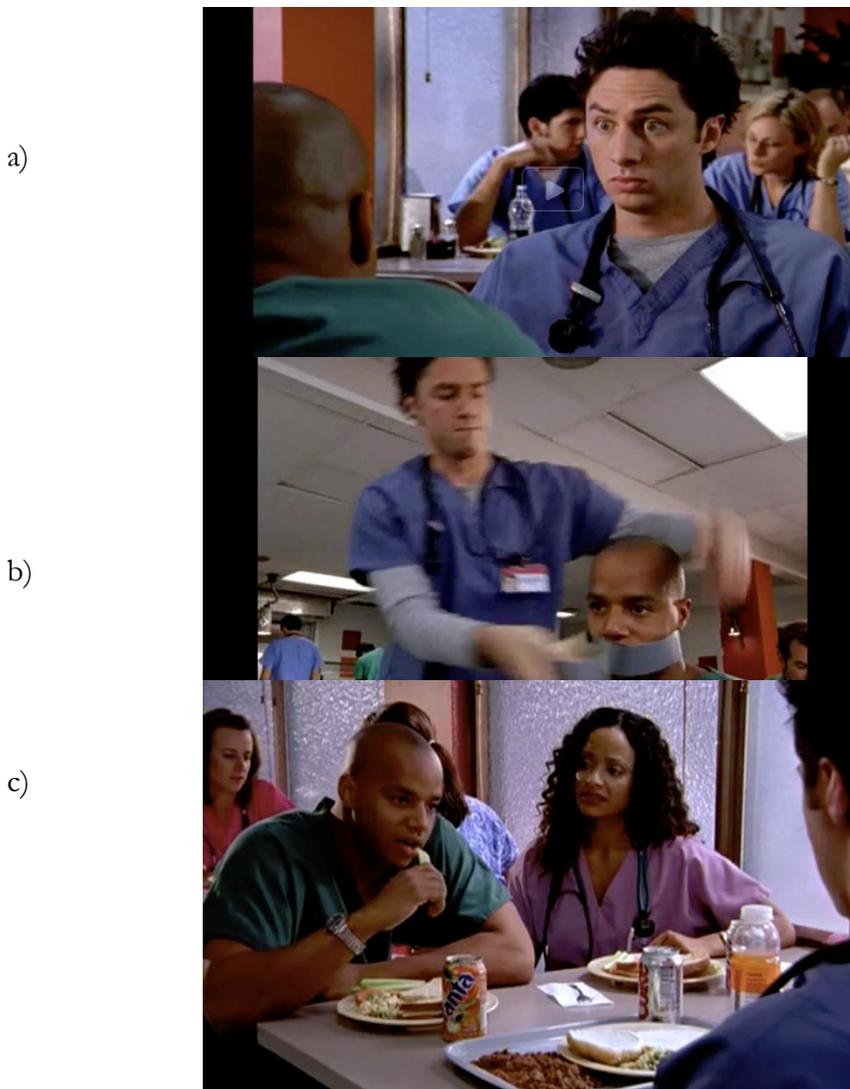
d)



(Fig. 7)

American Beauty, S. Mendes (1999)

Frame (7.b) is an anaphoric close-up of Lester's gaze, and frame (7.c) shows to the viewer what Lester is seeing. All the dancers but the blond one disappeared, and the spotlight oddly focuses on her. Clearly, the frame (7.c) is not satisfied by the narrative world of the previous frames and coerces the viewer to interpret that Lester is hallucinating. For a further proof, frame (7.d) depicts a scene that is satisfied with the world of the narration and ends the hallucination. In this example, the time progression is never interrupted. The hallucination in (7.c) chronologically follows the start of the performance in (7.a) and precedes the end of the performance in (7.d). Not every imaginative sequence is structured like this. In cinema, there are examples of non-veridical sequences that do not follow a constant temporal succession. We consider, for instance, sequence (8).



Scrubs, (Ep.04, S.2), B. Lawrence (2002)

Frame (8.a) is the anaphoric close-up of the main character's gaze (JD performed by Z. Braff). Frame (8.b) shows what JD is perceiving through his eyes, while frame (8.c) shows a state of the world that is contemporaneous to (8.b) but outside JD's visual

spectrum. This temporal inconsistency is what helps to understand that the world depicted in (8.a) and (8.c) does not satisfy frame (8.b), triggering the interpretation of the non-veridical frame as such.

Example (8) is a case in which the temporal succession in cinema is interrupted for interpretative reasons. We refer to this property as “stativity”. Just like stative verbs exist in natural language, stativity is found in comics when two consecutive pictures ($p_1 \neq p_2$) share their temporal coordinates ($t_1 = t_2$) (Abusch 2015). I argue that stativity holds in movies too.



(Fig. 9)

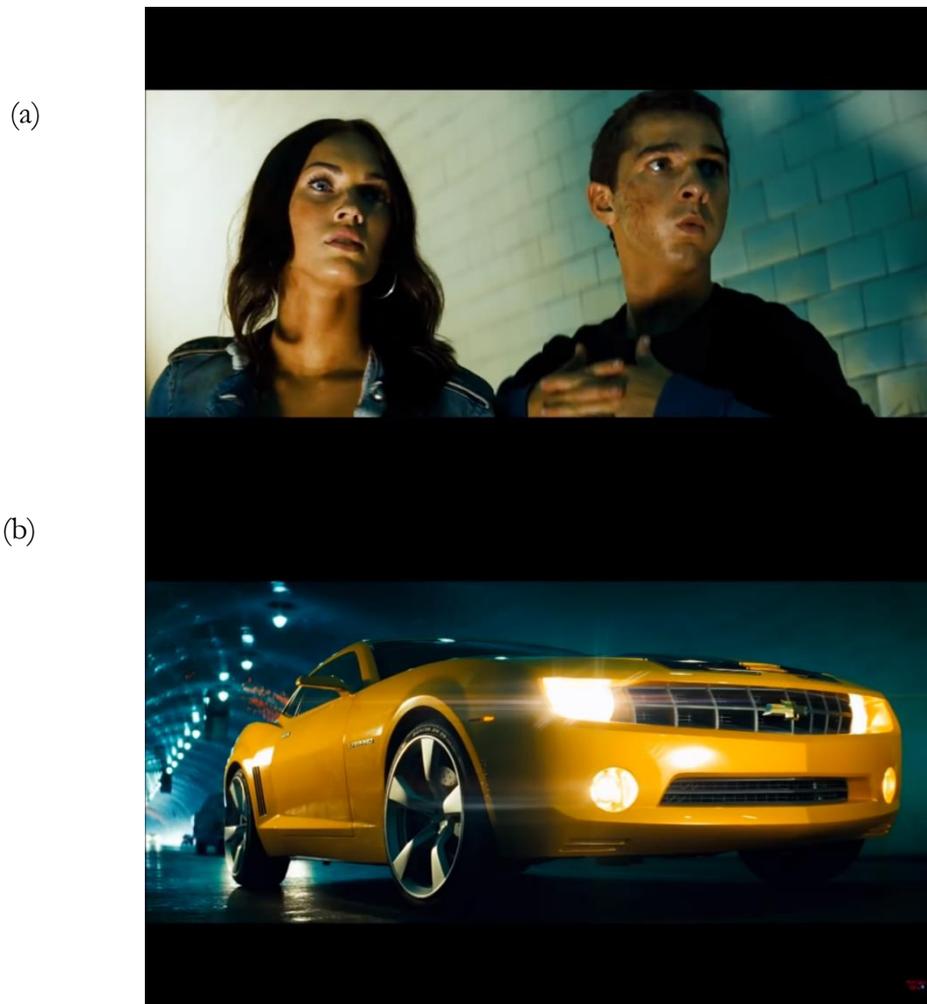
Snowpiercer, Bong Joon-ho (2013)

In sequence (9), the first shot (9.a) depicts the main character (Curtis, performed by C. Evans) who is forced to make a very important decision: choosing between the fulfilment of his mission and the life of his dearest friend. In (9.a) we see Curtis looking at his left, while (9.b) shows what he was looking at: his friend Edgar (performed by J. Bell) whose life is threatened by the enemy. Frame (9.a) and frame (9.b) are different but happen at the same time. In this case, stativity is used by the movie maker to trigger a precise interpretation of the filmic content: indecision and uncertainty.

3.3 Multimodality and filmic supplements

If the previous techniques owed their interpretative power to the property of iconicity, some cinematography features enrich the meaning of the filmic content thanks to the multimodality of this medium. The interface between different media of representation has a strong influence in how the meaning is constructed. This is true for natural language, whose content can be enriched with gestures, but also for cinema, whose

content can be enriched by the juxtaposition of a musical soundtrack or a spoken dialogue. Surprisingly, in the literature, there have been no attempts to account for cinema's multimodality and its implications for the interpretation of the filmic content. An example of a filmic sequence whose content is enriched by the interface between what is shown and the music in the background is showed in sequence (10).



(Fig. 10)

Transformers, M. Bay (2007)

In (10.a) the two main characters (Sam and Mikaela, performed by S. La Beouf and M. Fox), who were desperately looking for their car, finally find it. Frame (10.b) shows the car. This sequence is stative however, due to the presence of an uplifting musical motif playing in the background, not only the interpretation is temporally fluid, but it is also charged with a positive emotion (surprise).

Multimodality in cinema is also used to replicate some inferential typologies of natural language. There are cases, in fact, where the filmic sequence shows exactly what the character is describing by voice. The juxtaposition between frames and spoken language creates an informative statement, called supplement, like what co-speech gestures do in word-statements (Tieu *et al.* 2018). Co-speech gestures co-occur with spoken words and enrich their meaning by displaying a three-dimensional shape that resembles the intended denotations. In the same fashion, some filmic sequences show on the screen what the character is describing with words. In sequence (11), a chef (Master, performed

by K. Kobayashi) is cooking a dish and, while he describes by words each step of the recipe, the viewer is shown the corresponding action of the preparation.

(a)



(b)



(c)



(Fig. 11)

Midnight Diner: Tokyo Stories, J. Matsuoka (2019)

Another, less trivial example of supplements in cinema is given by those sequences where the filmic content suggests a positive (or negative) emotion and its corresponding soundtrack suggests a negative (or positive) emotion. The interplay between these two juxtaposed but contrasting media sparks a sense of alienation in the viewer which is coerced to interpret the scene in an unusual way. For example, in sequence (12), the candid music of Gioacchino Rossini's *Gazza Ladra* is juxtaposed to an extremely violent filmic sequence where a character kills another one.



(Fig. 12)
A Clockwork Orange, S. Kubrick (1971)

In this case, the use of a filmic supplement is motivated by the communicative aim of the artist (the movie maker), who wants to convey a specific (although controversial) interpretation of the sequence: violence is fun. This rhetoric use of different media in cinema reveals the extent to which meaning can be created and enriched in cinema. Not only it is generated throughout the use of some editing techniques, but it is also enriched by the exploitation of cinema's multimodality and by the intentions of the movie maker.

3.4 Violations of filmic conventions

Language and movies share the same properties (iconicity and multimodality), follow the same rules of discourse coherence and generate similar inferential typologies (supplements). However, filmic materials, unlike natural language, generate expectations, are conventionalized, motivated and violable (Cumming *et al.* 2017). The violation of filmic convention is the ultimate kind of enrichment that originates from the intentions of the movie maker. Often, the movie maker contravenes well-established semantic convention with the only purpose of coercing a specific interpretation.

In the case of the X-Constraint a violation would lead to confusion, or it could be also used to generate a particular effect of the viewer, or again, it could be used by the director to suggest and remark something about the story that is told.



(Fig. 13)

Arizona Dream, E. Kusturika (1998)

The narrative in sequence (13) depicts a young man (Axel, performed by J. Depp) (13.a) assisting his dying uncle (Leo, performed by J. Lewis) (13.b). In (13.a) the nephew directs his glance diagonally towards the viewers' left and slightly downward. Surprisingly, in (13.b), we reconstruct that he is not seated on the right of the dying character, as suggested by the first frame, but to the left, as we can clearly notice his hand on the dying character's shoulder. The violation of the X-Constraint is voluntary and adds a sense of confusion to the scene. Another example is shown in sequence (14):



(Fig. 14)

Batman, The Dark Knight, C. Nolan (2008)

We see from (14) that the 180° rule has been violated. The faces of both characters are filmed from the same point of view, even though the two characters are one in front of the other. It is a commonly held belief that this violation of the X-Constraint by the movie maker is intentional and leads to a specific interpretation not only of this specific sequence but of the whole movie. The face of the villain (Joker, performed by H. Ledger) in frame (14.a) seems to mirror the face of the hero (Batman, performed by C. Bale) in frame (14.b), in such a constructed way that seems to suggest that the two characters may share the same identity. Considering the content of the movie, in which the hero is never completely accepted as a positive character, we can see that with this violation, the movie maker manages to enrich the meaning of the filmic material by underlining the conflict of the main character.

4. Conclusions

In this paper, I discuss the idea that both linguistic categories and unique features may be found in films. Like language, filmic materials express their meaning through their iconicity and can enrich their possible interpretations thanks to their multimodality. Movies are interpreted with the use of discourse relations, like coreference and temporal succession, and replicate some inferential typologies of natural language (namely, supplements). However, filmic content is also subject to semantic and pragmatic conventions that, like language, can be subverted or violated. Instead of failing, the communicative aim of the movie maker is empowered by the sapient use of these violations, which can be exploited to add something more to the final interpretation of the filmic content.

The paper draws from the literature in cognitive linguistics and the theory of cinema and contributes to both. Firstly, it extends the application of Super Semantics to cinema, applying the concepts of iconicity and multimodality to films, and drawing a parallel between natural language's inferential typologies and those of filmmaking. Secondly, I expand Bordwell's critique of contemporary films by updating the list of classical editing techniques that intensify the narrative continuity and by stating that these techniques have also semantic value. Lastly, I depart from Branigan's subjective interpretation of films, by relying on linguistic rules that fix the interpretation of a sequence in a formal way.

References

- Abusch, Dorit (2013), «Applying discourse semantics and pragmatics to co-reference in picture sequences», in *Proceedings of Sinn und Bedeutung*, vol.17, n. 1, pp. 19-25.
- Abusch, Dorit (2015), *Possible Worlds Semantics for Pictures*, Manuscript, Cornell University, 2015.
- Abusch, Dorit, & Root, Mats (2017), *The formal semantics of free perception in pictorial narratives*, in Alexandre Cremers, Thom van Gessel & Floris Roelofsen (eds), *Proceedings of the 21st Amsterdam Colloquium*, Available online at <https://semanticsarchive.net/Archive/jZiM2FhZ/AC2017-Proceedings.pdf>.

- Adamson, Lauren B. (1996), *Communication development during infancy*, Boulder, 2019.
- Bateman, Joh A., & Schmidt, Karl H. (2012), *Multimodal Film Analysis: How films mean*. London 2011.
- Berliner, Todd and Dale, J. Cohen (2011), «The Illusion of Continuity: Active Perception and the Classical Editing System», in *Journal of Film and Video*, vol. 63 n. 1, 2011, pp. 44-63.
- Bordwell, David (2002), «Intensified Continuity Visual Style in Contemporary American Film», in *Film Quarterly*, 1 March 2002; 55 (3): pp. 16–28.
- Braningan, Edward R. (1984), *Point of View in the Cinema: a Theory of Narration and Subjectivity in Classical Film*, Berlin 2010.
- Chomsky, Noam (1957), *Syntactic Structures*, Paris, 2022.
- Cumming, Samuel, Greenberg, Gabriel, Kelly, Rory (2017), «Conventions of viewpoint coherence in film», in *Philosophers' Imprint*, vol. 17.
- Fitch, W. Tecumseh and Hauser, Mark D. (2004), «Computational constraints on syntactic processing in a nonhuman primate», in *Science*, vol. 303, n. 5656, pp. 377-380.
- Greenberg, Gabriel(2011), «The semiotic spectrum» in *Graduate School - New Brunswick Electronic Theses and Dissertations*, from <https://rucore.libraries.rutgers.edu/rutgers-lib/36033/#citation-export>
- Greenberg, Gabirel. (2018), «Content and Target in Pictorial Representation», in *Ergo: An Open Access Journal of Philosophy*, vol. 5, n. 30.
- Guasti, Maria Teresa (2016), *Language Acquisition*, Cambridge, 2016.
- Heim, Irene and Kratzer, Angelika (1998), *Semantics in Generative Grammar*, Malden, 1998.
- Katz, Jonag and Pesetsky, David (2011), «The identity thesis for language and music» from <http://ling.auf.net/lingBuzz/000959>.
- Kraft, Robert N.; Cantor, Phillip; Gottdiener, Charles (1991), «The Coherence of Visual Narratives», in *Communication Research*, vol. 18, n. 5, pp. 601-616.
- Kravanja, Peter and Coëgnarts, Maarten (2014), «Metaphor, bodily meaning, and cinema», in *Image and narrative*, vol. 15, pp. 1-4.
- Lakoff, George and Johnson, Mark. (1980), *Metaphors we live by*, University of Chicago Press, 2nd ed. 2003.
- Lewis, David (1970), «General Semantics», in *Synthese*, vol. 22, no. 1/2, 1970, pp. 18–67.

Metz, Christian. (1964), «Lecine□ ma: Langue ou langage?», in *Communications*, vol. 4, pp. 52-90

Morris, Charles (1938), «Foundations of the Theory of Signs», in *International encyclopedia of unified science*, vol.1, pp. 1-59.

Patel-Grosz, Pritty, Grosz, Patrick G., Kelkar, Tejaswinee, Jensenius, Alexander R. (2018), «Coreference and disjoint reference in the semantics of narrative dance», in *Proceedings of Sinn und Bedeutung*, vol 22, n. 2.

Potts, Chirstopher (2005), «The Logic of Conventional Implicatures», in *Journal of Linguistics*, vol 42, n. 2, pp. 490-495.

Potts, Christopher (2004), *The Logic of Conventional Implicatures*, Oxford Scholarship Online, 2007.

Rebuschat, Patrick, Rohmeier, Martin, Hawkins, John A. and Cross, Ian (2011), *Language and Music as Cognitive Systems*, Oxford Scholarship Online.

Pistoia-Reda, Salvatore (2020), «Una nota sulla pragmatica musicale», in *de Musica*, vol. 1, n. 24, pp. 173-178.

Schlenker, Philippe (2019), «What is Super Semantics?», in *Philosophical Perspectives*, vol. 32, n. 1, pp. 365-453.

Schlenker, Philippe, Chemla, Emmanuel, Schel, Anne M., Fuller, James, Gautier, Jean-Pierre, Kuhn, Jeremy, Veselinović, Dunja, Arnold, Kate, Căsar, Cristiane, Keenan, Sumir, Lemasson, Alban, Ouattara, Karim, Ryder, Robin and Zuberbühler, Klaus (2016), «Formal monkey linguistics», in *Theoretical Linguistics*, vol. 42, n. 1-2, 2016, pp. 1-90.

Stalnaker, Robert (1998), «On the Representation of Context», in *Journal of Logic, Language and Information*, vol. 7, pp. 3-19.

Tieu, Lyn, Pasternak, Robert; Schlenker, Philippe and Chemla, Emmanuel (2018), «Co-speech gesture projection: Evidence from inferential judgments», in *Glossa: a journal of general linguistics*, vol. 3, n. 1, pp. 109.

Tomasulo, Frank P. (1987), «Review of "Point of View in the Cinema: A Theory of Narration and Subjectivity in Classical Film"», in *The Journal of Aesthetics and Art Criticism*, v. 45, n. 3, pp. 309-312.