

# From Script to Screen: A Multimodal Analysis of Films

Under the supervision of

Amira Agameya  
Faculty of Arts – Cairo University

Amir Bolis

## Abstract

A film is the final product of a long process of screenplay writing, selection, and direction. This paper investigates the meaning-making resources of films as a multimodal composition, relying on two theories that incorporate the areas of linguistics and multimodality, namely: Reading Images and the Grammar of Visual Design and Systemic Functional Multimodal Analysis (SF-MDA). The sample under analysis was composed of ten screenplays and their films, spanning from 1990 to 2000, the era that is commonly known as modern cinema. A qualitative research design, analyzing selected film scenes and sequences, aimed at identifying the resources that make the filmic meaning, cohesion in films, and the degree of agreement between the film script as a mono-modal text and the film as a hybrid form of pictorial and auditory modes. Findings revealed that the film script does not account for everything that appears on the screen and that the meaning of a film is encoded in the screenplay and decoded in its visual and auditory resources of cinematography.

**Keywords:** reading images, multimodality, film, screenplay.

ملخص

يُعد الفيلم السينمائي نتاجاً لسلسلة من الخطوات الإنتاجية التي تتضمن أولاً الكتابة السينمائية للسيناريو، ثم الاختيار، وأخيراً الأخراج السينمائي. تهدف الدراسة الحالية الي التعرف علي مصادر صناعة المعني للفيلم السينمائي، معتمدة علي نظريتين تدمجان بين علم اللغة وعلم قراءة الأنماط المتعددة. تتكون العينة الحالية من عشر أفلام سينمائية، التي يتراوح أنتاجها من عام 1990 الي عام 2000، ماهو متعارفٌ عليه كعصر السينما الحديثة. تم إجراء تحليل نوعي لمجموعة من المشاهد والصور لدراسة ثلاثة محاور، ألا وهي مصادر صناعة المعني الفيلمي، وترابط النص الصوري، ومقدار التوافق بين السيناريو كنص لغوي والفيلم بكونه نتاجاً مركباً من أنماطٍ صورية وسمعية. أظهرت النتائج أن النص السينمائي لا يحوي كل ما يراه المشاهد علي الشاشة، وأن معني الفيلم السينمائي مُختزل في السيناريو ومُستعلن في الفيلم كنتاج نهائي.

**الكلمات المفتاحية:** قراءة الصور – تعددية الأنماط – الفيلم السينمائي – السيناريو.

Films are powerful tools for shaping people's thoughts and attitudes and are often described as even having the power to influence individual beliefs and values. Cape (2003) asserts the influence of films on both creating new thinking patterns in previously unexplored social phenomena, especially in children, as well as their ability to influence an individual's existing social boundaries based on what is shown on screen as the "norm". Wedding and Boyd (1999) state that films reflect existing cultural values of the society because of their ability to influence individual beliefs, especially through the dialogical sub-genre of films.

The study conducted a qualitative multimodal analysis of selected scenes from 10 screenplays. The multimodal analysis aims at identifying the meaning-making resources of films, the way in which the textual is translated into the visual, and the cinematographic techniques film directors make use of in order to realize the intended meaning of a film in light of its screenplay.

To do this, the study attempted to answer one research question that has not received much attention in film studies from a linguistic point of view: *How do the cinematographic resources of filmmaking contribute to the meaning-making of the narrative?* Cinematographic resources refer to film production and direction techniques such as picture, camera, lighting, color, the score (film soundtrack), framing, subject's composition, frame rate, and shot size.

The analysis is not in isolation from what is known as the screenplay (a monomodal text written for the purpose of the cinema). A screenplay as a genre is one kind of text written for the specific purpose of film production, in a formulaic page-structure, built up of separate shots, sequences and scenes.

Kress and van Leeuwen (2006) and O'Halloran (2008) regard the multimodal world as a world of semiotics. Both attempted to devise an intermediate ground between the visual semiotics of pictures, icons, portraits and drawings and the domain of language. Their theories were based on Halliday's Functional Grammar. Films, likewise, are visual stories with a wide range of semiotic signs. They share static portraits and still photography their semiotic landscape and differ in terms of dynamicity.

One of the challenges of linguistic analysis of multimodal compositions is the impracticality of presenting visual moving elements, such as videos, on paper and the arduous task of collecting a database comprising both the visual and the textual in one place. In normal circumstances, such a multimodal linguistic analysis may rather rely on static pictures and shots in order to account for as much accuracy as possible (Bednarek, 2015; Bo, 2018; Rohrbach et al., 2015). In the presentation of textual-multimodal data, an analyst can utilize graphs and pictures that accompany their textual parallels. Kress and Van Leeuwen (2006) provide this possibility, deploying Hallidayan Grammar to offer accurate reading of images and multimodal texts.

Second, the importance of the present study lies in the area of multimodality, especially since the Systemic Functional-multimodal (SF-MDA) approach is a well-established area by Kress and Van Leeuwen (2006) and O'Halloran (2008), who present this approach as an effective tool that is ready to accommodate the various types of multimodal discourses.

Films are dynamic pictures that combine many modes of communication, which run on screen simultaneously. Viewers of a film may engage with the narrative through the pictures (as was the case in classic cinema before the addition of sound and speech to the scenes), speech, sound effects, lighting, costumes, and other cinematographic effects.

The world of cinema is multimodal in nature, and the two most important modes of meaning-making in films are the visual mode of the representation of subjects, worlds, timing, actions and events, and the auditory mode that accompanies the visual composition.

The essence of filmmaking is *mise-en-scene* (Bordwell & Thompson, 2008), a term coined by the famous film theorist Andre Bazin (Bazin, 1967), which refers to everything the viewer can see and hear in a single shot. *Mise-en-scene* comprises all the cinematographic elements that create a film shot. These elements are characters, props, set design, costumes, makeup, color, lighting, blocking, frame rate, sound and music. These are not arbitrary constituents of the shot; they are elements that construe all the possible meaning-making resources of a shot in a film.

## **2. Literature Review**

A few theories deal with the cinema as an art without engaging directly with the genre of screenplay as the essence of filmmaking. These include the famous essays of Andre Bazin in a book entitled *What is Cinema*. Bazin tends to offer an approach to the meaning of films and dynamic pictorial compositions, comparing it to other forms of art such as photography, painting, and theatre.

Another theory on cinema is provided by Metz (1992), arguing that a film is a form of language but not a language system. While film makes meanings as a language of visual composition, it is not a language system in the sense that we may view a word as a picture, a sentence as a scene, or a number of sentences as a sequence. There is a line that divides the system of language and that of the cinema. The cinema is seen as rather a non-system language; an image is always speech, but not a unit of language.

### **2.1. A Theory of Multimodality from SFG**

The accounts the present study relies on are two highly related theories of linguistics, one that takes text and discourse as its field of study (SFG) and another that takes images and visual designs as its basis (Multimodality). Kress and Van Leeuwen (2006) introduced a theory that dates back to 1996, one that attempts a

reading of images, figures, maps, charts, and pictorial elements based on functional linguistics. An analyst may look at a picture the same way they may look at a text. For example, the system of color in a picture is the same as modality in a text. While *perhaps*, *certainly*, *more likely*, and *on an evidential basis* all construe a degree of evidentiality in a text, color (being saturated, differentiated, or modulated), contextualization, representation, depth, illumination, and brightness are all modality markers in a pictorial composition.

As the clause has participants, a picture also has participants. As the informational unit in a clause can determine salience of linguistic units, participants in a pictorial product have variable degrees of saliency. As the clause unfolds through a process, a pictorial product construes processes and actions. Both texts and visual compositions have cohesive ties that ensure their unity. Thus, instead of pictures, the multimodal analysis looks at videos (moving pictures). But Kress and Van Leeuwen's theory, however, does not consider the analysis of such moving elements. This, in itself, is a gap in the theory that would require further investigation – what the present study attempts to fill.

## **2.2. Film and Multimodality**

Bednarek (2014) demonstrates how multimodality and corpus linguistics can be integrated for the search on films and TV series. Like O'Halloran (2008), Bednarek establishes the need to approach TV and film texts multidisciplinary, being by nature hybrid texts. The modalities included in these texts are both aural and visual, incorporating semiotic systems on the Aural level (musical scoring, sound effects, dialogue) and the Visual level (moving images that utilize angle, color, brightness, etc.; transitions such as cuts, fades, and wipes; written language, gestures, postures, facial expressions).

The main focus of the study (p. 68) was offering techniques for combining multimodal resources with corpus linguistics, insisting on the possibility of mining movie scripts for visual information. One of these possibilities is considering the effect of synchronization on dubbing, of course through a rather qualitative analysis of particular segments/scenes of a movie. "Synchronization on dubbing" here refers to the way professional filmmakers match the newly added dialogues, music or any other effects to the original soundtrack of a picture that has already been shot. This typicality of a written mode to its multimodal realization is considered by Bednarek (2010), though it is not limited to movie scripts, nor is it targeting particular segments of the designated scripts such as ACTION, DIALOGUE, TRANSITIONS, etc. (for more details about the way monomodal corpora of TV and film narratives can be created see Bednarek (2015: p. 69-72). Finally, the gap the present study allocates is that Bednarek's study focuses on TV series rather than films, in addition to the analysis of the DIALOGUE alone, which is built up from fan transcripts and not based on original texts. Bednarek states that fan transcripts

may not be accurate data for multimodal analysis of TV and films because they only provide salient information that is important only according to the fan transcriber.

An account that is highly relevant to the current literature is Bateman and Schmidt's (2012) book *Multimodal Film Analysis: How Films Mean*, where they rely on the Systemic Functional Grammar theory, in addition to Kress and Van Leeuwen's (2006) grammar of visual design. Bateman and Schmidt view films as a sequence of moving images, which, integrated and constructed in certain ways, guide the viewer to specific interpretations. They establish the idea that there is information beyond the referential story events in films, focusing on the textual strands drawn from Halliday and Matthiessen (20014). Textual consideration of films has little attention from a linguistic stand point and those prior studies targeted rather the Interpersonal and Ideational meanings of films. One of their approaches, for example, adopts the zooming technique in films and how zooming combines shots together to make single statements, which, ultimately, make a film parallel to a text, where segments combine to form a cohesive 'whole'. Information in films, thus, should arise from the larger whole, where decontextualizing of the shots is not possible.

Bateman and Schmidt further applied the concepts of Distance, Shot Range/type, Framing, Camera Directionality, and other visual strands to identify meaning on the Textual level. Their analysis was linear, in which three columns constituted their analytical layout: a column for camera shot, one for the image track, and one for the dialogue and soundtrack of the films. The analysis depended on the dialogic segments of films, with no consideration of their original screenplays. I argue that in order to reach a solid textual analysis of the visual and the verbal elements, a multimodal analysis would have supported their approach, since these scripts already contain meanings that are implicit to the viewer, such as camera angle, shot types, etc. – i.e. the screenplay contains in itself the kind of information needed to direct the analyst towards a refined visual reading.

Also for Bateman and Schmidt (ibid), films make signals through, for example, camera movement, but a static shot is also a signal – in this case, the lack of signals in films can make signals. This is why they assert the principle that there is no fixed meaning for camera angle, directionality, zooming etc. in films, rather films mean in their context. Generalizations for Bateman and Schmidt are thus implausible when it comes to reading the visual information stacked in films.

Another study that uses SFG in a multimodal analysis of films is Baumgarten (2008), where a corpus of popular motion pictures and their German dubbed versions are used to explore relations between the visual and the verbal information in films. This contrastive study between English and German explores the ways the two languages relate verbal with pictorial information. Baumgarten, like Bateman and Schmidt (2012), dealt with a film as a text, which has signals and sign systems, relying on Kress and Van Leeuwen's (2006) grammar of visual design. Echoing

Bateman and Schmidt, films for Baumgarten are integrated texts where the parts interact with the ‘whole’, each affecting one another in the formation of the whole/here. Arrangement of visual and verbal components is systemic, even the dialogue and the sound effects drive the information salient, and that there is a degree of convergence between the visual and the linguistic resources. Films necessarily have two modes: a layer of visual information and another layer of verbal information, with the addition of sound and sound effects, all combining to generate one unified text that unfolds across time and space.

Baumgarten introduces the concept of ‘visual-verbal cohesion’, where cohesion of this interplay of modes happens internally: i.e., modes are internally related. For the record, English and German deploy different linguistic information to realize this visual-verbal cohesion. The results regarding the contrastive dimension are beyond the scope of the present study. The current analysis, rather, draws from this the possibility of examining static shots, by decontextualizing them for a qualitative analysis and recontextualizing them within their normal background – in this case, within their texts/films as a unified unit that unfolds systematically. Any analyst should only consider which verbal elements are to combine with which visual elements when it comes to selecting segments for the multimodal analysis. There is no reference to the screenplays in Baumgarten’s study and, like Bateman and Schmidt (2012), considering the screenplays would have been insightful, especially when it comes to examining the correlations between the verbal and the visual.

No previous study has examined films as visual designs from a linguistic perspective, and being at the heart of the process of film semiotics, the present study locates this as a gap worthy of filling.

### **3. Methodology**

As the technology of filmmaking advances, the elements of mise-en-scene receives modifications and, when necessary, changes, to the creation of a single film shot. For this, the present study mobilizes the concepts provided by StudioBinder (an online website that includes almost everything filmmakers may utilize to create, edit and modify their shots <https://www.studiobinder.com/>). The resources provided by StudioBinder are used in the current study as a reference to *shot size, framing, camera angle, depth of field, camera movement, frame rate, editing transitions, lighting, and composition*. The current study investigates how the textual resources of film scripts are translated into shots and scenes in light of the concept of mise-en-scene, along with the nine elements provided by StudioBinder online resources.

Through the investigation of the textual resources analyzed according to the Systemic Functional theory, together with the above-detailed filmmaking resources, the study may provide an understanding of how the mono-modal domain of film scripts is transformed into the multimodal world of film. This methodology should

target the questions of the degree of consensus between the textual and the visual (the mono-modal and the multimodal) and to what extent the mono-modal resources of a film script aid, shape, and determine how visual and auditory resources are to appear on the screen.

The study undertook a qualitative research design for the analysis of pictures, scenes, and sequences that were retrieved from the selected films.

### **3.1. Data Analysis**

The following sub-sections present the multimodal analysis of selected scenes extracted from the current sample. Through an analysis of a two-minute sequence, the study examines the possible meanings encoded in a film as a multimodal medium of representation, with the focus being directed towards *text, angle, movement, color, lighting, tone, and the score*. Because a film is a unified medium of visual and auditory representation, this section also presents an analysis of *cohesion* in films.











#### **3.1.1. Analysis of a Two-minute sequence**

*The Shawshank Redemption* is a 1994 drama, written and directed by Frank Darabont. The film is based on Stephen King's novella *Rita Hayworth and Shawshank Redemption* (1982). The section under analysis relates to the film's subplot of Brooks (starring James Whitmore), an institutionalized inmate who had spent fifty years in the Shawshank State Penitentiary before his parole.

There are five components under analysis (Table 3.1), namely: picture, action, dialogue, score and sound effects. The score is another term for the soundtrack in film. Action refers to the ACTION segments of the film script. Finally, dialogue is the speech of characters whether present in the film script or not – sometimes we may find pieces of dialogue that are not present in the film script and only present in the film itself. The red bold line in the picture column marks the end of one scene and the point of departure for the next. Finally the ~ sign refers to absence of elements, whether action, dialogue, or score and sound effects. The other two elements (picture and cinematography) are always present.










Table 3.1

*A Multimodal Analysis of a Two-minute Sequence (The Shawshank Redemption, 1994)*

Picture	Action	Dialogue	Score and sound effects	Cinematography
	He tosses Jake through the bars.	~	Scratches of the crow (Jake)	Close shot; low angle Natural lighting falling on subjects Slight tilt to the left
	The crow flaps away.	~	Jake crows as it flaps away	Close shot; low angle Static
	~	~	Low sighs and faint music	Close shot; eye level; Natural lighting on subject's face Zoom in to... Slight tilt to the right
	~	~	Continued	Close shot; eye level; static Subject splitting the background
	~	~	Continued	Medium close-up shot Shoulder-level shot Static
	TWO SHORT SIREN BLASTS herald the opening of the gate	LEFT GUARD Goodbye brooksi	Soft music, escalating	Wide shot; eye level Static
	It swings hugely open, revealing Brooks standing in his cheap suit,	~	Continued	Wide shot; eye level static
	wearing a cheap hat. Brooks walks out...	~	Continued	Wide shot; eye level Static
	~	~	Continued	Wide shot; eye level Static
	~	~	Continued	Wide shot; eye level Zoom in;



From Script to Screen: A Multimodal Analysis of Films

Picture	Action	Dialogue	Score and sound effects	Cinematography
	~	~	Piano plays	Full shot; hip level; Zoom in;
	Brooks is riding the bus, clutching the seat before him, gripped by terror of speed and motion.	~	Continued; Inside bus rattling	Medium shot; eye level; zoom-in;
	~	~	Continued Inside bus rattling	Medium close-up shot; zoom in
	~	~	Continued	Wide shot; eye level; static;
	~	~	Continued	Wide shot; eye level; static
	~	~	Continued	Wide shot; slight pan to the left;
	~	Brooks (V.O) Dear fellas.	Continued	Medium shot; eye level
	~	I Can't believe	Continued	Medium close-up shot; shoulder level;
	~	how fast things move on the outside	Continued	Medium close-up shot; shoulder level;

The sequence in Table 3.1 depicts one and a half minutes of screen time, a sequence of four scenes. The text below is the film script exactly as present in the shooting script of the film and it occupies three quarters of a page. There are four SLUGLINES heading each scene.

INT -- LIBRARY -- DAWN

Brooks stands on a chair, poised at the bars of a window, cradling Jake in his hands.

BROOKS

I can't take care of you no more.  
You go now. You're free.

He tosses Jake through the bars. The crow flaps away.

EXT -- SHAWSHANK PRISON -- MAIN GATE -- DAY (1954)

TWO SHORT SIREN BLASTS herald the opening of the gate. It swings hugely open, revealing Brooks standing in his cheap suit, carrying a cheap bag, wearing a cheap hat. Brooks walks out, tears streaming down his face. He looks back. Red, Andy, and others stand at the inner fence, seeing him off. The massive gate closes, wiping them from view.

INT -- BUS -- DAY (1954)

Brooks is riding the bus, clutching the seat before him, gripped by terror of speed and motion.

BROOKS (V.O.)

Dear fellas. I can't believe how fast things move on the outside.

EXT -- STREET -- PORTLAND, MAINE -- DAY (1954)

Brooks looks like a kid trying to cross the street without his parents. People and traffic a blur.

BROOKS

I saw an automobile once when I was young.  
Now they're everywhere .

*(The Shawshank Redemption, 1994)*

The setting is provided in the four slugline. Time is set between dawn and day. The four scenes take place in four different places: the library of Shawshank prison, the prison's main gate, bus, and street. Any shift in the setting marks a new scene. This is the first and main rule of scene breakdown in film scripts and all the information about the setting of a scene is condensed in the slugline: EXT or INT (outside or inside), exact place, and time of the day. The places that may be available in the slugline are names of countries, cities, towns, and villages, or exact buildings, buses, cars, etc., and any spatial environment that is separated from the outside by walls or any other confinements. Thus, the short form INT marks the inside of any possible spatial context, while EXT marks open areas such as streets, waters, and open yards.

Sometimes the calendar year is mentioned in the slugline as parenthetical information. There are three mentions of the same year (1954) in the above sequence. This mention of 1954 is tied to the sub-plot of Brooks, highlighting the fifty-year stretch of his imprisonment. Presence of elements on the script is not arbitrary, even when it is a mention of the year. When reading the film script, it becomes clear that the year 1954 is 19 times out of 22 related to the character of Brooks.

The amount of information condensed in the slugline shapes the entire scene and is usually considered the information that dictates lighting, costumes, and even camera angles. For example, in the above depiction, close-up and medium close-up shots are used in INT scenes (the library of the prison in scene 1 and the bus in scene 3).

The camera is a busy domain, representing busy visual compositions. Another example of this is using elements such as V.O. (standing for voice over), which allows dialogue and action to run together even when the character is doing something other than speaking. This makes it possible for us to see Brooks in Scenes 3 and 4 crossing the street and hear him talking without seeing him talking.

### **3.1.2. Text, Angle, and Movement**

A text is a mono-modal domain that makes meaning in a given context. A film, in contrast, is a multimodal configuration that carries meaning throughout a hybridity of modes that is made up of pictures, sounds, colors, lighting, movement, angles, and speech.

Certain words may be traced in the film script that suggest type of angle and type of camera movement. The analysis revealed a number of lexical items that relate to cinematography, some of which guide the reader to the type of camera angle and camera movement. These include certain technical words and phrases such as *pan*, *pan to*, *roll*, *close*, *close on*, *low angle*, *tight angle*, *angle on*, *angle shifts*, *angle shifts to*, *angle widens*, *angle slow push*, *angle reveals*, and *tilt*. In most cases these words and phrases appear in all CAPS.

Occurrence of these items clearly demonstrates a relation between the film script as a mon-modal work and its multimodal representation as film, but they do not account for all types of camera angles and camera movements. This is because their occurrence does not occupy much of the lexical space in the current corpus. In a corpus of 269,496 tokens compiled for the present study, the above specified words and phrases occur only 90 times combined. This implies that film scripts do not tend to accommodate all the cinematographic information that may guide the reader to camera angle and movement. The screenplay, therefore, does not carry all the information that relates to the technicality of the picture and cinematic organization of the film.

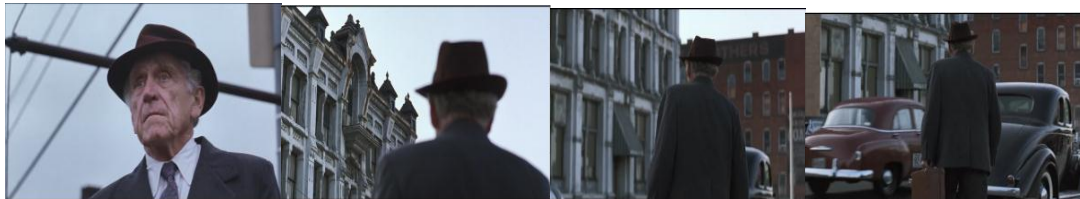
According to Kress and Van Leeuwen (2006), Bordwell and Thompson (2008), and the resources provided by StudioBinder, camera height or angles assign power relationships, either between subjects depicted in the frame or between represented subjects and the viewer. The specified sources confirm that a low angle suggests power of the represented subject. The analysis in Table 4.1, however, suggests that this view is not always the case in films. The represented subject in the above sequence is depicted as vulnerable. The third shot depicts Brooks crossing the street using low angles. The low angles are combined with a pan to the right and a tilt-up

(a pan is a camera movement on the horizontal axis, whereas a tilt is a movement on the vertical axis). The combination of the low angle with a simultaneous dual movement of a pan and a tilt creates uneasiness and heightens the character’s vulnerability.

This combination goes alongside Brook’s dialogue: *I saw an automobile once when I was a kid but now they are everywhere*, suggesting the character’s disorientation and inability to cope with the speed of the street. Below are the last four pictures of Scene 4, extracted from Table 3.1. But not all low angles are the same. In Picture 1, the low angle represents the front of the subject alone in the frame, with only street wires and a fraction of a building in the upper-right corner. Picture 2 represents the back of the character in relation to a building. The character’s dialogue gives a clue of this relation: *The World went and got itself in a big damn hurry*, where the character is being diminished gradually against the background of the building that mounts him.

Figure 3.1

*Angle and Movement as Meaning-making Resources of Film*



1 Low angle	2 Low angle	3 Low angle	4 Hip-level angle
Static	Tilt down	Tilt down	Tilt down + Pan left

This inter-sectioning of dialogue, camera angle, and camera movement are the resources that make meaning in films: the subject in the frame is overwhelmed by the size of the buildings and the speed of the street as he describes his situation to Andy and Red (fellow prisoners) with a voice-over monologue that is meant to be sent as a letter.

**3.1.3. Cohesion in Film**

In the same way a text is a unified unit of discourse, unified through what Halliday and Hassan (1985) call cohesive ties, a film is also a unified unit. In a hybrid visual and auditory composition such as a film, there are elements that bring together its discrete units. The current analysis identified eight basic elements that could establish filmic cohesion: subjects, type of shot, camera angle, camera movement, color, lighting, tone, and the score.

**3.1.3.1. Subjects**

There is no story without characters. They are the primary cohesive ties of any narrative. The sequence of shots in [Figure 3.2](#) is coherent, first, through the two

main characters (Glass and his son). Given Bordwell and Thompson's (2008) elements of mise-en-scene (setting, lighting, color, prop, costume and makeup, and staging), it appears that most of these mise-en-scene elements pertain to characters. The setting is where and when characters are depicted – a farmhouse in an Indian territory in the nineteenth century. The costumes and the makeup of characters tie the story to a particular setting. The colors of costumes and of the screen in general and the lighting of the scene bind the characters to certain modes and feelings, which in turn project the same feelings onto the viewer. For example, in Picture 15 the near monochromous colors suggest the state of desolation in the aftermath of burning Glass's farmhouse and the near loss of his child.

Unlike any other visual cohesive device, characters are equally present both in film script and the film itself. The question related to subjects as pictorial cohesive units is one that addresses the presence and absence of subjects. Equally important is the size of the shot within which the subjects in a frame are depicted and the camera angle and movement.

### *3.1.3.2. Type of Shot, Angle, and Movement*

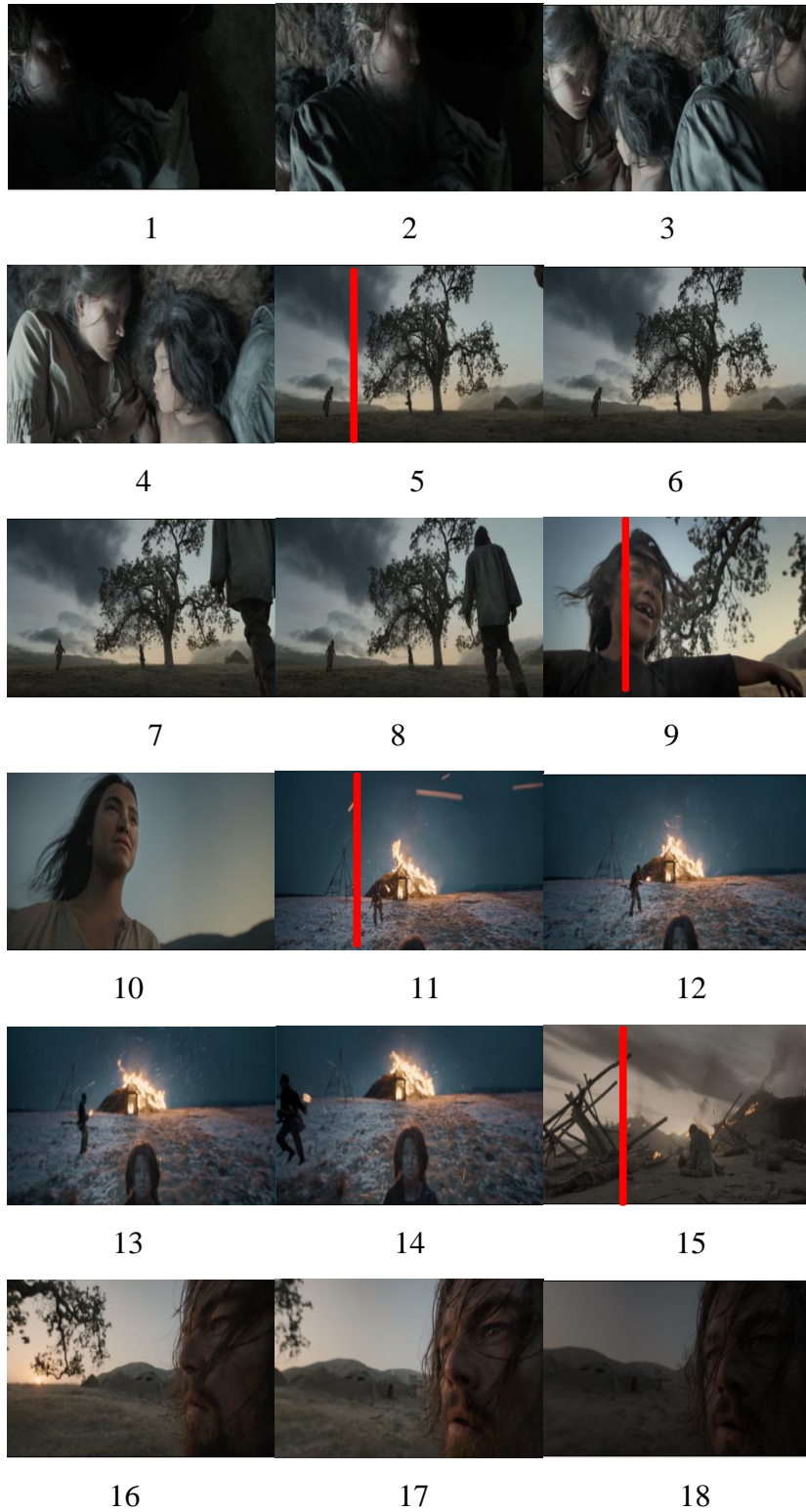
The type of shot also creates filmic cohesion. The close-up and medium close-up shots in the sequence in Figure 4.1 are limited to Glass, his boy, and his wife, creating familial ties and linking the role of the protagonist to the narrative. For example, in Pictures 9 and 10, Glass's boy is framed with the same type of shot and the same low angle as his Indian mother in two short consecutive scenes. This means that Picture 10 refers back to Picture 9 in an anaphoric relation, relating a boy to his mother through the use of the same shot size and the same camera angle.

The majority of the close-up shots in the film point at the protagonist, Glass, his suffering, his past reflections, and his future intentions. The close-ups, in other words, are the windows to the character's feelings and intentions.

Camera movement suggests the type of action on the screen as well as the world in which the characters are involved. Another notion that is related to camera movement is speed of motion. Speed is controlled by what is known as the frame rate. Frame rate is the speed at which individual still photos are displayed.

Figure 3.2

*Analysis of One Minute of Screen Time (The Revenant, 2017)*



The slow movement of the camera and the characters in the above sequence from *The Revenant* (2017) is a strong cohesive device that takes the viewer to the world of Glass's hallucinations and flashbacks. The same slow movement of the camera takes place in two sequences other than the above one, which in turn project Glass's unreal world to the viewer. Everything moves slowly, including the camera, whenever Glass reflects back to past events or dreams of his past life. This means that the scenes which depict the mental, oneiric world of the protagonist are all tied up by the slow motion of the camera and the characters.

### 3.1.3.3. Color, Lighting, and Tone

Color, lighting and tone of the pictorial composition are common elements between both still photography and painting on the one hand, and film on the other. Kress and Van Leeuwen (2006) regard color as a marker of modality. The meaning of modality here is the degree of realism in a picture or a painting. In visual compositions such as film, color and lighting differentiate subjects and elements on the screen, adding emphasis to one element and depriving another from its pictorial focus. Tone, in comparison, refers to the degree of brightness and contrast within the frame. Kress and Van Leeuwen (2006, pp.160-161) associate color with modality, where color in static pictures and portraits assigns degrees of reality. What is real or unreal is dictated by color, where color saturation marks out degrees of modality (Figure 3.3).

A final note on the analysis of color, lighting and tone is that the term reality refers to film reality (internal reality of the film) and not realism. Film scenes are depicted as real or unreal in the context of the film through the subtlety of color, lighting and tone. Realism and reality are not the same thing (Metz, 1992). The notion of film internal reality, as not being general reality as we know it, allows filmmakers to convey to the viewer the reality of the scenes.

Figure 3.3

#### *Degrees of Modality through Light Saturation*







5



6



7

#### 3.1.3.4. The Score

The score/soundtrack of a film is one strong cohesive device that ties events together. Multiple scenes can be unified through the soundtrack, hence, enhancing filmic cohesion as one unit. Music is related to certain events and characters. The score serves well as a cohesive device that ties bits and pieces of the narrative. The message of a film is encoded and unified through the score of the film.

In *The Pianist* (a 2002 film by Roman Polanski), Chopin's pieces in the film score run as a motif from the opening of the film to the end, a motif that points at the theme of the suffering of the main plot of the character of Wladislaw Szpilman and the late sub-plot of Wilm Hosenfeld. The interplay of picture and score merges the two plots together in a cohesive whole by alerting the viewer that the story is also about Wilm Hosenfeld.

Other elements can also function as cohesive devices throughout a particular film. In the case of *The Revenant*, fire, mist, smoke, and frost are elements that appear almost in 90% of the film, either occupying the background or working as mystifying objects that define the world of the narrative. In the case of *The Pianist*, there are the costumes, fire, ravaged buildings, corpses, and instruments, all pointing back to earlier parts of the narrative and forward. In the *Shawshank Redemption*, the viewer is always exposed to bars, walls, barbed wire, books, and guards.

#### 4. Findings and Discussion

The film elements that constitute the core of the analysis of filmic discourse has been shown to be picture, action, dialogue, score and sound effects, and the pictorial resources such as camera angle, camera movement, the framing of the subject and their composition in the picture, color, tone, and lighting. Without these, it would be implausible to offer a thorough analysis of a film.

Pictures, scores, lighting, colors, speed of motion, angle and movement were shown to be signs that make specific meanings, i.e., a language that is read through the lens of the directors, gaffers, cinematographers, and, later on, the viewer. Findings suggest that films are also systems in the sense that a beat is a segment of film composition the same way a word, for example, is to the sentence. It is the



building of the film from its smallest segment, such as a beat, to its largest structure, the act, that makes the ultimate meaning. This building up of scenes from pictures to beats, sequences from scenes, and acts from sequences is a systematized process. Any change of the structure of a scene or a sequence is a change in the meaning of the composition, i.e., a change in the internal language of the film. This suggests that a film is a constituency.

The earliest film theorists such as Bazin (1972; 2004) and Metz (1992) regarded cinema as a new form of art, although different from earlier forms such as painting, theatre, and still photography in its dynamicity and relatedness to the notion of realism. Because of their closeness to the notions of realism and reality, films have the ability to shape social beliefs and behaviors of individuals (Cape 2003; Wedding & Boyd, 1999), which in turn affect cultural values due to the degree of realism film stories provide to viewers.

The four leading dimensions of the semiotic analysis of any discourse are discourse, genre, style, and modality (Kress & van Leeuwen, 2006, p.91). The reading of semiotics of a film as a visual story will depend chiefly on the genre. The current analysis revealed that the film script as well consists of semiotic signs that are dictated by the genre. Two effective tools allowed the hidden semiotics of the film script to become clear: elements of *Mise-en-Scene* and *StudioBinder*. When approaching a script in light of the components of *mise-en-scene* and the cinematographics of *StudioBinder*, the reader would be able to accentuate film signs that pertain to the rendering of the moving image.

But application of the pure cinematic constituents of Bordwell and Thompson (2008), in addition to the cinematographic elements that guided the present analysis are best applied to the multimodal world of film rather than to the screenplay as a text. The film elements that constitute the core of the analysis of filmic discourse have been shown to be picture, action, dialogue, score and sound effects, and the pictorial resources such as camera angle, camera movement, the framing of the subject and their composition in the picture, color, tone, and lighting. Without these, it would be implausible to offer a thorough analysis of a film.

Findings also suggest that the picture is the smallest building block in the cinema, i.e., other constituents of the film, such as beat, scene, and sequence, may not offer the adequate analysis the picture do. It is the basis for any film production and analysis. Pictures build beats, beats build scenes, scenes build sequences, and sequences build acts. To deprive the film and the script of the concept of *system* is to ignore the building up of the visual story from its smallest unit, i.e., picture, to its largest unit. This stratification of the elementary units of film and their development into larger units has been a well-established process since the early cinema. Such stratification is not obvious only in the final film; it has its traces in the film script as well.

Modern cinema is based on montage (an editing technique where a series of pictures or short shots are sequenced to a condensed space and time) and mise-en-scene (the arrangement of scenes and subjects on a stage or within a filmic frame). It is, in fact the cutting down and the joining of the discrete pictures, beats, and scenes that renders the final story. Without the systemic strata of the motion picture, montage would not be possible.

## 5. Conclusion

The present study attempted a multimodal analysis of ten films, deploying pictures that ran parallel to their action segments (the parts of the film script that constitute characters' actions and the story's events). This allowed the present analysis to demonstrate that there is not always agreement between the textual mode of the script and the visual and auditory modes of the produced film. Screenplays, thus, are subject to a process of molding by directors and cinematographers – a process that would fit the desired story in the filmic production and not necessarily following a certain shooting script. Elements of the screen outweigh the screenplay it represents. There will be auditory and visual gaps on the page, or rather missing pictorial and auditory information such as the score of the film and secondary subjects whose presence in the filmic scene is not necessarily mentioned in the script.

The study also designated cohesive elements that tie filmic scenes and sequences. A film is one unit throughout subjects, camera angles and movements, lighting, color, and the staging of props, characters, places and times. These visual cohesive tools, in addition to the standardized linguistics and formats of the screenplay, suggest that a film is a language *system* as well as a *langugae*, and not only a *langugae* in the Metzian sense (Metz, 1992). It is this systemic organization of the verbal and visual story that makes the ultimate meaning of a film in modern cinema.

## 6. Limitations and Suggested Research

Because cinema is a global industry, generalizing the present findings is limited to Western film scripts and films, in particular to those written in English and produced by international film corporations. The findings may not also be applicable to cinematic works before the 1990s due to the fact that the time span of the present sample ranges from 1994 to 2000.

It is perhaps sociolinguistics that would be a strong candidate for further research on films – how films shape the ideology and behavior of social groups across different cultures. This would require the selection of samples of film scripts and their films across a given number of cultures.

## References

- Bateman, J., & Schmidt, K. H. (2013). *Multimodal film analysis: How films mean*. Routledge.
- Baumgarten, N. (2008). Yeah, that's it!: Verbal reference to visual information in film texts and film translations. *Meta: journal des traducteurs/Meta: Translators' Journal*, 53(1), 6-25.
- Bazin, A. (2004). *What is Cinema? vol. I* (Vol. 20). Berkeley, CA: University of California Press.
- Bednarek, M. (2015). Corpus-assisted multimodal discourse analysis of television and film narratives. In *Corpora and discourse studies* (pp. 63-87). Palgrave Macmillan, London.
- Bo, X. (2018). Multimodal Discourse Analysis of the Movie "Argo". *English Language Teaching*, 11(4), 132-137.
- Bordwell, D., Thompson, K., & Smith, J. (2008). *Film art: An introduction* (Vol. 8). New York: McGraw-Hill.
- Cape, G. S. (2003). Addiction, stigma and movies. *Acta Psychiatrica Scandinavica*, 107(3), 163-169.
- Halliday, M. A. K., & Hasan, R. (2013). *Cohesion in English* (No. 9). Routledge.
- Halliday, M. A. K., & Matthiessen, C. M. (2014). Halliday's introduction to functional grammar. *Routledge. Journal of Pragmatics* 22: 325-373.
- Kress, G. R., & Van Leeuwen, T. (2006). *Reading images: The grammar of visual design*. Psychology Press.
- Metz, C. (1992). *Film language: A semiotics of the cinema*. University of Chicago Press.
- O'Halloran, K. L. (2008, July). Multimodality around the world: Past, present, and future directions for research. In *Plenary Paper Presented at the 35th International Systemic Functional Congress*.
- Rohrbach, A., Rohrbach, M., Tandon, N., & Schiele, B. (2015). A dataset for movie description. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3202-3212).
- Wedding, D., Boyd, M. A., & Niemiec, R. M. (1999). *Movies and mental illness: Using films to understand psychopathology*. Boston.

