Examining Film Engagement Through the Visual Language of Comics

4/28/2017

R. Brad Yarhouse

ABSTRACT

As a meta-medium, comics are a synthesis of a number of communication systems involving a sophisticated dance of Images, words, and symbols. These elements build relationships that contribute to a comic reading experience. Based on more than the juxtaposition of elements, comics weave a network of patterns that can be analyzed and expressed. New studies led by researchers such as Neil Cohn suggest a structure similar to language can be found in comics. Film also contains a variety of media elements creating sequential relationships within the moving image experience. Like comics, a viewer's participation with film can be passive or involve a variety of levels of engagement through inference. By examining film media within the context of the visual grammar and theories proposed for comics, we may gain more insights into how engagement in film is achieved.

Introduction

Before there were movies, before photos, there were drawings. They provided a means of describing the world in a method that was universal and yet culturally specific (Wilkins 2016). Placing images into sequences allowed for a method of visually expressing movement, time, and narratives. Sequences of images are a communication tool that has been used by many cultures throughout human history. Often mentioned in this regard are Egyptian and Mesopotamian reliefs, the medieval Bayeux tapestry, and Mayan text image pairings found on ceramics and stuccoed walls among other places (Wichmann and Nielsen 2016).

In recent history, sequential imagery often takes the form of comics and movies. Comics often are a synthesis of a number of communication systems, a sophisticated dance of images, words, and symbols. These elements build relationships that contribute to a comic reading experience. Based on more than the juxtaposition of elements, (Cohn et al. 2012) (Rollings and Adams 2003) (Vaughan et al. 2015) (Wilkins 2016) (Merriam-Webster n.d.) (Yus 2008) sequential art weaves a network of patterns that can be analyzed and expressed. Many theorists and artists have suggested that comics are a visual language. (Eisner 1985) (McCloud 1993) (Hatfield 2005) But a growing body of research suggests that comics, rather than being a language, use a visual language (Cohn 2013). Moving pictures also work in a sequence, not only by means of editing but by creating the illusion of movement frame-by-frame. They incorporate multimodal elements such as verbal/written language, sign/body language, and visual moving images that build relationships within the movie experience and help communicate the vision of the authors. Comics, movies[1], and other sequential systems rely on communicating the creator's intent through a structure or grammar (Cohn, 2013: 4) understood by both the creator and the user. This structure acts as the language to help convey ideas from author to the perceiver. By

itself, it doesn't provide emotional involvement or commitment. In comics, the more the user is requested to participate, the more the user engages in the medium moving it from a passive viewing to a participatory experience (McCloud, 1995: 70). Like comics, can participation within a sequential language be one of the factors in films engagement? By examining moving media within the context of the visual grammar and inferential theories proposed for comics, we may gain more insights into how engagement in movies is achieved. Within the course of this paper, we will briefly discuss several theories that are being applied to narrative comics such as, visual language grammar, multimodal relationships, and inference. We will examine the opening to the 1993 film*, Blue* Directed by Krzysztof Kieślowski to see how these ideas can be applied.

Visual Language Grammar

Recent research by Neil Cohn, from Tilburg University in the Netherlands, into linguistics and sequential images suggest a structure similar to language can be found in comics and other sequential imagery. Cohn's theory and research on visual languages are grounded in cognitive theory and seek to provide research data on the building blocks of visual narrative communication. His theory for a visual language is not merely a metaphor or analogy. As he states in the *Visual Language of Comics* (2013), "The notion of a "visual language" is here meant to be on par with verbal and signed languages as a primary human ability for the expression of concepts using a grammatical system." "Humans use only three modalities to express concepts: creating sounds, moving bodies, and creating graphic representations." (Cohn, 2013: 3) The concept of a drawn language like any language requires a grammar built upon a sequence of elements that are constructed with a significant relationship to one another. A network of relationships that we hold in our memory and refer to for contrasting and comparing

creating a global view that builds the narrative structure (67) and orders meanings. (70) He lists visual language grammar as having several core components:

- **Orienter** (O)—provides superordinate information, such as a setting
- **Establisher** (E)—sets up an interaction without acting upon it
- **Initial** (I)—initiates the tension of the narrative arc
- **Prolongation** (L)—marks a medial state of extension, often the trajectory of a path
- **Peak** (P)—marks the height of narrative tension and point of maximal event structure
- **Release** (R) —releases the tension of the interaction

These categories constitute phases in Cohn's method relating to phrases in sentence syntax. Phases make up a narrative arc. Phases are a rule that follows this ordering:

(Establisher) – (Initial (Prolongation)) – Peak – (Release)

Cohn states, "Together these categories form phases, which are coherent pieces of a constituent structure. Just as phrases make up a sentence in syntax, phases make up an Arc in a narrative. Each Arc can be considered as a "visual sentence," meaning that a longer comic book or graphic novel may contain many Arcs throughout the whole story." (70)

In the following example {Fig. 1} from the comic *Paper Girls* (Vaughan et al. 2015) the first panel acts as an Orienter/Establisher determining the location and sets up the interaction. Panel 2 is the Initial, the paper flying through the air. The third panel provides a Peak—a moment of completion in the narrative phase.

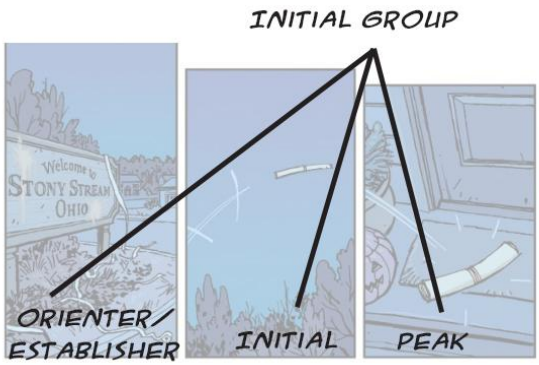

Fig. 1. Copyright Image Comics 2015.

Initial groups can also be subsets of a larger context. By continuing to breakdown *Paper Girls*, this hierarchy is apparent. Reviewing the full page, {Fig. 2} the Initial group is joined by a *release* image, revealing that a girl is throwing papers. The same image initiates the next sequence, which is quickly released with the last image of a boy, dressed as Freddy Krueger asking for a newspaper. The final two images act as a group and are in fact the Peak group for the page. They sum up nicely the reason for the series of comic panels we have read. They, along with the first Initial group, create an overarching arc, which connects together to give us a narrative structure for the page.

Cohn recommends beginning with identifying the Peak(s), (which are the peak moments of crucial action) then identifying the Initials, (which initiate the peaks and always are before the Peak(s). The Release is the completion of the sequence; in this case, there are two releases. The first set of images introduces us to the girl, the next two images to the antagonists.

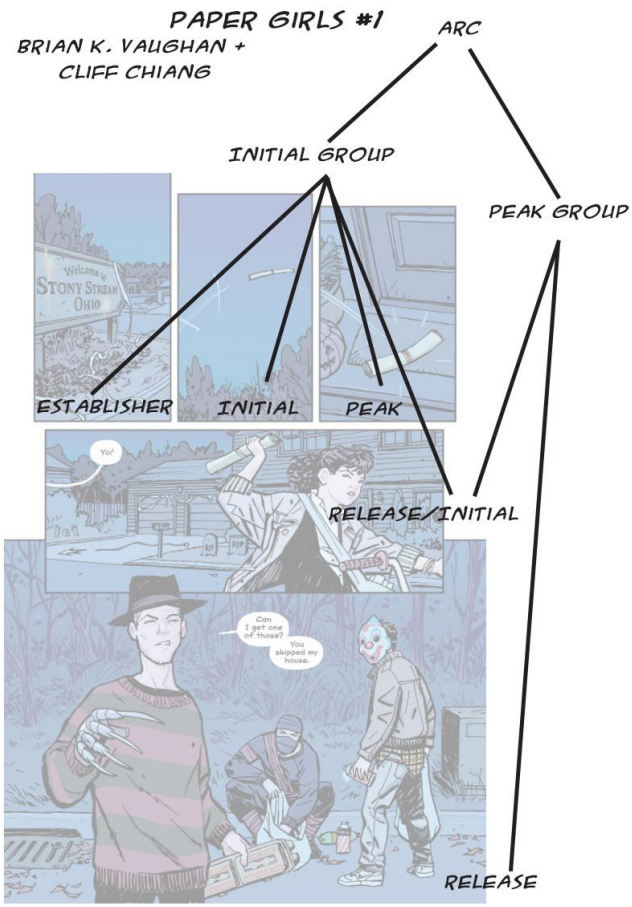

Fig. 02 Visual Narrative Grammar diagramed for Paper Girls Copyright 2015 Image Comics.

Cohn and others have published a growing number of papers researching visual language and cognition. Concepts such as the relationship between narrative structure and coherence (Cohn and Bender 2017), the study of attention within ordering of images (Foulsham, Wybrow, and and Cohn 2016), Neurological evidence of language structure in visual narrative comprehension (Cohn et al. 2014), these and other research are part of a growing body of evidence supporting a visual language.

Multimodal Architecture

Comics are a synthesis of communication systems. Written text, which follows a

syntactic grammar can take the forms of captions, exposition, onomatopoeias, and word balloons and combines with a visual language grammar made up primarily of pages, drawn panels, symbols, and individual elements within drawings. Within a comic we often see relationships:

- Between panel to panel
- Between texts and images
- Between text to text (word balloons or text boxes)
- Within an image with other parts of the image
- Between page and other pages

The relationships between multimodalities, in comics, text and images, create a rich user experience. For example, a word balloon's tail relates the text within to the drawn character that is indicated to be speaking. This is a relationship between visual structure, (the comic panel, page, or space it occupies), the Morpheme (the sign of word balloon and indicative tail), and written text (the words in the balloon)[2] (Cohn, 2013: 35). Together they react with other elements such as other characters within and without the frame to create a narrative. Scott McCloud in his 1993 book *Understanding Comics* describes a hierarchy of relationships between words and images breaking down these relationships into a series of categories.

- **Word Specific**: Pictures illustrate, but don't significantly add to a largely complete text.
- **Picture Specific**: Words do little more than adding a soundtrack to a visually told sequence.
- **Duo-Specific**: Panels in which both words and pictures send essentially the same message.
- **Additive**: Where words amplify or elaborate on an image or vice versa.
- **Parallel:** Words and pictures seem to follow very different course--without intersecting.
- **Montage:** Words are treated as integral parts of the picture.[3]
- **Interdependent:** Words and pictures go hand-in-hand to convey an idea that neither could convey alone.

(McCloud, 1993: 153)

This hierarchy of relationships gets to the heart of an exchange of information that is occurring out of sight of most comic readers. While McCloud astutely observed these, they are part of broader relationships that go beyond comic panels and word balloons. Cohn has widened these relationships expanding them into a multimodal grammar for all visual languages (Cohn 2016). By doing this he provides the start of a multimodal architecture that can be applied to not just words and images but to many multimodal relationships from simplistic to complex.

- **Autonomous**: non-multimodal. For example, an image or series by itself or text in a book without images.
- **Dominant:** one modality is semantically dominant using grammar (visual or otherwise) the other modality does not have a grammar.
    - **Vis-Dominance**: the visual dominates over other elements. Other elements do not use grammar. Examples of text that is paired with images in this manner would include Onomatopoeia (sounds are written like a word such as Bam!) or a drawn street sign with the word ‘Stop’ on it. {Fig 3.}
    - **Verb-Dominance**: language is dominant over visuals. Visuals can offer context without conveying important information. (A three-panel strip with the same image in each but different words) other types include illustrations, emoticons, or emoji at the end of a sentence. {Fig 4.}
    - **Co-Dominance**: A balance of semantic dominance but at least one modality lacks a sequential grammatical structure. Single panel comics fall into this where the image is singular and the text follows a grammatical structure. But both must be present in order for meaning. {Fig.5}

    - **Substitutions:** replacing a modality with another in a component. I ♥ NY or in comic panels replacing a gunshot with an "action star" for a panel. {Fig.6}

- **Assertive:** A grammatical structure in both modalities (syntax, narrative) while the meaning is controlled by one.
    - **Verb-Assertive**: Words guide the semantics. Meaning is handled by the verbal. Sequential visuals now supplement the text giving more expressive meaning beyond just placeholders. McCloud would call this Word-specific. {Fig.7}
    - **Vis-Assertive**: Visuals have a narrative structure and guide the meaning of a sequence, beyond that in the text. Dialogue is mostly redundant with the visuals. McCloud would have called this Picture-specific. {Fig.8}
    - **Co-Assertion**: Both modalities use grammatical structures and both modalities semantically contribute in non-negligible ways. The sum is greater than the parts. This is similar to Eisenstein's Montage theory (described later) and relies on McCloud's well-known ideas and language.
        - **Parallel:** Words and pictures seem to follow very different course--without intersecting. {Fig.9}
        - **Duo-Specific:** Panels in which both words and pictures send essentially the same message. {Fig.10}
        - **Additive:** Where words amplify or elaborate on an image or vice versa. {Fig.11}
        - **Interdependent:** Words and pictures go hand-in-hand to convey

an idea that neither could convey alone. {Fig.12}

- **Absorption:** semantics of visual being absorbed by the text. Instead of showing the visual, the text describes it separate from the visual accompanying the text. "Meanwhile back at the ranch…" might accompany an image that does not show the actual location but perhaps people in the interior of the ranch (Cohn 16).

What these categories briefly tell us is that multimodality, which on the surface appears simple, is a complex concept. These elements combined together share complicated relationships with one another, requiring the reader to decode and then deduce the unseen by calling on various levels of inference. Inference is a means of engagement in the comic experience.

Fig. 03 An example of Vis-Dominance. Calvin and Hobbes, Bill Watterson, August 29, 1988

Fig. 04. An example of Verb-Dominance. Garfield, Jim Davis, 2009

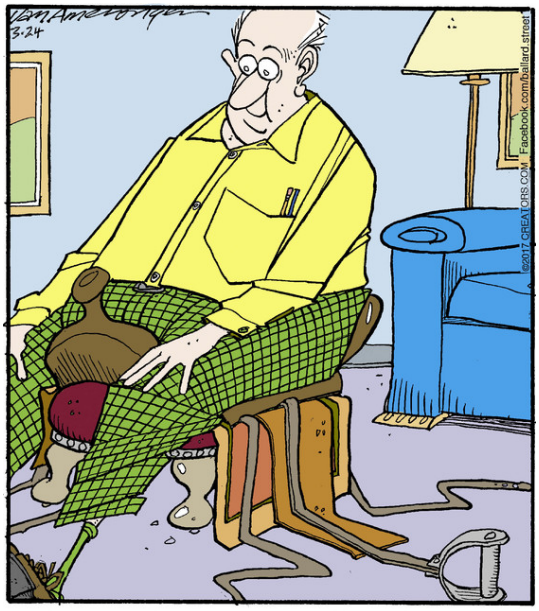

Fig. 5 An example of Co-Dominance. Ballard Street, Jerry Van Omerongen, 2017

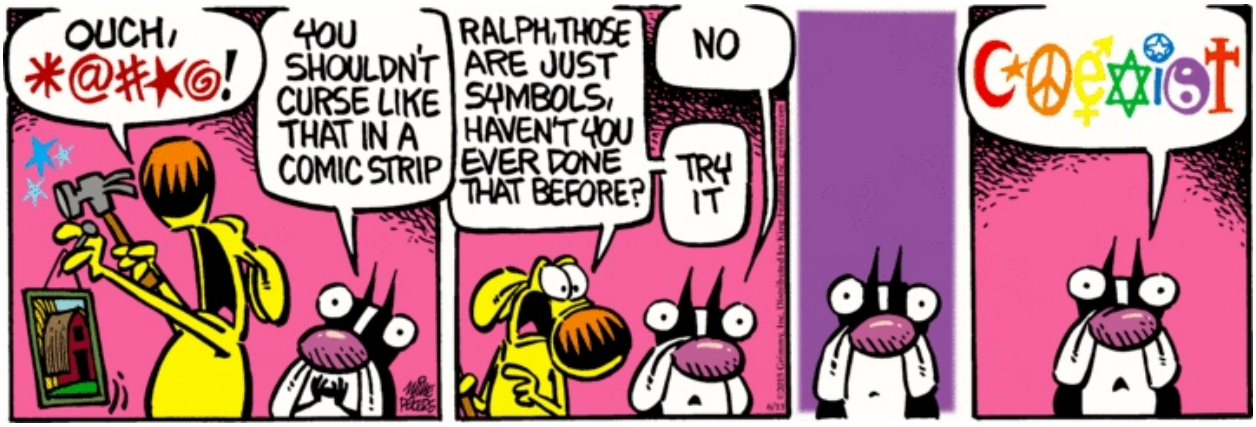

Fig. 6 An example of Substitution. Mother Goose and Grimm, Mike Peters, June13, 2015

Fig. 7 An example of Verb-Assertive, or Word Specific. Understanding Comics, Scott McCloud, 1993

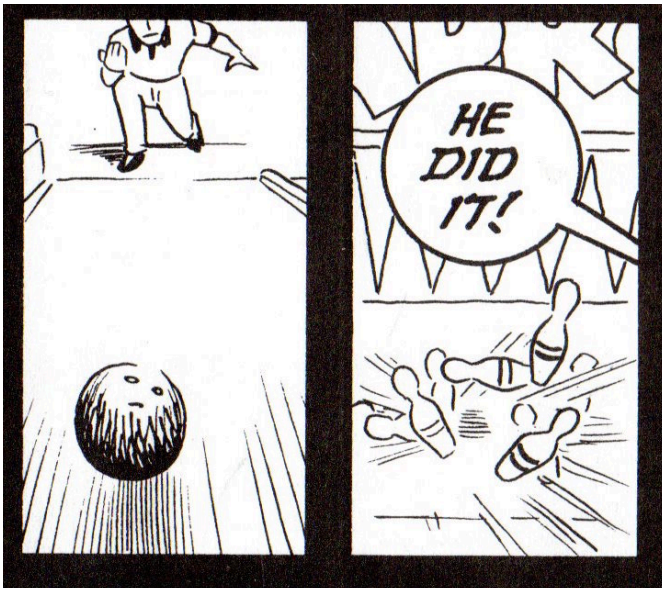

Fig. 8 An example of Visual-Assertive, or Picture Specific. Understanding Comics, Scott McCloud, 1993

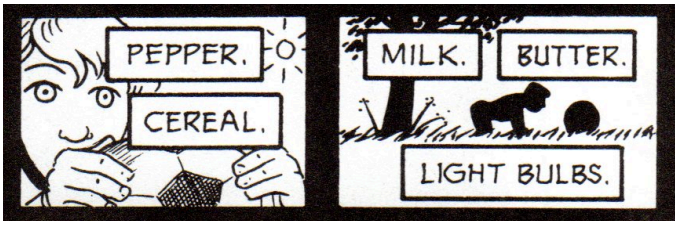

Fig. 9 An example of Co-Assertion: Parallel. Understanding Comics, Scott McCloud, 1993

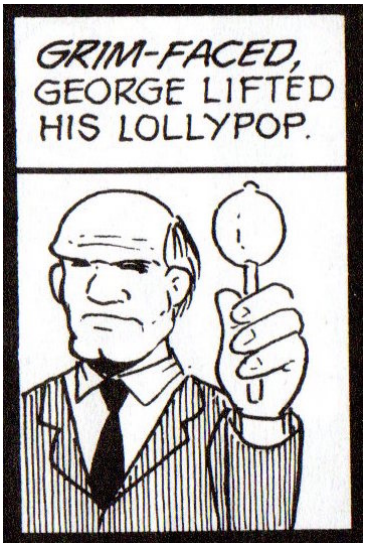

Fig. 10 An example of Co-Assertion: Duo Specific. Understanding Comics, Scott McCloud, 1993

Fig. 11 An example of Co-Assertion: Additive. Understanding Comics, Scott McCloud, 1993

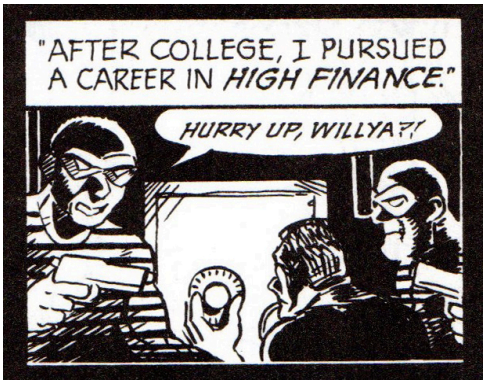

Fig. 12 An example of Co-Assertion: Interdependent. Understanding Comics, Scott McCloud, 1993

Engagement

Media engages us in a variety of ways. From the use of interesting points of view, interesting settings; that can involve physical, temporal, environmental, emotional, and ethical dimensions, (Laramee 2002) (Rollings and Adams 2003) compelling characters that we can empathize with, interactive actions or hooks in games and other media that require the user to make a "decision that relates to the game, and thus keeps them playing" (Howland 2002)[4] Beyond these, in comics and film inference has long been thought to play an important role in engagement. (Bordwell 1985) (Branigan 1992) (Chatman 1978) (Eisenstein 1957) (Magliano, Dijkstra, and Zwaan 1996) (McCloud 1993) (Saraceni 2015). Inference drives the participation required in comics by the viewer to generate the experience of reading a comic versus watching a comic.

Scott McCloud in *Understanding Comics* (1993) used a Gestalt-based theory for how comics functioned based on closure or inference. He stated, "Closure in comics fosters an intimacy surpassed only by the written word, a silent, secret contract between creator and audience." (McCloud, 1993: 69) Closure occurs in the brain and allows us to complete

incomplete data. Our brains fill in the missing material (McCloud 1993). Cohn states in *Linguistics and the Study of Comics*:

> "A primary focus of semantic research of visual language has focused on inference, the drawing of non-provided meaning from the existing forms. Inference has been a motivating notion in discourse studies particularly, as how in the sentences "The fireman sprayed the water on the house. Smoke rose from the building." the reader derives the meaning that the house was on fire and it went out, though such concepts are never mentioned overtly in the text." (Cohn 2013)

Closure is multimodal and affects all the senses. In the world around us, we constantly piece together visual elements to complete an unseen picture. Seeing an object in the dark can lead to a correct inference or fallacy. Hearing a popping noise at night can lead us to think we are hearing gunfire or in the cultural context, a celebration with fireworks. As David Bordwell states, "Sensory stimuli alone cannot determine a percept since they are incomplete and ambiguous. The organism constructs a perceptual judgment on the basis of nonconscious inferences."(Bordwell, 1985: 31) Engagement within comics is in part a binary dance between the gestalt principle of inference or closure and a visual language grammar structured towards building the narrative. Comics rely on a variety of inferences to fuel the user's engagement.

Comics often use a type of inference called a *Bridging Inference* (Haviland and Clark 1974) (McNamara and Magliano 2009) an inference that requires the reader to fill in the missing information in order to continue comprehending the material. Inference has been written about by Eisenstein in 1942 while describing Montage theory, Eisenstein stated: "This property consisted in the fact that two film pieces of any kind, placed together, inevitably combine into a new concept, a new quality, arising out of that juxtaposition." (Eisenstein 1957)

Levels of Inference

Inference has various levels of engagement based on a relationship between the organizing structure and practices of the medium and the level of inference required to comprehend. It is a range that was first discussed by Scott McCloud in *Understanding Comics* that he called panel transitions. McCloud provided for relationships between panels, which were next to each other, Cohn has described this as limiting but admits that "Panel transitions are especially intuitive since they follow how we experience the reading process: moving from one panel at a time." (Cohn, 2013: 67) McCloud ((Schulz 1952-1974)1993) broke these down into:

- **Moment-To-Moment**. Requiring little closure.
- **Action-To-Action**. Distinct cause and effect relationships.
- **Subject-To-Subject**. Staying within a scene or idea but requiring the user to bridge. (See above)
- **Scene-To-Scene**. Similar to Subject-To-Subject but requiring more user effort to imaginatively bridge the narrative gap.
- **Aspect-to-Aspect**. Time is suspended and a wandering eye examines different aspects of a place, idea or mood.
- **Non Sequitur**. No logical relationship between panels whatsoever.

McCloud himself considered his terms an inexact science at best; nevertheless, engagement within a comic's structure seems to be influenced by the level of inference required by the user in order to culturally comprehend the material. In research that was conducted in 2015 and reported in the journal *Neuropsychologia*, Neil Cohn and Marta Kutas ran a study involving 36

comic readers and 120 novel visual sequences from the *Complete Peanuts* by Charles Schulz (1952-1974) to determine how inference affected the brain recording EEG data. The goal was to determine the effects of looking at a series of comic strips with four criteria for measuring inference. The scale is listed as:

A. **Impoverished.** Requiring much inference to understand the relationship between panels.

B. **Implied**. The panel provides a clue that something will occur.

C. **Expected**. Whereas the anticipated image is provided but causes discontinuity with the final image.

D. **Explicit**. A panel is provided that "solves" the final panel without inference. {see fig.13}

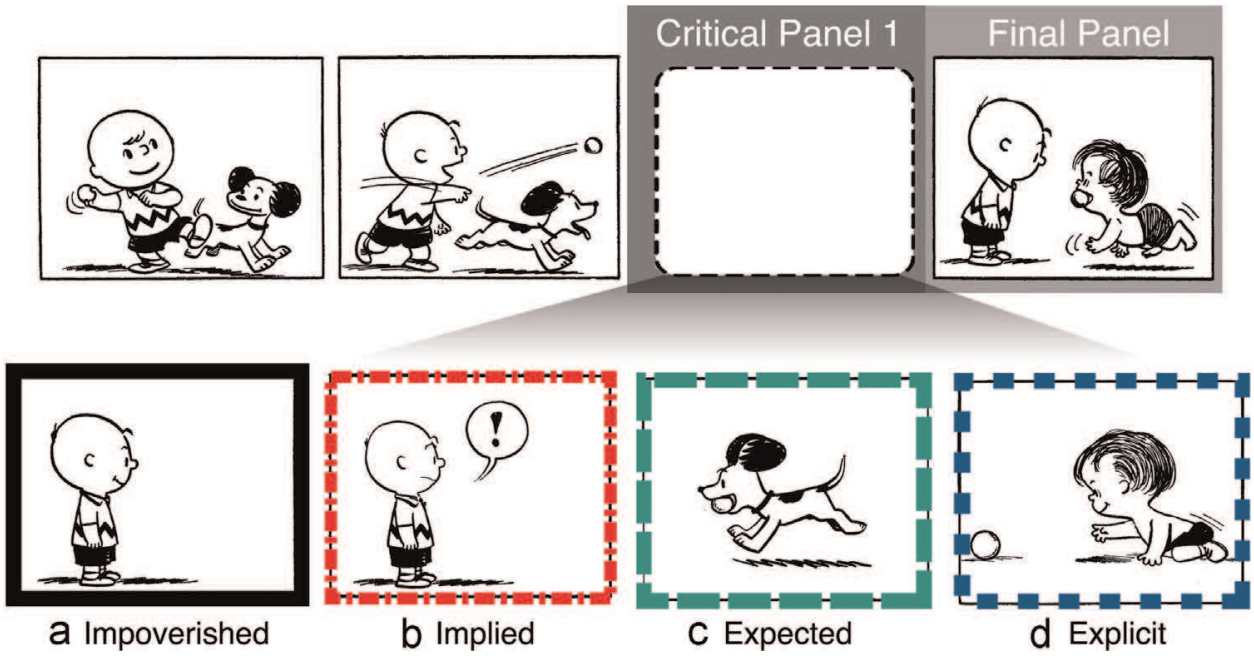

Fig.13 Examples from experimental stimuli using a single sequence-frame with multiple critical panels.

(Cohn, Kutas, 2015) Images Copyright Peanuts Worldwide LLC.

The research suggested that users found the 'explicit' panel to be the easiest to comprehend but that impoverished or implied sequences may contribute to a more satisfying narrative for a reader (Cohn, Kutas, 2015: 276).

The more inference required within the understood organizing structure, the greater the engagement—to a point. There seems to be a 'sweet spot' between too much reliance on inference, which leads to incomprehension, and too little inference leads to pacing issues (time). Inference can be slippery; it relies on partial information that must be filled in by the user in order for the complete message to communicate. This includes a level of cultural understanding. A medium such as comics relies on inference to be the engine that drives the user's experience. The visual language grammar that Neil Cohn has proposed, along with multimodal relationships, can be seen as the organizing structure of the comic's medium while inference provides a means of engagement.

Applying Sequential Narrative Grammar and Inference to Krzysztof Kieslowski's *Blue*

In a series of opening scenes, we are introduced to Krzysztof Kieslowski's film *Blue* (1993) and the events that haunt the rest of the film. We see at first a seemly-unrelated sequence of events, which are only constructed into story on reflection, not while in the process of moment-to-moment viewing. When watching this for the first time, the audience is trying to construct the narrative based on a familiarity with narrative and film language. We are first watching the characters in an attempt to identify who is the protagonist, to whom are we to relate? What is the location? What is the story? Kieslowski knows this and uses misdirection to play with our assumptions.

He offers up Anna first, the girl amongst the streaks as our protagonist, then reveals a

mother's voice and a father standing as Anna goes to pee in the woods. A few seconds later, we are introduced to a boy playing with a ball and stick toy, juxtaposed against a similar street noise soundscape stitching together this shot to the previous shots. Is he our protagonist? The boy provides us with our first emotional release, he achieves placing the ball on the stick of a ball and stick children's toy and smiles with pleasure. We share this moment with him. Then the sound of a crash occurs off camera. He along with us hears the sound as he observes the crash off screen. We share his desire to race to the location, to save anyone whom can be saved. Anna has moved from the temporary seat as protagonist to victim, the boy has taken the 'lead' as protagonist. The scene fades to black as he races to the scene.

The next scene opens with a disorienting image of a fluff of cotton in extreme close-up. The use of focus in this shot visually echoes the shot from under the car when Anna returned from peeing, this shot is set up so that the fluff is in extreme focus with the background out of focus while a figure moving towards us from the background. This shot seems to have no relationship to the previous car crash scene. The next shot is an extreme close-up of a single eye in which we can see the doctor reflected in the pupil in focus talking to the mother, Julie, (played by Juliette Binoche) as it turns out, is the only person to have lived through the accident. We step back to reveal Julie for the first time in a hospital bed. The doctor informs her that her husband is dead. She speaks one word "Anna?" The doctor replies, "Yes, your daughter too." Julie hides her head and the scene ends.

The sly misdirection during the opening scene unsettles us, who is the protagonist? Who do we connect with? When introduced to Julie in the hospital bed the audience catches up momentarily with the story. The audience is led to the inference that she was the 3rd unseen person in the car.

Each element of change introduces a reaction with the audience, and pushes against the previous element or foreshadows the next elements. The viewer is being asked to participate in putting the story together. It's not being spoon fed, in fact, the opposite is happening, it is being misdirected, manipulating the audience into thinking several things while really another event is being set up: the crash, the tragedy, and the transformation of Julie's life.

In this example, relationships are being established between a variety of media.

- Between shot to shot
- Between sound and shot
- Between voice and shot
- Between film scenes and other film scenes
- Between assumptions and narrative goals

If each of the shots within the scene at the opening of *Blue* act as building blocks comprised of a variety of relationships then, by breaking down the relationships and the organization of the sequence we should be able to see the language engine that is allowing the director to communicate with us.

Summing up and applying theories toward Film

In the course of examining comic narrative, we have discussed three types of relationships that visual narrative languages and, in particular, comics engage in: first, Visual Language Grammar, second, Multimodal Architecture, and finally inference. Visual Language Grammar will be used to break down the scenes and to examine the relationships between the clips to one another and the building of groupings. Use of Multimodal Architecture and inference will be noted.

Diagramming Blue

The first scene in *Blue* can be broken down into 16 discrete shots with the 2nd scene in Blue as 3 shots. Examining the first five shots of *Blue* we are seeing a series of montage images. Montage as a type of inference can be used to either provide clarity or to misguide the user. But as Eisenstein stated, the sum is greater than the whole (Eisenstein 1957). Within the opening to *Blue*, we see Kieslowski using the concept of montage to misguide the user.

Shot 1, opens with the tire on the pavement, or what Cohn would describe as an Orienter. It sets up the road as a location for the scene. The sound in this shot is road noise corresponding to the image and will continue throughout the 16 shots interrupted by specific noises. Also note that the sound leads the visual; the visual identifies the sound and supports the sound. In Multimodal Architecture, there is no grammar being used for the sound. Grammar would consist of a series of sounds that build story) While the sound leads the visual, it is the visual that gives it context so in a sense it is *Vis-Dominant*. Note that later we visually return to the tire and so this is also an Establisher shot as indeed is this whole grouping. It establishes that the tire will be important. But it disguises it so that it's not apparent, it looks like it is giving us reference to the location (the road), which it is, but it is also telling us more.

The 2nd shot is an Establisher for this sets up the action. The action, in this case, is not disclosed but hidden. The wrapper is a curiosity to the first-time viewer. What is it, why is it here? It continues to reinforce the use of the color blue, but like many images in this film it serves two purposes. The 2nd is as an Initial it sets up the next shot and only gains context from its relationship with the girl in the next shot, an *Implied Inferential* relationship. Later in the film, it will be referenced and becomes a memory both for the audience and the main character. Shot 3

is through the back window showing the swirling street light glare. We can just make out hints of the little girl. This establishes the girl, Anna. Shot 4 is an Initial shot, a reverse of the previous shot. It shows what the girl is seeing and is a point-of-view shot. The relationship between 3 and 4 can be seen as an *Implied Inferential* shot. While a point-of-view shot is typical in western film, research into people seeing film or television for the first time has shown that they had greater difficulty connecting POV shots with its intended character (Schwan and Ildirar 2010).

Shot 5 is a repeat of shot 3 through the back window but Anna is closer to the camera. This acts as a Prolongation shot in VNG. This opening scene makes up our first set of images and as a group they act as an Establisher. {Fig. 14}

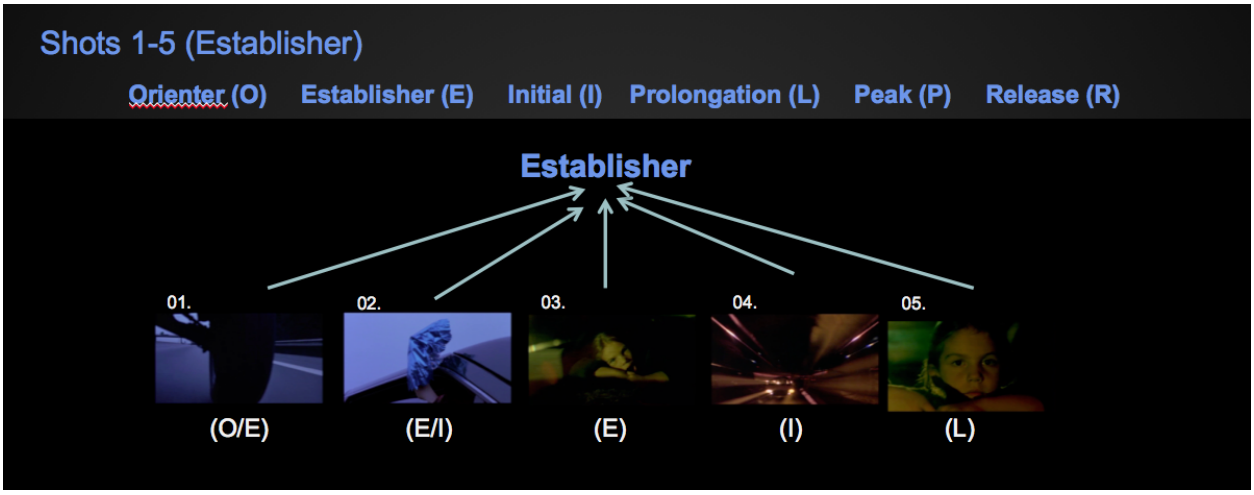

Fig. 14 This series of shots from *Blue* constitute the establisher. Copyright Miramax Pictures, 1993.

Shot 6 is a Peak shot, but it also introduces us to the father figure. Note how much story gap there is between shot 5 and 6? This is an *Impoverished Inferential* relationship. The viewer is required to fill in the story making inferences from the last scene.

Shot 7 is a Peak. The girl goes back to the car completing the need to pee and the introduction of her. But it is also an Establisher. It establishes the drip of brake fluid setting up

the coming accident. Note the voiceover, which is an example of *Multimodal Co-assertion: Interdependent* relationship, the scene without it would not function, therefore they are dependent on each other. Together they tell us who the girl is, (image of the girl coupled with the voice calling her Anna). By themselves, they don't work. We also infer that the female voice is her mom.

Shot 8 is a Release. The car leaves. This closes the 1st Peak group. {Fig. 15}

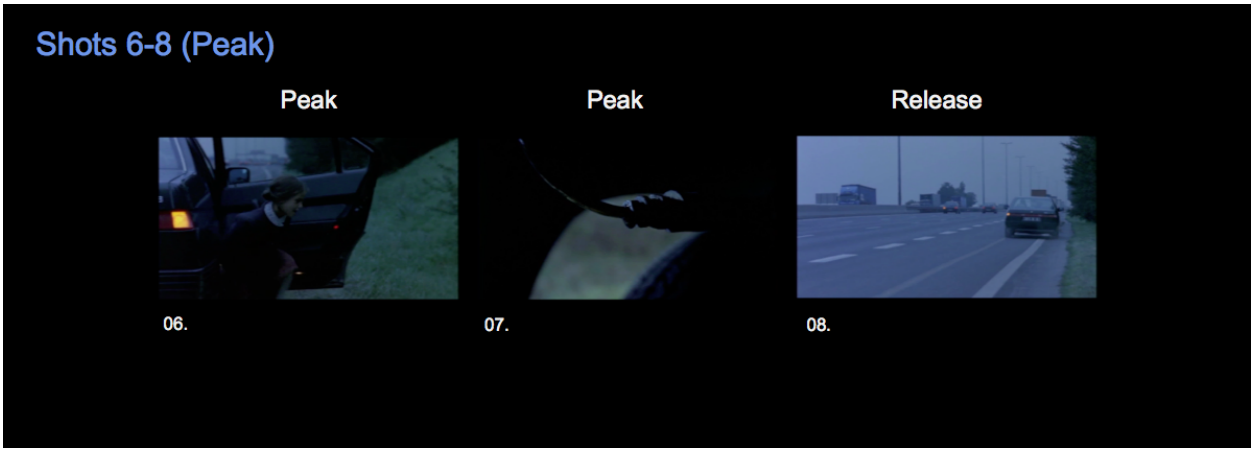

Fig. 15 This series of shots from *Blue* constitute the peak. Copyright Miramax Pictures, 1993.

Shot 9 disrupts the flow. It's a location change and an apparent change of scene but we will see that that is not the case. This is an Establisher; it introduces us to the boy who is the observer. This is an example of an *Impoverished Inferential* relationship.

In shot 10 the boy looks toward something he hears, a car. This acts as an Initial. The sound leads. In shot 11 the boy is out of focus trying to get the toy ball and stick to work. The car comes in from the background. This is a small Peak; it ties the previous segment with this new one with the boy.

Shot 12 has the boy look down. This prolongs the scene, a Prolongation. In shot 13 the boy gets ball onto the stick, success! (A Release) But at the same time off camera, we hear the

crash. Even though this happens in the same scene, this is *Multimodal Co-assertion: Parallel*, two things happening that are not seemly connected. (Sound and image are separate) This is a Peak in Cohn's VLG. Shot 14 is also a Peak. The car has hit. We see the ball fall out indicating a child on board. Not only a Peak of this segment 9-14 but also is a Peak for 1-14. {Fig. 16} In shot 15 the boy picks up skateboard and runs. This is a Release.

We end this scene with shot 16, a Wide shot of the boy running toward car. Another Release shot.

Fig. 16 This series of shots from *Blue* constitute an Initial. Copyright Miramax Pictures, 1993.

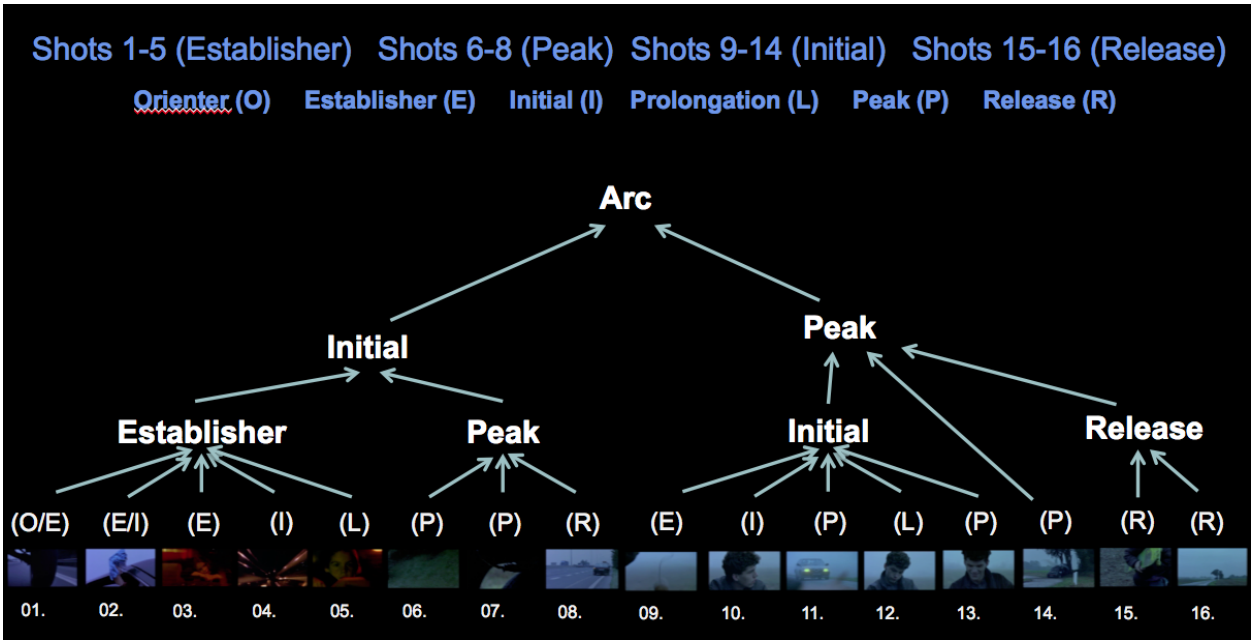

Fig. 17 This completes the narrative arc. Copyright Miramax Pictures, 1993.

Conclusion and further investigations

This paper only begins to examine *Blue's* narrative sequential shots and it's meant to be an example more than a full analysis. What does breaking down a scene such as this in this manner provide for us? Visual Language Grammar provides a tool for analysis of sequential systems. I have chosen to include Multimodal Architecture concepts and ideas on inference in the service of peering under the hood of sequential systems. This gives us the start of a structure for viewing film narrative from a different vantage point. Getting a fresh point of view on any well-known subject is always valuable.

Each of these concepts being explored in comic theory is in their infancy and much is left to expand upon and in particular, Multimodal Architecture and inferential relationships. Applying these ideas to film may bring more understanding of how relationships between many

types of narrative systems function. How do we create media that provides a similar experience as found in other mediums? Such as creating VR that gives the user a comic reading experience versus a passive experience? Can filmmakers create comic-based superhero films that provide richer experiences using narrative inferential structure versus just tacking on word balloons, bright colors and triple biceps? Consider weakened narrative systems such as 'motion comics' which in the process of pairing film techniques to comic art lose the comic reading experience and result often in poor-looking animations. How could understanding the relationships between Visual Language Grammar, Multimodal Architecture, and inferential relationships inform this medium? Does each visual narrative medium use various degrees of visual grammar, multimodal architecture, and inference? What ranges of these concepts help provide narrative experiences that are engaging and what constitutes non-engagement? Like baking, the balance of these ingredients provides a means of communication from creator to the audience. Understanding how visual grammar builds phrases, how the network of relationships between modalities communicate with one another and us, and how much inference to ask our audience to participate in could be the difference between a soufflé and a pound cake.

Works Cited

Bordwell, David. 1985. *Narration in the Fiction Film*. Madison, WI: University of Wisconsin Press.
Bordwell, David. 2011. "Common Sense + Film Theory = Common-Sense Film Theory?" David Bordwell's website on cinema. Retrieved March 13, 2017 (http://www.davidbordwell.net/essays/commonsense.php).
Branigan, Edward. 1992. *Narrative Comprehension and Film*. London: Rutledge.
Chatman, Seymour. 1978. *Story and Discourse*. Ithaca, NY: Cornell University Press.
Cohn, Neil. 2012. "Comics, Linguistics, and Visual Language: The past and future of a field." Visual Language Lab. Retrieved May 01, 2015 (http://www.visuallanguagelab.com/P/NC_Comics&Linguistics.pdf).
Cohn, Neil. 2013. *The visual language of comics: Introduction to the structure and cognition of sequential images*. London: Bloomsbury Academic.
Cohn, Neil. 2016. "A multimodal parallel architecture: A cognitive framework for multimodal interactions." *Cognition* 146:304-323.
Cohn, Neil and Patrick Bender. 2017. "Drawing the line between constituent structure and coherence relations in visual narratives." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 43(2):289-301.
Cohn, Neil, Ray Jackendoff, Phillip Holcomb, and Gina Kuperberg. 2014. "The grammar of visual narratives: Neural evidence for constituent structure in visual narrative comprehension." *Neuropsychologia* 64:63-70.
Cohn, Neil and Marta Kutas. 2015. "Getting a cue before getting a clue: Event-related potentials to inference in visual narrative comprehension." *Neuropsychologia* 77:267-278.
Cohn, Neil, Martin Paczynski, Ray Jackendoff, Phillip Holcomb, and Gina Kuperberg. 2012. "(Pea)nuts and bolts of visual narrative: Structure and meaning in sequential image comprehension." *Cognitive Psychology* 65(1):1-38.
Dickey, Michele. 2005. "Engaging by design: How engagement strategies in popular computer and video games can inform instructional design." *Educational Technology Research and Development* 53(2):67-83.
Eisenstein, Sergej M. 1957. *The Film Sense*. New York, NY: Meridian Books.
Eisner, Will. 1985. *Comics & Sequential Art*. Tamarac, Florida: Poorhouse Press.
Foulsham, Tom, Dean Wybrow, and Neil and Cohn. 2016. "Reading without words: Eye-movements in the comprehension of comic strips." *Applied Cognitive Psychology* 30:566-579.
Hargraves, Orin. 2014. "The truth about clichés: Why the stigma against them isn't entirely fair." Salon. Retrieved March 6, 2017 (http://www.salon.com/2014/07/20/the_truth_about_cliches_why_the_stigma_against_them_isnt_entirely_fair/).
Hatfield, Charle. 2005. *Alternative Comics: An Emerging Literature*. Jackson, Mississippi: University of Mississippi Press.
Haviland, Susan and Herbert Clark. 1974. "What's new? Acquiring new information as a process

in comprehension." *Journal of Verbal Learning and Verbal Behavior* 13:512-521.
Howland, Geoff 2002. "Balancing gameplay hooks. In F. D." Pp. 78-84 in *In Game design perspectives*, edited by Francois Laramee. Hingham, MA: Charles River Media.
Kieslowski, Krysztof 1993. "Blue." France: CED Productions, Eurimages, France 3 Cinema, MK2 Productions, CAB Productions, Studio Filmowe TOR.
Laramee, François. 2002. *Game design perspectives*. Hingham, MA: Charles River MEdia.
Magliano, Joseph, Katinka Dijkstra, and Rolf Zwaan. 1996. "Generating predictive inferences while viewing a movie." *Discourse Processes* 22:199–224.
McCloud, Scott. 1993. *Understanding comics: the invisible art*. Northampton, MA: Kitchen Sink Press.
McLuhan, Herbert M. 1964. *Understanding media: the extension of man*. New York, NY: McGraw-Hill.
McNamara, Danielle and Joe Magliano. 2009. "Toward a comprehensive model of comprehension." *Psychology of Learning and Motivation* 51:297-384.
Merriam-Webster. n.d. "Merriam-webster." Retrieved March 21, 2017 (https://www.merriam-webster.com/dictionary/engagement).
Peterson, Mary, Barbara Gillam, and Sedgwick H.A. 2007. *In the Mind's Eye*. Oxford, NY: Oxford University Press.
Rollings, Andrew and Ernest Adams. 2003. *on Game Design*. Indianapolis, IN: New Riders.
Saraceni, Mario 2015. "The Graphic Novel Relatedness: aspects connectivity in comics." Pp. 167-179 in *The Graphic Novel: An introduction*. Cambridge: Cambridge University Press.
Schank, Roger and Robert Abelson. 1977. *Scripts, plans, goals and understanding an inquiry into human knowledge structures*. Hillsdale, NJ: Lawrence Earlbaum Associate.
Schank, Roger and Robert Abelson. 2008. *Scripts, plans, goals and understanding an inquiry into human knowledge structures*. New York, NY: Psychology Press.
Schulz, Charles. 1952-1974. *The Complete Peanuts*. Seattle, WA: Fantagraphic Books.
Schwan, Stephan and Sermin Ildirar. 2010. "Watching Film for the First Time: How Adult Viewers Interpret Perceptual Discontinuities in Film." Jstor.org. Retrieved March 07, 2017 (http://www.jstor.org/stable/41062454).
Vaughan, Brian, Cliff Chiang, Matthew Wilson, and ared and Fletcher 2015. "Paper Girls." Berkley, CA: Image Comics.
Wichmann, Soren and Jesper Nielsen 2016. "Sequential Text-Image Pairing among the Classic Maya." Pp. 283-309 in *The visual Narrative Reader*, edited by Neil Cohn. London: Bloomsbury Academic.
Wilkins, David 2016. "Alternative representations of space: Arrernte Narratives in Sand." Pp. 253-281 in *The visual Narrative Reader*, edited by Neil Cohn. London: Bloomsbury Academic.
Yus, Francisco. 2008. "Inferring from comics: a multi-stage account." *Quaderns de Filologia. Estudis de Comunicacio* 3:223-249.

[1] For this paper I will be using the word movies in place of other words such as film, cinema, etc.

The reason being that moving pictures occur in all sorts of delivery mediums: television, Internet, Blu-ray, phones, etc., and are comprised of not just film stock.

[2] Cohn has defined word balloons and tails as part of Morphemes, using the same strategies as morphemes in verbal and sign languages, which act as for example as a prefix, suffix or infix.

[3] This is not the same use of the word Montage as Eisenstein formulated in *The Film Sense*. (1957)

[4] Examples of hooks include action hooks, resource hooks, tactical and strategic hooks, and time hooks. (Dickey, 2005)