

## **2. LITERATURE REVIEW**

### **2.1. PREVIOUS RESEARCH**

Based upon previous research, this particular research has several similarities, differences, as well as new aspects. Research by Karahan (2023) identifies how rhythm controls the viewer's cognitive and emotional engagement in the film *Raging Bull* (1980). The author highlighted how fast-paced editing creates tension, while slow-paced editing gives time for the viewer to take in information of the surroundings. The research analyzes rhythm through different film elements such as script, cinematography, sound, color, acting, and editing as the main influence. However, by discussing many different film elements, it also causes the research to become too wide and less focused. Ultimately, the current study will only focus upon the rhythmic editing aspects, as it is the main driving force of the film's rhythm.

Similarly, Das (as cited in Chetia & Konwar, 2023) found that fast cuts create more cognitive load, while slower cuts function as reflective moments. While their study is based on a psychological view, the similarity lies in the subtle exploration of tension, suspense, and release. The strength of their research is that it answers a scientific "why" by proving how emotions can be regulated through editing by using psychological theories. However, this text becomes a more theoretical basis, as it does not present a shot-by-shot breakdown or analysis for a specific film. In contrast to its use of psychological theories, this new research will use a formalist approach to analyze how rhythmic editing constructs moments of tension and release.

Another study by Sembiring, Putra, Nabil, & Alfathoni (2025) discusses how editing in general (rhythm, color grading, and sound) can elicit emotional responses in the viewer. It becomes relevant to the current study, as it highlights how editing is also affected by movements of image and sound. They identified several different elements, but each case is not explained thoroughly. To keep the current study more focused, the author will only focus on rhythmic editing aspects and how it forms tension and release.

Yogaprayuda & Risang (2022) use a more limited scope by micro-analyzing the rate of cutting used to increase tempo in *The Photocopier* (2021). They conclude that rhythm is varied and the rate of cutting has a certain intention. Its strength lies in the use of mix methods to find out quantitative information through a questionnaire asking the viewer about the film's rate of cutting. Additionally, they conducted an interview with the editor of the film to gather qualitative data. A limitation faced in this study is that it only analyzes the first 26 minutes of the film. Although the new research will also be conducting mezzo analysis of a specific scene, a scene from each act will be chosen throughout the film to see different outcomes. Additionally, a micro analysis of the shot-to-shot impact and a macro analysis of how it affects the overall plot will also be done.

Next, Saputra (2024) shifts the focus onto how rhythmic editing conveys mood in the horror genre. Quite similarly, *The Siege at Thorn High* utilizes rhythmic editing as a way to fit into the conventional action-thriller film by keeping a chaotic rhythmic movement. The author of the previous research uses Pearlman's theory and horror genre conventions to identify whether or not the first 10 minutes successfully created the desired mood. Through careful analysis, they found that fast-pacing used in *Wanalathi* (2022) was unsuccessful in creating tension as there is no set-up (tension) for the pay-off (release), which goes against the genre conventions. However, the limitation to the analysis lies in the fact that it only analyzes the first 10 minutes of each film, making it a mezzo-level analysis and less varied. This new research will discuss three scenes to see the different ways in which rhythmic editing forms tension and release, through a micro, mezzo, and macro level analysis.

In addition, there are several previous research that focuses on the film *The Siege at Thorn High* (2025). However, they are more concerned with the themes of the film, rather than the film form and style. Nadia, Soraya, & Ariska (2025) highlights the representation of youth violence and racial discrimination in the film, analyzed through Peirce's semiotics in strengthening their research. The limitation found is its main analysis focus towards the subjects and dialogues, omitting other visual symbols present at times. Meanwhile, Safitri, Heriawan, & Toni (2025)

analyzes the representation of Chinese ethnicity through Edwin's character using critical discourse analysis. Using it, they are able to highlight the representation expressed through various POVs (subject, object, audience, and media) to validate its claims. Its limitation lies on gathering the filmmaker's opinions only through social media, which could be improved if they had access to a personal interview.

Lastly, Rahmawati, & Nofiaturrehmah (2026) focuses on the poor education system and social condition, also analyzing it through Peirce's semiotics. In their discussion section, only three shots are analyzed from the film, with an additional example being the film poster. Thus, it could have been further explored by utilizing more examples from the film. After looking through previous research of *The Siege at Thorn High* (2025), it is clear that the importance of this new research lies in its discussion of film style. Previous research has only explored the themes of the film, while its visual and technical aspects have not been discussed. With the film's rapid pace, this new research will discuss how rhythmic editing forms tension and release through *The Siege at Thorn High* (2025).

## **2.2. THEORETICAL FOUNDATIONS**

### **2.2.1 RHYTHMIC EDITING**

Rhythmic editing is the act of manipulating time, movement, and energy through changes in timing, pacing, and trajectory phrasing (Pearlman, 2025, p. 13). Many editors have described the process of adding rhythm to a film as an intuitive process. However, Pearlman argues that this intuition can only be achieved with a certain amount of experience in the editing suite. This can be manifested through implicit learning from past edits and watching films, the ability to make judgments, sensitivity towards detail, creativity, and rumination. Other than that, an editor needs to be in touch with the rhythm of their daily lives and body. Once an editor understands these three types of rhythm, only then can they genuinely follow the rhythm of an actor's performance, scene, and the overall film.

### 2.2.1.1 TIMING

Timing has to do with reaching dramatic effectiveness through editing choices (Dancyger, 2019, p. 404). It involves the decision-making of where to place shots to achieve maximum impact. This can be achieved through three aspects, which are the frame, duration, and placement of the shot (Pearlman, 2025, p. 104). Through frame, the start and end frame of a shot can deliver different meanings. Pearlman gives three scenarios where a child is holding a glass that is about to fall. In the first scenario, the shot is cut before the glass leaves the child's hand, implying it as an accident. The second scenario shows the child's hand gripping and slowly letting go of the glass, making it seem intentional. For the third scenario, the editor cuts a few frames after the glass fell, leaving the shot with the child's frozen hand, implying that the child is in fear of consequence.

Other than that, timing can form rhythm through the choice of duration. The feeling of how long or short a duration of a clip is would depend on the relative duration of the shot before or after it. For example, a 10-second shot may feel short if juxtaposed with a 60-second shot, but long if juxtaposed with a series of 1 second shots. Other than the relativity of duration itself, it also depends on the information, movement, or change occurring within the shot (Pearlman, 2025, p. 106).

Moreover, timing also has to do with the placement of a shot in the context of when a certain information is revealed in the story to create tension and release. Pearlman (2025) highlights the importance of the editor's role in shaping the viewer's experience, who usually does not need to strictly follow the script (p. 108). The progression of a plot and understanding of a character's traits heavily depends on the order of shots to establish certain emphases or trajectories. This emphasis can also be portrayed through repetition of shots, and ending it with a different shot type to express a change.

### 2.2.1.2 PACING

Pacing is known as the tempo in film that is created through the rate of movement in a single or series of shots. It is the metronome in expressing how fast or slow a film progresses (Hullfish, 2024, p. 110). Pearlman (2025) refers musical tempo back to the heartbeat, where 60-80 beats per minute is defined as normal, and anything below or above it would be considered slow or fast (p. 110). Dancyger (2019) also describes pace as a way to formulate emotional response (p. 403). Fast pacing suggests intensity, while slow pacing is the opposite. Furthermore, Pearlman explains that the pace of a film can be identified through 3 aspects: the rate of cuts, change or movement within a shot, and overall change.

The rate of cutting or how often a cut occurs per second, minute, or hour is a determinant of pace (Pearlman, 2025, p. 111). It can occur as a pattern, such as increasing the rate of cutting as the plot reaches its climax. In other cases, it can occur without a specific pattern in duration. Pearlman demonstrates an example through cutting a conversation scene, where the arc of movement (beginning, middle, end) shapes the temporal element. If the editor cuts before the arc of movement is completed, it would make the cut feel sharper. On the other hand, letting the entire arc of movement flow would make the pace feel slower.

In relation to this, Salt (2009) developed the concept of Average Shot Length (ASL) which is used to calculate a film's rate of cutting by dividing the total film duration with the total number of shots. On a smaller scale, it can be calculated the same way for a scene, dividing the total duration of a scene with the total number of shots in it. This ASL becomes the baseline to determine the moderate pace of a film or scene, as it acts as the average. By connecting this to Pearlman's comparison of tempo with the heartbeat, the ASL would be the normal, while anything below it would be fast, and anything above would be slow.

Pace can also be determined through change or movement within a shot, or a shot's internal rhythm. Pearlman (2025) gives an example where two different 5-second shots may feel fast if lots of actions are happening in each of it (p. 112). Differently, if each shot has less changes within it, it would feel slower. By increasing the internal rhythm, more effort is needed to comprehend the shots. Not

only that, but pace also depends on the overall change in events, movements, or emotions throughout the film (Pearlman, 2025, p. 114). Pearlman describes the presence of fast pacing through several examples, the first one being a shot from *A Night at the Opera* (1935). As seen in Figure 2.1, there are four separate actions taking place in one frame, done to increase pace while also maintaining humor.



Figure 2.1. A shot from *A Night at the Opera* (1935).

(Source: Pearlman, 2025, p. 114).

A second example is seen through *Gosford Park* (2001), a film that feels fast paced due to frequent movements in characters, camera, and emotion, while the plot remains slow. Figure 2.2 shows a shot from the film of a person entering a car. The shot feels fast-paced because of the movements in characters taking place at the same time. It is heavily raining, a man helps a woman get into the car, and two people are waiting on either side of the vehicle. A third example is seen in Figure 2.3, where in *His Girl Friday* (1940), fast pacing is achieved through rapid-fire dialogues and quick change of events (Pearlman, 2025, p. 117). Interestingly, the film contains little to no camera movements and infrequent cuts.



Figure 2.2. A shot from *Gosford Park* (2001)

(Source: Pearlman, 2025, p. 116).



Figure 2.3. *His Girl Friday* (1940).

(Source: Pearlman, 2025, p. 116).

The fourth example is taken from Aespa's *Supernova* music video. This medium typically utilizes fast cuts and camera movements without needing a significant change in plot or character for it to feel fast paced. Figure 2.4 shows a sequence in the music video where the four clips are respectively shown one after the other. Each clip has a duration of 2 seconds and uses camera movements to keep the flow going. Overall, the rate of cutting continues to get faster, maintaining the energy throughout the entire music video (Pearlman, 2025, p. 117).

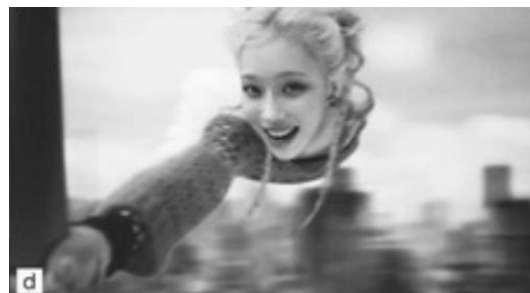


Figure 2.4. A sequence from Aespa's *Supernova* music video.

(Source: Pearlman, 2025, p. 117).

### 2.2.1.3 TRAJECTORY PHRASING

Trajectory phrasing refers to the assembling of different movements to create a specific flow of energy through certain choice of takes and cuts. The “energy” meant by Pearlman (2025) here refers to Laban and Bartenieff’s study of movement, which specifically talks about effort (p. 118). It explains that the way a person exceeds effort is the way they intend to deliver it through movement. For example, a punch can be aggressive, playful, or limping, and to do either one will require a different sort of effort. There are three ways in which trajectory phrasing can be achieved, which are through linking or colliding trajectories, selecting energy trajectories, and stress.

Linking trajectories is seen through movement, such as when one shot and the next uses the same screen direction to create a smooth cut. On the other hand, colliding trajectories may be seen as two shots with different screen directions that may create a feeling of shock. Other than movements, Pearlman (2025) explains Eisenstein’s ideas of collision through juxtaposing two opposing ideas that creates a new and foreign understanding for the viewer (p. 121). She notes some of the examples to be contrasts in dark and light, shot size, depths, graphic directions, and scales. In simpler terms, linking trajectories can be seen as continuity in editing, meanwhile colliding trajectories express a discontinuity in editing.

Selecting the right energy trajectories is also significant in shaping the energy of a scene, as it can describe the emotions and inner lives of the characters. For example, Pearlman illustrates a scenario where a couple is having an argument on the sofa and eventually the woman gets up to go to the kitchen, and the man follows. If there were two shots of each character, acted out with different energies (e.g., hesitantly and confidently), it would create different outcomes through different combinations. As a result, each choice would affect the way the viewer perceives the character’s lives and motivations (Pearlman, 2025, p. 122).

Lastly, trajectory phrasing can be shown through stress, which is emphasis or accent shown through energy from a certain shot. For example, a 2-second close-up shot of a scream and a 2-second wide shot of a sigh requires different efforts and produces different emphasis (Pearlman, 2025, p. 127). Even though it has the same

temporal property, the 2-second close-up of the person screaming would create a higher energy of stress. From this, it is up to the editor on which type of accent and where to place these accents in the film in to create the right energy.

### **2.2.2 TENSION AND RELEASE**

Tension and release are reoccurring cycles of questions, ideas, and information that form “an experience of the real” or in cinema as “cinema of the real” (Pearlman, 2025, p. 134). It is important to remember that the root word of ‘cinema’ is ‘kine’ which means movement. In creating this layer of reality, the rhythm of a film should be able to form cycles of tension and release as a way to synchronize with the viewer’s thoughts and experience. Three cycles operate simultaneously at different levels: movement of events at the macro level, movement of emotions at the mezzo level, and movement of images and sound at the micro level. While she uses these terms in describing each movement, it is also used to describe the structural duration of the film, through shot, scene, and the overall film.

#### **2.2.2.1 MOVEMENT OF EVENTS (MACRO LEVEL)**

Movement of events is related to changes in the macro level of a film. Pearlman states that the first frame of a film would cause the viewer to raise anticipation of a future event. This is where tension is built and may linger (eventually turning into suspense) for either a short or long time, with minimal or plenty of energy (Pearlman, 2025, p. 136). The moment the event takes place, tension is released. Additionally, Bordwell (2024) states that this process is done to keep the viewer engaged with the plot (p. 53).

From a bigger picture, the typical plot structure shapes an overall cycle of tension and release through five distinct stages: set-up, change, escalation, climax, and resolution (Pearlman, 2025, p. 13). It starts with the status quo (set-up) which introduces conflict or a disruption of the equilibrium (change). As the conflict develops and tension continues to rise, the narrative reaches an escalation stage. It reaches a climax where tension is highest, and the conflict may be solved or

unsolved, resulting in a new status quo (resolution). This pattern can also be applied on a mezzo level of a scene, or on a micro level to individual shots.

An example of its application on a micro level can be seen through a scene in *Matewan* (1987) as the character catches up to a train (as seen in Figure 2.5). Through the physical movement and expression portrayed by the actor, there is anticipation for him to catch the train. This creates tension and suspense, as questions arise as to whether or not he will make it, and what is at stake. At the end, this period of suspense is ultimately controlled by the rhythm of the edits (Pearlman, 2025, p. 136).



*Figure 2.5. A scene from Matewan (1987).*

*(Source: Pearlman, 2025, p. 136).*

#### **2.2.2.2 MOVEMENT OF EMOTIONS (MEZZO LEVEL)**

The mezzo level has to do with creating an effective emotional experience, with rhythmic editing choices as the stimulus for this. Editors may choose specific takes of an actor's performance, in which the intensity of their energy aligns with the intended emotional outcome of the scene. Pearlman (2025) highlights an internal tool that editors use to trigger this through the embodied simulation theory (p. 138). It talks about how humans respond to movements they see as though they are doing it themselves. Hence, the more relevant the emotion or experience is to them, the more intensely their brain responds to it. Amelia, Hakim, & Depita (2024) mention that this movement can be triggered through a relatable conflict that the character is facing, which also makes the plot more engaging (p. 442).

Furthermore, Pearlman (2025) gives an example on how cycles of tension and release are formed through a scene in *Notorious* (1946) using embodied simulation. The conflict focuses on Alicia, as she has to grasp Sebastian's wine cellar key to prove his involvement in a crime. The scene takes place in his office, with the key on his desk. It starts with a shot of Alicia's eyes and cuts to a shot that approaches the key, making it seem that she is close to grabbing it. However, to increase tension and hold suspense, the editor reveals that she is far away from it. Then, Sebastian is revealed to be near, though he does not see her or the keys. Tension arises as Alicia has to get across the room to grab those keys without getting herself caught and killed. There is a continued anticipation for her to successfully grab the keys to relieve the tension, and fear that she might not succeed (p. 139).

While embodied simulation is an internal tool, Pearlman also identifies an external tool used to stimulate movement of emotions. In her example of *Broadcast News* (1987), Pearlman (2025) explains how characters can contribute rhythm through their individual rhythmic functions (p. 153). The character Tom Grunnick functions as a "rest" in between the intense tension. On the other hand, Blair Litton drives the emotional tension of the scene by mentioning that they are running out of time which causes time pressure. This rhythmic function can be expressed through contrasting characteristics, such as difference in gestures portrayed by actors.

### **2.2.2.3 MOVEMENT OF IMAGE AND SOUND (MICRO LEVEL)**

The third movement occurs on a pre-cognitive level as it may be rapidly occurring through individual images and sound on a shot-to-shot basis. It runs on the micro level, such as a punch being thrown, raising ideas on how it will end, leading to the next shot where the tension is released. There is no time to think about how it may connect, but rather to just feel it in the moment. The same plot can engage or disengage rhythmically. An engaging film tells the journey of how one thing leads to another, and the speed or energy in which it is delivered. Leeuwen (as cited in Pearlman, 2025, p. 141) explains how rhythm can maximize viewer involvement

(tension) and passively abandon an event (release) by withholding and revealing information.

In forming movements of tension and release in image and sound, the editor has to reconsider the time, movement, and energy of the shot in relevance to the intended emotion. The editor must figure out different combinations of rhythmic tools through timing, pacing, and trajectory phrasing. With the right combination, tension is achieved through anticipating moments, visual and aural imbalance, while release happens through conflict resolution and stable visual and sound (Pearlman, 2025, p. 142).

With that in mind, movement of image and sound affects tension and release at the micro level. Pearlman (2025) emphasizes that these micro choices do not work alone, but the three movements work simultaneously across different levels (p. 143). Using the key scene in *Notorious*, tension is present during the shot revealing Alicia’s distance from the keys and juxtaposing the next shot of the husband nearby (micro level). The mezzo level is more concerned with tension and release in emotion of the scene delivered through Alicia’s facial expressions. She is determined to grab the keys and yet fear lingers as she realizes her husband is nearby. On the macro level, the event that unfolds presents a narrative stake.

### 3. RESEARCH METHOD

#### 3.1. DATA COLLECTION TECHNIQUES AND METHODS

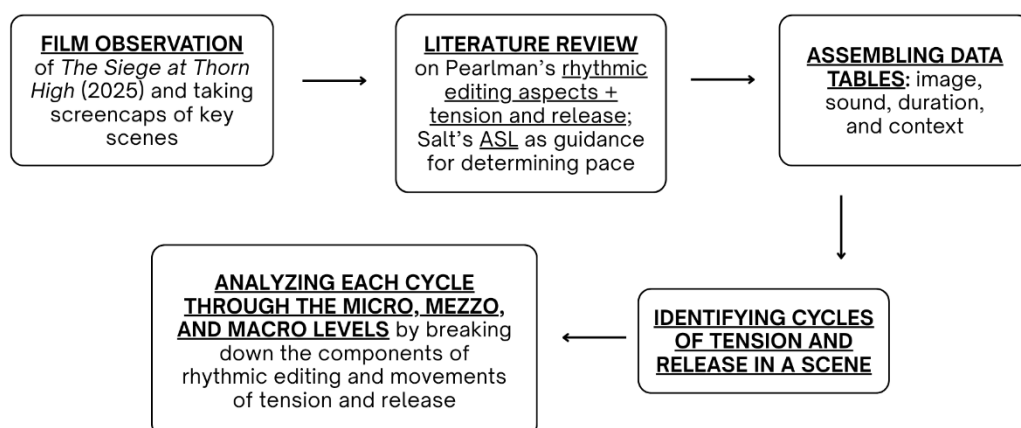


Figure 3.1. Research scheme.

(Source: Author)