



## APPENDICES



Scene 1: A big guy exercising				
Time	Moving images	Indonesian dialogue	English subtitles	Sound and sound effect (sfx)
03:05 – 03:07		<b>Fico:</b> <i>[grunting]</i>	[No subtitle]	Voice: Fico grunts while exercising, expressing exhaustion  SFX: Knockout boxing-style rock music
03:08 – 03:10		<b>Fico:</b> <i>Udah ah. Cukup.</i>	[No subtitle]	Voice: Flat-intonation  SFX: Knockout boxing-style rock music
<b>Vector shift:</b> Fico → tea pot and glass  <b>Shot size and angle:</b> Close shot → Medium shot. Frame 1 uses close shot, representing personal relationship between Fico and audience, and enabling audience to closely watch Fico exercising and see his physical condition. Frame 2 uses medium shot, representing a shift from personal relationship to social relationship between Fico and audience. Both frames are shot in frontal, eye-level			<b>Back-translation:</b>  [no subtitle]	

<p>angle.</p> <p><b>Gaze:</b> No direct gaze at audience.</p> <p><b>Composition:</b> Given (tea pot) – New (Fico)</p>	
<p><b>Speech-Text-Image Relation</b></p>	<p>This scene is a continuation of preceding scene which depicts Fico practicing boxing, yet eventually hurting himself in a ridiculous way. The first frame illustrates Fico doing sit up, shot in close shot which enables audience to observe him closer. However, in the second frame, it is revealed that Fico does not actually do the sit up. The frame shows that a woven mat is nailed on the wall, with properties such as tissue pack and plastic bottle glued to the mat. Fico leans against the mat-covered wall and shows ‘sit up’ gesture as if he is doing real sit up. Feeling tired, Fico eventually ends his exercise and walks to have a glass of water.</p> <p>Based on the analysis of visual aspect, the vector shifts from Fico (frame 1 and 2) to teapot (frame 2). After exercising, Fico walks towards a table (positioned in the left) and has a glass of water. Fico does something to the tea pot; he walks towards it and grab it. Thus, Fico becomes the actor, the participant “from which the vector emanates, or itself, in whole or in part, forms the vector” (Kress &amp; van Leeuwen, 2006, p. 63). Meanwhile, the tea pot becomes the goal of Fico’s action, the participant “to whom or which the action is done” (Kress &amp; van Leeuwen, 2006, p. 64). The vector shift helps illustrating what Fico does after ‘exercising’.</p> <p>In presenting the situation, the scene employs two different frame sizes. In the first frame, the situation is presented in close shot, enabling audience to be ‘closer’ to Fico and observing him. Since in the first frame it hasn’t been revealed that Fico does not really do the sit up, the close shot allows audience to closely observe Fico’s physical condition. However, in the second frame, the situation is presented in medium shot, thus revealing the ‘truth’ that Fico does not actually do sit up. Audience are not as close to Fico as they were in the first frame; however, the revelation—what Fico actually does is still able to make audience laugh. Moreover, the scene is shot in frontal, eye-level angle. Frontal angle implies involvement—what is depicted is the part of represented participants’ and audience’s world (Kress &amp; van Leeuwen, 2006). Meanwhile, the eye-level angle implies the absence of power difference involved, positioning represented participant and audience in the equal level (Kress &amp; van Leeuwen, 2006). Furthermore, Fico does not look directly at the camera (audience), implying that there is no gaze at audience. The absence of direct gaze at audience implies that what is seen in the scene is offered as information, positioning the represented participants in the scene as the object audience are viewing (see Kress &amp; van Leeuwen, 2006). Since there is no gaze at audience, they are ‘offered’ to see and observe the situation. These visual aspects (frame size, gaze, angle) allows audience to be involved in the situation by closely</p>

‘watching’ Fico exercising and knowing what Fico does after exercising—drinking a glass of water to quench the thirst, just like everyone else.

The scene employs left – right layout composition, as dominantly depicted in the second frame. According to Kress and van Leeuwen (2006:181), elements positioned at the left side is a *given* information and elements positioned at the right side is a *new* information. Thus, the tea pot becomes a given information, while Fico becomes a new information. Based on the situation, tea pot represents information audience may have already know—drinking as something people commonly do after exercising. On the other hand, Fico represents new information audience have to focus on. In addition to this, the tea pot is depicted as a blurred object while, in contrast, Fico is depicted clearly. In other words, Fico is more salient than the tea pot. Saliency in image can build a hierarchy of importance among the participants or elements in an image, despite the placement of the elements (Kress & van Leeuwen, 2006). Thus, in the second frame, Fico is more important than the tea pot because he is more salient than the tea pot. This may be due to the fact that Fico exercises in an unusual way, which becomes a *new* information. Furthermore, by being new information, audience are expected to focus on Fico, whose action contains humour.



There is only one line of dialogue in the scene. The dialogue line “*Udah ah. Cukup*” (loosely translated as “Okay. Enough”) corresponds to what is depicted in the scene and, more or less, conveys the same meanings as in the moving images. This correspondence signalizes elaborative relationship between the dialogue and the image. According to Barthes (1977, in Kress and van Leeuwen, 2006), the relationship between the verbal element and the image is elaboration when the verbal element elaborates the meanings in the image, and vice versa, restating the same meanings yet in different way. However, the dialogue is not subtitled and this may leave audience confused. Even though the spoken verbal mode exists in the scene, however, audience would need the written verbal mode (in this case, subtitles) to function as the translation of the message rendered through the spoken verbal mode. The absence of subtitles results on meanings distributed to mainly visual and audio mode and implies that reduction has been done to the subtitling.

Scene 2: Junior high school student extorted money by a big guy				
Time	Moving images	Indonesian dialogue	English subtitles	Sound and sound effect (sfx)
04:26 – 04:28		<b>Iqbal:</b> <i>Wei, wei, wei, wei! Gue barusan dipalak, gue malah untung goceng.</i>  <b>Student 2:</b> <i>Ha? Di mana?</i>	[04:26 – 04:27] Someone just tried to extort money from me, yet I ended up making profit from him  [04:27 – 04:28] Where?	Voice: Iqbal uses high intonation, expressing his excitement.  SFX: (no SFX)
04:29 – 04:31		<b>Iqbal:</b> <i>Noh, noh, noh, sama gajah Afrika, noh!</i>  <b>Student 2, 3, 4:</b> <i>Ayo! Ayo! Ayo!</i>	[04:29 – 04:32] There. The big guy over there.  [04:32 – 04:33] Come on. Let's go!	Voice: Iqbal points at a big guy, shouting. Student 2, 3, and 4 yell excitedly.  SFX: (no SFX)
<b>Vector shift:</b> Student 2, 3, 4 ← → Iqbal  <b>Shot size and angle:</b> Medium shot represents regular social relationship between the students and audience, implying that the scene is a representation of common daily life. The angle is oblique and eye-level, putting audience in the equal position to the students, yet detaching audience from them.  <b>Gaze:</b> No direct gaze at audience, implying that audience only observe the situation			<b>Back-translation:</b>  <b>Iqbal:</b> <i>Seseorang mencoba merampas uang saya, tapi saya justru dapat untung dari orang tersebut.</i>  <b>Student 2:</b> <i>Di mana?</i>  <b>Iqbal:</b> <i>Di sana. Pria besar itu.</i>	

<p>portrayed in the scene.</p> <p><b>Composition:</b> Center (Iqbal) – Margin (student 2, 3, 4)</p>	<p><b>Students 2, 3, 4:</b> <i>Ayo pergi!</i></p>
<p><b>Speech-Text-Image Relation</b></p>	<p>The scene illustrates a junior high school student, Iqbal, who is extorted by a big guy (Fico), yet eventually gets money from him. Iqbal tells his friends about what happened to him and, since he told them that he got money instead of losing one, his friends become excited and want to get money as well.</p> <p>Based on the analysis of visual aspects, the vector shifts from Iqbal’s friends to him, and vice versa. In other words, both participants (Iqbal and his friends) emanate and receive vector. However, since the vector is realized through the eyeline or gaze of the participants in the scene, the participants become reactor and phenomenon at the same time. Reactor refers to the participant who does the looking and is able to show facial expression (Kress &amp; van Leeuwen, 2006). Meanwhile, phenomenon refers to the participant whom the reactor looks at (Kress &amp; van Leeuwen, 2006).</p> <p>The situation in the scene is presented in medium shot. Medium shot reflects the distance between represented participants and audience, which is social (see Kress &amp; van Leeuwen, 2006). This implies that the relationship between audience and participants (junior high school students) in the scene is social. Furthermore, participants do not show direct gaze at audience, implying that audience only observes the situation. The absence of direct gaze at audience implies that what is seen in the scene is offered as information, positioning the represented participants in the scene as the object audience are viewing (see Kress &amp; van Leeuwen, 2006). The situation is also shot in oblique, eye-level angle, putting audience in the equal position to the participants, yet detaching them from the participants. When the image is presented at eye-level, the represented participants and audience is in an equal relation (Kress &amp; van Leeuwen, 2006). However, oblique angle represents the image shown as something audience are not involved in (Kress &amp; van Leeuwen, 2006). Those aspects (frame size, gaze, and angle) presents the situation in the scene as a representation of common daily life; the situation may happen in real life and it is common, yet audience is not involved in the situation, merely observing it.</p> <p>The scene employs center – margin composition, represented by Iqbal and his friends. According to Kress and van Leeuwen (2006), when an element is placed at the center of the image, that element becomes the center or the core information. In contrast, the margins are supporting elements positioned around the center which depend on the center element (Kress &amp; van Leeuwen, 2006). Based on the situation, Iqbal represents a center information, implying that what is encountered by him (extortion) is something which should be focused on. On the other hand, Iqbal’s friends represent the margins, implying that they are just supporting elements for the center since Iqbal is already the core information. The positioning of Iqbal as a center information and his friends as the margins reflects what audience might have seen in real</p>

life. It is common that when a person brings a new information, there might be possibilities that his/her peers would surround him/her to get the information. The person conveying the information automatically becomes the focus or the center of attention, while the others merely listen to him/her.

Regarding the relationship between visual mode and verbal mode, the dialogue lines uttered by the characters in the scene corresponds to the gestures and facial expression shown by the characters. This implies that the dialogue in the scene functions to elaborate the meanings in the image by restating what is conveyed by the moving images (see Kress & van Leeuwen, 2006). From the perspective of non-Indonesian audience, subtitles as verbal mode in this scene also functions as an elaboration to the visual mode. According to Barthes (1977, in Kress & van Leeuwen, 2006), verbal element may function to elaborate the image, and vice versa, restating the same meaning in different way. The subtitles serves as the target language transcription of the source language dialogue. On the other hand, the image also elaborates the subtitles, providing illustration of the situation based on the subtitles.

Scene 3: The robbers entering the bank				
Time	Moving images	Indonesian dialogue	English subtitles	Sound and sound effect (sfx)
08:33 – 08:35		<b>Robbers:</b> <i>Tiarap! Tiarap! Tiarap!</i>	[08:33 – 08:34] Everyone stay on the ground!	Voice: Robbers yell, forcing bank customers and tellers to stay on the ground.  SFX: people screaming in panic, gun shots, surf-rock music in background
08:35 – 08:37		<b>Ernest:</b> <i>Tiarap semuanya!</i>	[08:35 – 08:37] Do not move!	Voice: Robbers yelling fiercely  SFX: people screaming in panic, gun shots, surf-rock music in background
<b>Vector shift:</b> Robbers → Bank customers and bank tellers  <b>Shot size and angle:</b> Medium shot (frame 1); medium shot (fr. 2, panel 1); long shot (fr. 2 panel 2). The first frame is shot in medium shot size. However, in the second frame, the first panel is shot in medium close shot and the second panel in long shot, with bigger panel size. The angle in the first frame is oblique, low angle. In the second frame, the first panel is shot in oblique, eye-level angle, while the			<b>Back-translation:</b>  <b>Robbers:</b> <i>Semuanya tiarap!</i>  <b>Ernest:</b> <i>Jangan bergerak!</i>	





<p>second panel is shot in oblique, medium-high angle.</p> <p><b>Gaze:</b> No direct gaze at audience</p> <p><b>Composition:</b></p> <ul style="list-style-type: none"> <li>- Frame 1: Center (Ari) – Margin (Doors)</li> <li>- Frame 2 – panel 1: Given (Ernest) – New (Bank customers)</li> <li>- Frame 2 – panel 2: Given (Bank customers) – New (Ernest)</li> </ul>	
<p><b>Speech-Text-Image Relation</b></p>	<p>The scene illustrates three robbers (played by Ari Kriting, Kemal, and Ernest) entering the bank, threatening bank customers and bank tellers inside. The robbers are gun-armed. They start shooting and tell the customers and tellers to stay to the ground.</p> <p>Based on the analysis of visual aspects, the vector shifts from the robbers to the bank customers and bank tellers. The vector is realized through the rifle and the gaze of the robbers. The participants in the scene (the robbers) do not show physical contact with the bank customers and bank tellers, merely pointing their rifles at them. Since the vector is realized through gaze, the narrative process in this scene is relational (see Kress &amp; van Leeuwen, 2006, p. 67). In relational process, the represented participants who emanate the vector is called reactor, instead of actor, and the participant who is aimed at the vector is phenomenon, instead of goal (ibid.). Thus, the robbers become the reactor and the bank customers and the bank tellers become the phenomenon, the people whom the robbers look at.</p> <p>The scene is shot in two different frame size: medium and long shot. The first frame and the first panel of the second frame uses medium shot, while the second panel of the second frame uses long shot. Moreover, the situation in both frames is presented at oblique level. However, different vertical angles are used to depict the situation. The first frame uses low angle. Meanwhile, the first panel of the second frame use eye-level angle, and the second panel use medium-high angle. Medium shot reflects the distance between represented participants and audience, which is social (see Kress &amp; van Leeuwen, 2006). Meanwhile, long shot represents impersonality (ibid.). In the first frame and the first panel of the second frame, audience are positioned close enough to the robbers that they can see the robbers clearly. In the second panel of the second frame, the audience is positioned further, enabling them to see the robbery from a wider perspective. The long shot allows audience to see the frightened bank customers and bank tellers, as well as the robbers. At oblique level, audience are not involved in the situation depicted in the scene (see Kress &amp; van Leeuwen, 2006). However, the</p>

low angle used in the first frame implies that the represented participants (the robbers) have power over the audience (or at least, they have more power than the audience). According to Kress and van Leeuwen (2006), when an image is presented at low angle, the represented participants in the image have power over the viewer. The first panel in the second frame uses eye-level, implying that the relationship between the robbers and the audience is equal (see Kress & van Leeuwen, 2006). By being in an equal relationship, audience are able to identify the robbers as characters who has no power difference to them, despite the fact that the robbers commit crime. In addition, the panel is shot in medium shot and at oblique angle that helps audience observe and recognize the robbers in a close distance, without involving them in the situation. At last, the second panel of the second frame uses medium-high angle, allowing audience to see the situation from higher perspective. Medium-high angle positions audience as having more power than the represented participants (the robbers and the bank customers). When an image is presented at high angle, the represented participants in the image have less power than the viewer (Kress & van Leeuwen, 2006). Integrated with long shot size and oblique angle, the situation in the second panel is seen from the third-person omniscient point of view, positioning audience as God-alike interactive participant who only knows and observes the situation without being involved in it. This assumption is also supported by the fact that the represented participants do not show direct gaze at the audience. The absence of direct gaze at audience implies that what is seen in the scene is offered as information, positioning the represented participants in the scene as the object audience are viewing (see Kress & van Leeuwen, 2006).

For this scene, two different layout compositions are employed. The first frame uses Center – Margin layout composition. The robber (represented by Ari Kriting) is positioned at the center of the image, between the doors. An element placed in the center is the core information, while elements placed in the left or right side of the central element are called Margin (see Kress & van Leeuwen, 2006). Integrated with the shot angle, the composition positions the robbers as the central information or characters whom audience should focus on and acknowledge the robbers' power over them. The second frame uses Given – New layout composition, as illustrated in both panels. According to Kress and van Leeuwen (2006), elements positioned in the left is a *given* information and elements positioned in the right is a *new* information. However, in the first panel of the second frame, the robbers (represented by Ernest) are positioned as the given information. This might be due to the correspondence to the first frame which 'introduces' the robbers and characters or represented participants in the scene. Meanwhile, in the second panel the robbers are positioned as new information. The positioning might be due to the plot of the story. The first frame and the first panel of the second frame functions to introduce the represented participants (the robbers). However, the second panel of the second frame does not function as an introduction to the participants; rather, it shows what the participants do. Thus, the first frame and the first panel of the second frame can be verbalized in a sentence as "These are the robbers", while the second panel of the second frame can

be verbalized as “The robbers rob the money from the bank.”



Observing the relation between the verbal mode and the non-verbal mode, both dialogue and subtitles in this scene corresponds to the non-verbal modes. The dialogue “Tiarap!” (Stay to the ground) is translated as “Stay to the ground!” and “Do not move!” which are corresponds to the illustration provided, depicting the robbers entering the bank and urging bank customers and bank tellers to stay to the ground. However, it is worth to note that in the second frame, the subtitles “Do not move!” do not have similar meanings to the dialogue. This implies that there is a transformation of subtitles which, according to Chuang (2009) may occur due to the selection of meanings from the non-verbal elements. Nevertheless, either the verbal mode or the non-verbal mode express the same meaning. According to Barthes (1977, in Kress & van Leeuwen, 2006), verbal element may function to elaborate the image, and vice versa, restating the same meaning in different way. Thus, in this scene the relationship between the verbal mode and the non-verbal modes is elaboration.

Scene 4: Old man singing with his guitar				
Time	Moving images	Indonesian dialogue	English subtitles	Sound and sound effect (sfx)
28:00 – 28:02		<b>Mudy (old man):</b> [singing] <i>Dulu, dulu, dul, dul, dul, dul, dul</i>	[no subtitles]	Voice: Mudy sings a dangdut song while playing guitar. The Indonesian word ‘dulu’ is repeated in a funny way.  SFX: guitar strumming
28:02 – 28:04		<b>Mudy &amp; Ari Kriting:</b> [singing together] <i>Dulu aku suka padamu, dulu aku memang suka. La, la, la.</i>	[no subtitle]	Voice: Mudy and Ari sing the dangdut song together.  SFX: guitar strumming
<b>Vector shift:</b> Ari → Mudy (old man)  <b>Shot size and angle:</b> Medium close shot → Medium shot. The angle is oblique and eye-level. Both frames are shot in tilted angle.  <b>Gaze:</b> No direct gaze at audience.  <b>Composition:</b> Given (old man) – New (Ari Kriting)			<b>Back-translation:</b> [no backtranslation]	

<p><b>Speech-Text-Image Relation</b></p>	<p>The scene illustrates Mudy, an old man who comes at the moment of bank robbery. Ari Kriting, one of the robbers, comes to him and asks what he does, initially urging him to leave but only to ask the old man to sing for him. Playing his guitar, the old man then sings a dangdut song which amazes Ari. Eventually, Ari Kriting joins him singing and dancing with delight.</p> <p>Based on the analysis of visual aspects, the vector shifts from Ari to Mudy. The vector is realized through Ari's gaze to Mudy. Since the vector is realized through gaze, the narrative process in this scene is relational (see Kress &amp; van Leeuwen, 2006). In relational process, the participants who emanate the vector is called reactor, instead of actor, and the participant who is aimed at the vector is phenomenon, instead of goal (ibid.). Thus, Ari becomes the reactor and Mudy becomes the phenomenon, the person who makes Ari amazed.</p> <p>The situation in the scene is initially presented in medium close shot (frame 1), enabling audience to be closer to Mudy and Ari. However, in the second frame the situation is presented in medium shot, a little further than the one in the first frame but still allows audience to observe the situation closely. Medium shot reflects the distance between represented participants and audience, which is social (see Kress &amp; van Leeuwen, 2006). In this scene, the relationship between the participants and audience is social. By being social, the represented participants (Ari and Mudy) are recognized as characters or people who are common in daily life, even though audience do not necessarily know them well. Moreover, participants do not show direct gaze at audience. The absence of direct gaze at audience implies that what is seen in the scene is offered as information, positioning the represented participants in the scene as the object audience are viewing (see Kress &amp; van Leeuwen, 2006). This implies that audience are not further involved in the situation (for instance, joining the group singing), merely observing. The situation is also shot at oblique, eye-level angle. Oblique angle represents the image shown as something audience are not involved in (Kress &amp; van Leeuwen, 2006). Meanwhile, an image shot at eye-level reflects the equality between the represented participants and audience (Kress &amp; van Leeuwen, 2006). This implies that the angle used puts audience in the equal position to the participants, yet detaching them from the participants. Overall, the aspects (frame size, gaze, and angle) present the situation in the scene as a representation of common daily life; it is common to see two people singing dangdut song together. However, audience are not involved in the situation; they just observe the situation.</p> <p>The given – new layout composition is employed in this scene, represented by Mudy and Ari. According to Kress and van Leeuwen (2006), elements positioned in the left is a <i>given</i> information and elements positioned in the right is a <i>new</i> information. Based on the situation, Mudy represent a <i>given</i> information. On the other hand, Ari represents a <i>new</i> information. Mudy represents a character whom audience can generally recognize in real life—an old man wearing</p>
--	---

traditional Javanese clothes and *blangkon*. Moreover, the camera is not focused on Mudy in frame 1, shifting the focus on Ari since Mudy represents a character or information which is more likely already known. Positioned in right, Ari represents *new* information. With the camera focus on him, Ari is expected to be the focus of audience. Furthermore, Ari represents a robber, a character who may not be something new in audience's lives but, at least, do not appear regularly.

Regarding the text-image relationship in this scene, it can be said that what is conveyed in the lyrics is not reflected in the moving images. The song sung by the characters tells about a person who used to love a girl but eventually no longer loves her. The gesture (dance) and facial expression shown by the characters clearly do not reflect the feelings in the song. However, the gesture and the facial expression corresponds to the sound (in this case, the music arrangement). This means that each modes (verbal and non-verbal) add different meanings to the scene, indicating that the text-image relationship in this scene is extension. According to Barthes (1977, in Kress & van Leeuwen, 2006), when new or different meanings are added to complete the message, the relationship between the verbal and the non-verbal mode is extension. However, from the perspective of non-Indoensian audience, subtitles as written verbal mode is not shown this scene. Due to the absence of subtitles, the meanings in the dialogue (in this case, the song lyrics) are distributed to the visual and audio mode. In this scene, the audio mode plays important role since the scene illustrates two singing men. The catchy guitar strums and energetic vocal arrangement reflect the mood of the song, even though there is not a translation for the lyrics sung. Moreover, what is depicted through visual mode (happy expressions and dancing gestures) intensifies the mood in the scene. In general, the visual and audio mode in this scene function to substitute the function of subtitles to convey the meanings. In addition to this, the absence of the subtitles implies that reduction has been done to the subtitling.

Scene 5: Mungil				
Time	Moving images	Indonesian dialogue	English subtitles	Sound and sound effect (sfx)
30:16 – 30:18		<b>Mungil:</b> <i>Taaaa!</i>	[30:16 – 30:18] Tada!	Voice: Mungil speaks in a girly, excited way. The voice is reverberated  SFX: surf-rock music in background
30:19 – 30:21		<b>Mungil:</b> <i>Halo! Saya Mungil!</i>	[30:19 – 30:21] Hello, I'm so cute and tiny.	Voice: Mungil introduces himself in a girly way. The voice is reverberated.  SFX: surf-rock music in background
<p><b>Vector shift:</b> Mungil → Mungil's face (frame 1); Mungil → audience (frame 1 &amp; 2)</p> <p><b>Shot size and angle:</b> Close up → Medium long shot. The first frame is shot in frontal, eye-level angle, while the second frame is in oblique, medium-low angle.</p> <p><b>Gaze:</b> Direct gaze at the audience.</p>			<p><b>Back-translation:</b></p> <p><b>Mungil:</b> <i>Tada!</i></p> <p><i>Halo. Saya mungil dan menggemaskan.</i></p>	



<p><b>Composition:</b></p> <ul style="list-style-type: none"> <li>- Frame 1: Given (blank space) – New (Mungil)</li> <li>- Frame 2: Center (Mungil) – Margin (blank space)</li> </ul>	
<p><b>Speech-Text-Image Relation</b></p>	<p>The scene introduces Mungil, a bank robber who later joins the other robbers and forms a group with them. Unlike the other robbers, Mungil is effeminate and dresses provocatively. Regardless her behavior, Mungil is one of the strongest and smartest robbers, armed with a rifle and mastering martial art.</p> <p>Based on the analysis of visual aspect, in the first frame the vector shifts from Mungil’s index fingers to his face. The vector is realized through his gesture (he uses both his index fingers to point at his face). Mungil does something to his face (he points at his face); therefore, Mungil becomes the actor, the participant “from which the vector emanates, or itself, in whole or in part, forms the vector” (Kress &amp; van Leeuwen, 2006) and his face becomes the goal, the participant the vector aims at (Kress &amp; van Leeuwen, 2006). However, in the second frame, Mungil maintains his gesture (his index fingers still point at his face) and looks at the audience. In this frame, the vector shifts from Mungil to the audience. The narrative process in this frame is relational because the vector is realized through gaze or eyeline (see Kress &amp; van Leeuwen, 2006). Thus, in the second frame Mungil becomes the reactor and the audience becomes the phenomenon.</p> <p>The situation in the first frame is shot in close shot and at frontal, eye-level angle. Close shot (or close up) represents intimate distance between the represented participants in the image and the viewer (see Kress &amp; van Leeuwen, 2006). In addition to this, frontal angle implies involvement—what is depicted is the part of represented participants’ and audience’s world (Kress &amp; van Leeuwen, 2006). Moreover, when the image is presented at eye-level angle, the image represents equality between the represented participants and the viewer (Kress &amp; van Leeuwen, 2006). The integration of these aspects (frame size, horizontal and vertical angle) implies that the relationship between Mungil and the audience is personal and equal; Mungil is neither more powerful nor less powerful than the audience. The frame size used in this frame allows the audience to be involved and know Mungil personally. In line with this, Mungil shows direct gaze at the audience. When there is a direct gaze or direct address to the viewer, the image represents a demand (see Kress &amp; van Leeuwen, 2006). However, in the second frame, the situation is shot in medium-long shot and at oblique angle. The size shift positions the audience further than they were in the first frame and the oblique angle used represents detachment (see Kress &amp; van Leeuwen, 2006). Thus, in the second frame the audience are not involved anymore in the interaction with Mungil, but they are able to see Mungil’s appearance from top to toe. However, Mungil still demands the audience to interact with him. In addition to this, Mungil is presented at low angle, representing his power over the viewer (see Kress</p>





& van Leeuwen, 2006). In general, it can be said that Mungil demands the audience to get to know him and be closer to him, trying to establish an equal, intimate relationship with the audience. Nevertheless, the audience are eventually positioned further, no longer involved in the interaction with Mungil, and are seen to have less power than Mungil. This might be due to the fact that Mungil is a weird character and in real life, some people still have think that effeminacy is a stigma. Thus, after interacting with Mungil in a very close distance and observing him, the audience are positioned away from him in order to end the involvement in the interaction with him, even though Mungil still demands an interaction. Moreover, Mungil is a robber who has more power than the audience, regardless the effeminacy.

This scene employs two different layout compositions: given-new and center-margin composition. The first frame uses given-new composition. According to Kress and van Leeuwen (2006), elements positioned in the left is a *given* information and elements positioned in the right is a *new* information. Mungil is positioned in the right part of the screen; therefore, Mungil represents a *new* information. In addition to this, Mungil is also depicted as a salient element. As a new, salient information, Mungil is expected to receive more attention from the audience, since the scene is about introducing Mungil as a new character in the film. Moreover, in the second frame Mungil is positioned at the center of the frame. According to Kress and van Leeuwen (2006), if an element is placed in the center, the element becomes the 'center' or the core information. In the second frame, Mungil is no longer a new information. However, he is still the element the audience have to focus on.

In this scene, the subtitles provides the translation of Mungil's dialogue. However, there is a problem in the translation. In the second frame, Mungil says "Halo! Saya Mungil!" which can be loosely translated as "Hello! I am Mungil" or, alternatively, "Hello! My name is Mungil!" but the dialogue is translated as "Hello! I'm so cute and tiny!" Indonesian word 'mungil' is an adjective referring to something which is tiny and pretty, equivalent to English word 'cute'. This transformation, according to Chuang (2009), may be the result of the selection of meanings from the moving images. However, in the context of the film, the word 'mungil' refers to the name of a character. Since subtitles are used as the guide for non-Indonesian audience to comprehend the film, the subtitles become the main verbal mode. Observing the relation between the verbal and the non-verbal mode, what is conveyed in the dialogue corresponds to what is depicted in the scene. Mungil's greetings fits his gesture and facial expressions. The correspondence indicates that the relationship between the dialogue and the moving images in this scene is elaboration. When both verbal and non-verbal modes convey the same meanings, yet restating them in a different way, the relationship between the modes is elaboration (see Barthes, 1977, in Kress & van Leeuwen, 2006). Moreover, from the perspective of non-Indoensian audience, it can be assumed that what is presented through the image and sound also elaborates the meanings in the subtitles. Mungil's gesture and voice intonation corresponds to the subtitles, regardless the problem in the translation. This means that for this scene, both



	dialogue and subtitles elaborate the meanings in the moving images.
--	---

Scene 6: The shooting				
Time	Moving images	Indonesian dialogue	English subtitles	Sound and sound effect (sfx)
31:25 – 31:26		<b>Police officer:</b> <i>Tembak!</i>	[no subtitles]	Voice: A police officer orders to shoot, speaking through a megaphone  SFX: police siren, gunshot sound, surf-rock music in background
31:26 – 31:27		[no dialogue]	[no subtitles]	Voice: [no speaking character]  SFX: breaking window panes, gunshot sound, surf-rock music in background
<b>Vector shift:</b> Police officer → Window panes  <b>Shot size and angle:</b> Both frames are shot in close shot, enabling audience to see the details of the shooting. The angle in the first frame is oblique, with the object (rifle) positioned a little higher than audience. In the second frame, the object (window) is shot in oblique, eye-level angle.  <b>Gaze:</b> No direct gaze at audience			<b>Back-translation:</b> [no backtranslation]	

<p><b>Composition:</b></p> <ul style="list-style-type: none"> <li>- Frame 1: Given (blank space) – New (Police officer)</li> <li>- Frame 2: Center (Window panes) – Margin (Window jamb)</li> </ul>	
<p><b>Speech-Text-Image Relation</b></p>	<p>The scene illustrates the cops surrounding the bank, preparing to attack the robbers. They eventually start firing and shooting the bank windows.</p> <p>The vector shift in this scene is illustrated in two frames. The vector shifts from the police officer to the window. Since the police officer does something which physically affects the window, the narrative process in this scene is action process, positioning the police officer as the actor, the participant “from which the vector emanates, or itself, in whole or in part, forms the vector” (Kress &amp; van Leeuwen, 2006, p. 63) and the window as the goal, the participant the vector aims at (ibid.). The situation can be verbalized as a following sentence: “The police officer shoots the window.”</p> <p>Both frames constituting the scene use close shot, enabling audience to observe the shooting closely. The situation is also shot at oblique angle. Oblique angle represents detachment (see Kress &amp; van Leeuwen, 2006); therefore, the audience are not involved in the situation depicted in the scene; merely observing the situation. In addition to this, two different vertical angles are employed to depict the scene. In the first frame, the situation is depicted at medium-high angle. Based on Kress and van Leeuwen (2006), the use of low angle implies that the represented participant (in this case, the police officer) has power over the viewer. Meanwhile, in the second frame the situation is depicted at eye-level angle, implying that the relationship between the represented participant (in this case, the window) and the viewer is equal (Kress &amp; van Leeuwen, 2006). In general, audience are able to see the shooting closely. They can see the police officer as a participant who is more powerful than them, meanwhile the bank (which represents bank customers and bank tellers, including the robbers inside) is positioned as having similar or equal power to the audience. In addition to this, the absence of direct gaze to viewer represents offers (see Kress &amp; van Leeuwen, 2006). Thus, the audience are only offered to observe the situation and are not involved in the shooting.</p> <p>There are two different layout compositions used in this scene. The first frame employs given – new composition. The police officer is positioned in the right part of the frame, thus becoming a new information (see Kress and van Leeuwen, 2006). Relating to the next frame, the police officer shoots the window which, logically, is positioned at the left. Thus, the window becomes a given information. This implies that the window is something the audience already know and the police officer, since he is a new information, is expected to gain attention from the audience. That way, the audience will</p>

focus on what the police officer does. However, the second frame uses center – margin composition. The window (particularly, window panes) is positioned at the center of the frame. According to Kress and van Leeuwen (2006), the element placed in the center is the ‘center’ or the core information. By being the core information, the window should be the focus of the audience, allowing them to see the result of the police officer’s action.

The dialogue line uttered in this scene (“Tembak!”, which means “Fire!”) corresponds to what is illustrated. The scene illustrates police officer firing at a bank, in which imperative order “Fire!” becomes the most logical dialogue line to accompany the action. Since the dialogue conveys the same meanings as in the illustration, the relationship between the dialogue and the moving images is elaboration. According to Barthes (1977, in Kress & van Leeuwen, 2006), the relationship between the verbal element and the image is elaboration when the verbal element elaborates the meanings in the image, and vice versa, restating the same meanings yet in different way. However, the dialogue is not translated into subtitles. This results on the distribution of meanings to non-verbal modes, since the scene is devoid of subtitles (the written verbal mode), which serve as the guide for non-Indonesian audience. The absence of subtitles may also indicate the reduction done in the subtitling. The Indonesian command “Tembak!” is not translated to English. However, audience may understand the situation depicted in the scene by grasping the meanings from the image (the police officer shooting the window) and the sound (police siren, gunshot sound, breaking glass). Without the non-verbal modes, non-Indonesian audience might find it difficult to understand the situation.



Scene 7: Receiving ammunition				
Time	Moving images	Indonesian dialogue	English subtitles	Sound and sound effect (sfx)
33:57 – 33:58		<b>Mungil:</b> <i>Tangkap, mang!</i>	[no subtitles]	Voice: Mungil shouts in a girly manner  SFX: gunshot sound, surf-rock music in background
33:59 – 34:00		[no dialogue]	[no subtitles]	Voice: [no speaking character]  SFX: gunshot sound, surf-rock music in background
<b>Vector shift:</b> Mungil → Mudy  <b>Shot size and angle:</b> The scene is shot in medium shot. Both frames constituting the scene are shot at oblique angle. However, in the first frame the participant is presented at medium-low angle, while in the second frame the participant is presented at eye-level angle.  <b>Gaze:</b> No direct gaze at audience			<b>Back-translation:</b> [no backtranslation]	

<b>Composition:</b> Given (Mudy) – New (Mungil)	
<b>Speech-Text-Image Relation</b>	<p>The scene illustrates two robbers, Mungil and Mudy (who previously disguised as an old man), shooting at the police officers. In the preceding scene, the police officers surround the bank and are commanded to shoot the bank in order to threaten the robbers. However, the bank robbers counterattack the officers instead, shooting them away. As Mudy runs out of ammunition, Mungil throws him an extra ammunition to use.</p> <p>Based on the visual analysis, the vector shifts from Mungil to Mudy. The vector is realized through the action of throwing the ammunition. Since the vector is realized through action which involves physical contact (Mudy receiving ammunition from Mungil), the narrative process in this scene is action process, positioning the police officer as the actor, the participant “from which the vector emanates, or itself, in whole or in part, forms the vector” (Kress &amp; van Leeuwen, 2006, p. 63) and the window as the goal, the participant the vector aims at (ibid.). The situation depicted in the scene can be verbalized into the following sentence: “Mungil gives ammunition to Mudy.”</p> <p>Both frames in this scene use medium shot and oblique angle. The medium frame size allows audience to observe the situation in a distance that is close enough for them, but not too close to the robbers. In addition to this, since the oblique angle represents detachment (see Kress &amp; van Leeuwen, 2006), the audience are not involved in the shooting; they merely observe the situation. In the first frame, Mungil is depicted at medium-low angle. In contrast, in the second frame Mudy is depicted at eye-level angle. The difference in vertical angle indicates power difference between both represented participants. Based on Kress and van Leeuwen (2006), an object shot at low angle in the image represents the power the object has over the viewer. In contrast, an object shot at eye-level angle in the image represents the power equality between the object and the viewer (Kress &amp; van Leeuwen, 2006). Since Mungil is depicted at low angle, he has more power than the viewer. However, the power is imposed on Mudy, rather than on audience. On the other hand, Mudy is depicted at eye-level angle, implying that he has less power than Mungil. This might be due to the fact that Mudy runs out of ammunition and, without it, he cannot counterattack the police officers. In contrast, Mungil apparently has more ammunition that he is able to give some to Mudy. Furthermore, both participants do not shot direct gaze at the audience. According to Kress and van Leeuwen (2006), the absence of direct gaze to viewer represents offer. Thus, in this scene, the audience are only offered to observe the situation, not further involved in it.</p> <p>This scene employs given – new layout composition, realized through two frames constituting the scene. According to Kress and van Leeuwen (2006), elements positioned in the left is a <i>given</i> information and elements positioned in the right is a <i>new</i> information. In the first frame, Mungil is positioned at the right part of the frame, while in the second frame</p>

Mudy is positioned at the left part of the frame. Since both frames depict the same setting of place and the second frame is the continuation of the first frame, it can be said that Mudy is on Mungil's left side. Thus, Mudy becomes the given information and Mungil becomes the new information. By being new information, Mungil is expected to gain more attention from the audience. In relation to the power difference, the layout composition used helps focusing the audience to Mungil as the more powerful participant who has more ammunition than Mudy.

Similar to the previous scene (scene 5), the dialogue line "Tangkap, mang!" (loosely translated as "Catch it, bro!") corresponds to the action done by the characters. This means that the relationship between the dialogue and the moving images in this scene is elaboration, since both verbal and non-verbal modes convey the same meanings, yet in different way (see Barthes, 1977, in Kress & van Leeuwen, 2006). In this scene, subtitles do not appear in the screen. It implies that the dialogue is not translated and, therefore, non-Indonesian audience do not have written verbal guide to understand the meanings in the dialogue. This results on the distribution of meanings to non-verbal modes. In line with this, the absence of the dialogue also indicates the use of text reduction in the subtitling. Since the audience cannot rely on verbal mode, they rely on what is depicted in the screen to understand the situation. However, the dialogue uses a culture-specific form of address (Sundanese word "Mang") which might be confusing for non-Indonesian audience without any verbal guide to explain the meanings of the culture-related item.



Scene 8: Identifying logo				
Time	Moving images	Indonesian dialogue	English subtitles	Sound and sound effect (sfx)
54:52 – 54:54		<b>[written text on logo]</b> <i>RUMAH SAKIT JIWA CINTA WARAS</i>	[54:52-54:54] Loving Sane Mental Hospital	Voice: [no dialogue]  SFX: computer sounds
54:55 – 54:57		<b>Boy William:</b> <i>Itu logo rumah sakit jiwa.</i>	[54:55-54:56] That's the logo of a mental hospital	Voice: Boy's voice is at middle pitch level.  SFX: [no sound effects]
<p><b>Vector shift:</b> Arm → Logo (frame 1); Police officers → Logo in computer monitor (frame 2)</p> <p><b>Shot size and angle:</b> The first frame is shot in close shot, while the second frame is shot in medium shot. The angle in the first frame is frontal, with the computer monitor positioned at eye-level. However, in the second frame the situation is shot in oblique angle, with the participants in the scene positioned a little lower than</p>			<p><b>Back-translation:</b></p> <p><b>Caption:</b> <i>Rumah sakit jiwa yang penuh kasih dan masuk akal</i></p> <p><b>Boy William:</b> <i>Itu logo sebuah rumah sakit jiwa.</i></p>	

<p>audience.</p> <p><b>Gaze:</b> No direct gaze at audience</p> <p><b>Composition:</b></p> <ul style="list-style-type: none"> <li>- Frame 1: Given (Arm) – New (Logo)</li> <li>- Frame 2: Given (Logo) – New (Police officers)</li> </ul>	
<p><b>Speech-Text-Image Relation</b></p>	<p>The scene illustrates the police officers trying to reveal the identity of the robbers. Based on the recording provided by the bank CCTV, the officers find out that all the robbers wear the same exact bracelet. As they zoom the recording in, they see a logo of a mental hospital in the bracelet, assuming that the robbers are actually the patients of the mental hospital.</p> <p>Based on the visual analysis, the vector in the first frame is realized through the arm of the robber, aiming at the logo which appears on the right part of a computer monitor. In this case, the vector aims at the logo but the source of the vector remains unknown (the arm only shapes the vector). According to Kress and van Leeuwen (2006), the representation is called ‘event’; there is a goal—the object the vector aims at—but the actor is unknown. As a goal, the logo is shown and as a vector, the arm directs audience to focus on the logo. However, who or what shows or projects the logo to the screen remains unknown. In the second frame, the vector shifts from the police officers to the logo appearing on the monitor. The vector is realized through the officers’ gaze. Since the vector is realized through gaze, the narrative process in this scene is relational, positioning participant who emanates the gaze as the reactor and the participant who is looked at as the phenomenon (see Kress &amp; van Leeuwen, 2006). In this case, the police officers become the reactor, and the logo on the monitor becomes the phenomenon, the object seen by the police officers. The situation can be verbalized into the following sentence: “The police officers look at the logo on the monitor.”</p> <p>The scene is illustrated in different frame sizes and angles. The first frame uses close shot and frontal, eye-level angle. Close shot enables audience to observe the represented participants closely. In this case, the close shot enables audience to see the logo shown on the monitor clearly and read the text on the logo easily. In addition to this, frontal angle implies involvement—what is depicted is the part of represented participants’ and audience’s world (Kress &amp; van Leeuwen, 2006). Moreover, when an object is shot at eye-level angle, it represents the power equality between the object and the viewer (Kress &amp; van Leeuwen, 2006). The integration of frame size and shot angle in the first frame involves audience in seeing the logo on the monitor closely. In addition to this, the monitor is placed in an equal status to the audience, implying that audience can interact with or see what is shown on the monitor as if they were watching their own</p>



television. In other words, the monitor is not positioned as a more technologic item which makes the audience feel inferior. In the second frame, the audience are positioned a little further due to the shift of frame size. The situation in the second frame is presented in medium shot and at oblique, medium-high angle. By using medium shot, the audience can observe the situation inside the police van, not only a particular participant or object. Moreover, medium shot reflects the distance between represented participants and audience, which is social (see Kress & van Leeuwen, 2006). In addition, oblique angle represents detachment (Kress & van Leeuwen, 2006) and high angle used to portray an object represents the viewer's power over the object (Kress & van Leeuwen, 2006). Furthermore, in both frames there are no gaze directed at the audience. As Kress and van Leeuwen (2006) stated, the absence of gaze represents an offer. The integration of these aspects (frame size, shot angle, and gaze) in the second frame offers the audience to interact with the represented participants in the scene (the police officers). However, the interaction is only in the form of observation. On the other hand, the audience are not involved in the situation and have more power than the represented participants in the scene. This can be assumed that the audience are not involved in the process of finding information regarding the identity of the robbers; they can leave the job to the officers.

The scene uses given – new layout composition, positioning the elements on the left and right part of the frame. For instance, in the first frame the logo is positioned at the right part of the screen and the arm of the robber is positioned at the left part of the screen. According to Kress and van Leeuwen (2006), elements positioned at the left is a *given* information and elements positioned at the right is a *new* information. Since the logo is placed at the right part of the frame, the logo becomes a new information. In contrast, the arm of the robber becomes a given information. Moreover, the logo is shown in full-color, while the arm is shown in faded, black and white color. The difference in color scheme makes the logo more salient because the logo appears colorful. Being a new information, the logo becomes the element which is expected to gain more attention from the audience. In the second frame, the logo is no longer a new information since it is placed at the left part of the frame (the monitor is on the left). Audience have recognized the logo and, therefore, it is not necessary to introduce the logo as a new information; what becomes a new information is the police officers. Here, the audience is expected to focus on the police officers who finally reveal the identity of the robbers.

In this scene, the subtitles serve as a translation for the dialogue and a verbal element which appears in the image. In the first frame, for instance, the caption on the hospital logo is written in Indonesian. The caption is translated and the translation is shown in form of subtitles. However, there seems to be a failure in the translation in which proper name is translated. The name of the hospital whose logo is shown is “Cinta Waras.” Apparently, the name of the hospital is translated into English, appearing as “Loving Sane Mental Hospital” in the subtitles. The translation is also incorrect because, when back-translated to Indonesian, the phrase becomes “Rumah Sakit Jiwa yang Penuh Kasih dan Masuk

Akal.”

In the second frame, the subtitles serve as the translation of the dialogue. The scene illustrates police officers looking at the logo appearing on the monitor. One officer (Boy William) says that the logo appearing on the monitor is the logo of a mental hospital. In this frame, the subtitles do not function to add new meaning since the same logo has been shown in the previous frame, and it can be assumed that audience already recognize the logo. According to Barthes (1977, in Kress & van Leeuwen, 2006), verbal element may function to elaborate the image, and vice versa, restating the same meaning in different way. Thus, the relationship between verbal and non-verbal mode in this scene is elaboration.

Scene 9: At the porch of the mental hospital				
Time	Moving images	Indonesian dialogue	English subtitles	Sound and sound effect (sfx)
59:14 – 59:16		<b>[left banner]</b> <i>PINTU KAMI SELALU TERBUKA UNTUK ANDA MEMPEROLEH KESEHATAN JIWA</i>  <b>[right banner]</b> <i>TUHAN TIDAK PERNAH MENCIPTAKAN PRODUK GAGAL</i>	[no subtitles]	Voice: Hospital patients laughing and yelling.  SFX: James Bond-ish orchestral music in background
59:17 – 59:19		<b>[right banner]</b> <i>TUHAN TIDAK PERNAH MENCIPTAKAN PRODUK GAGAL</i>	[59:17-59:18] God never creates a failure	Voice: Hospital patients laughing and yelling.  SFX: James Bond-ish orchestra music background
<p><b>Vector shift:</b> Vector is realized through upper line of both banners → Ge (frame 1); Vector is realized through camera panning → Bintang (whole scene)</p> <p><b>Shot size and angle:</b> Medium shot → Medium-close shot. The first frame is shot in medium shot and at oblique, eye-level angle. The second frame is also shot in medium-close shot. However, the situation is presented at frontal, eye-level angle.</p> <p><b>Gaze:</b> No direct gaze at audience</p>		<p><b>Back-translation:</b> <b>[right banner]</b> <i>Tuhan tidak pernah membuat kegagalan</i></p>		

<p><b>Composition:</b></p> <ul style="list-style-type: none"> <li>- Frame 1: Center (Ge) – Margin (doors)</li> <li>- Frame 2: Given (banner) – New (Bintang)</li> <li>- Whole scene: Given (Ge) – New (Bintang)</li> </ul>	
<p><b>Speech-Text-Image Relation</b></p>	<p>The scene depicts the identity of the robbers as patients of a mental hospital before they rob the bank. Taking place at the hospital porch, a paramedic (Ge Pamungkas) is seen entering the hospital, passing Bintang who stands next to a banner, hugging the door. Bintang is one of the hospital patients who later robs the bank with seven other patients.</p> <p>Based on the visual analysis, there are two vectors in this scene. The first vector is realized through the upper line of two banners seen in the first frame. The upper lines create a vector aiming at Ge who is walking through the door. Ge becomes the object the vector aims at; however, what or who emanates the vector remains unknown. When the representation involves only a goal, it is called ‘event’ (see Kress &amp; van Leeuwen, 2006). In the first frame, audience do not have idea from which the vector emanates. However, the vector helps directing the focus of the audience to the goal—Ge. The second vector is realized through camera panning to right. Initially the camera shoot Ge who is walking through the door. Bintang is in the frame, positioned at the right part of the frame. As Ge enters the building, the camera pans to the right, showing Bintang who stands on the right side of a banner. Similarly, Bintang becomes the object the vector aims at; however, the element emanating the vector remains unknown. Thus, Bintang becomes the goal and the vector directs audience to look at Bintang.</p> <p>The frame size used in this scene shifts from medium shot to medium-close shot. In the first frame, the situation is shot in medium shot and at oblique, eye-level angle. Medium shot, according to Kress and van Leeuwen (2006) reflects the distance between represented participants and audience, which is social. In addition to this, the use of oblique angle implies that the viewer is detached from the situation depicted in the image (ibid.), while eye-level angle represents equality between the represented participants and the viewer (Kress &amp; van Leeuwen, 2006). Ge also does not direct his gaze to the audience, implying that he offers something to the audience—to recognize him and see what he is doing. The integration of these aspects (frame size, shot angle, and gaze) positions Ge at the equal level to the audience and presents him as a character whom audience can see in their life—a paramedic working at a hospital, yet the character does not have closer relationship with the audience. Furthermore, audience is not involved in the situation. This means that audience are offered to see Ge and know what he is doing, but do not have access to interact further with him. In the</p>

second frame, the situation is shot in medium-close shot and at frontal, eye-level angle. Positioned closer to the represented participant, audience can see Bintang in a closer distance. In addition to this, according to Kress and van Leeuwen (2006), frontal angle implies involvement—what is depicted is the part of represented participants' and audience's world. Moreover, Bintang is shot at eye-level angle, positioning him at equal level to the audience. Again, there is not gaze directed at the audience in this frame. These aspects indicate that audience are offered to get involved in the interaction with Bintang by observing him in a closer distance.

The scene uses given – new composition layout and center – margin composition layout. The first frame, for instance, uses center – margin composition layout, positioning Ge between two margins represented by two banners and two doors. When an element is placed at the center of the image, that element becomes the center or the core information (Kress & van Leeuwen, 2006). In contrast, the margins are supporting elements which depend on the center element (Kress & van Leeuwen, 2006). By employing this layout composition, Ge becomes the core information—the center, while two banners bordering his left and right side become the margins. Being the center, Ge is expected to gain more attention and focus from the audience since he is the core information. The banners also carry information, represented by the text written on them; however, since they are positioned as the margins, the information contained in the banners become less important than the core information. Unlike the first frame, the second frame employs given – new layout composition. When an element is positioned at the left, the element becomes a *given* information and, in contrast, when an element is positioned at the right, the element becomes a *new* information (see Kress & van Leeuwen, 2006). The banner is positioned at the left, thus becoming the given information. Meanwhile, Bintang is positioned at the right, becoming the new information. The banner contains a text “Tuhan tidak pernah menciptakan produk gagal,” translated loosely as “God never creates failed products.” Positioned as the given information, the banner represents an information which is assumed to have been recognized or known by the audience that God never creates failed products. Meanwhile, Bintang (who represents a person with mental illness) is positioned as the new information. The fact that Bintang has mental illness becomes something new that audience should focus on. However, this is not due to Bintang being a new failed creation; rather, it is because the given information Bintang is one of God's creations and, regardless his mental illness, the given information gives a sense of both positivism and humor.



Nevertheless, it is important to see that the two frames depicted is a unity. In other words, both frames are shot in one shot, merged by camera panning and differentiated only by the frame size and the angle. Thus, it can be assumed that both frames, when positioned side-by-side constitute a unity. As a whole scene, it also indicates the use of layout composition, just like the individual frames. In the whole scene, the layout composition used is given – new composition, positioning Ge as the given information and Bintang as the new information. As a paramedic, Ge represents a character or



a person whom audience are already familiar with. In real life, paramedics are common to see, especially in hospitals or clinics. In contrast, Bintang represents a person with mental illness. Positioned as the new information, Bintang represents a character whom audience are not really familiar. In other words, this character does not appear regularly in audience's life. Moreover, it is important to remember that some people regard mental illness as a flaw which distinguishes those with the illness from people without mental illness.

Regarding the relationship between the verbal and the non-verbal mode, the subtitles and the moving images convey different meanings. The moving images, for instance, illustrate the situation in a mental hospital, showing one of the mental hospital patient. On the other hand, the subtitles, corresponding to the verbal element in the moving image, says "God never creates a failure." Both express different meanings, yet, when integrated, express meaning which is more complex. The subtitles added to the scene function to add new meanings to complete the message in the scene. With the addition of the subtitles, what may be possibly conveyed in the scene is that even though some people (represented by the patients of the mental hospital) have flaws, it does not mean that they are a failure, for God, who predetermined their existence, never creates a failure. From this perspective, it can be seen that the relationship between the subtitles and the moving images is extension, for the subtitles add new meanings to the image which complete the message (see Barthes, 1977, in Kress & van Leeuwen, 2006).



Scene 10: The doctrine				
Time	Moving images	Indonesian dialogue	English subtitles	Sound and sound effect (sfx)
1:03:51 – 1:03:53		<b>Pandji (voice on radio):</b> <i>Kalian adalah penjahat paling aneh.</i>	[no subtitles]	Voice: Pandji's voice is at high pitch level, evoking a doctrine. The voice comes from the radio.  SFX: Dark and gloomy guitar music
1:03:53 – 1:03:55		<b>Pandji (voice on radio):</b> <i>Kalian adalah penjahat paling aneh.</i>	[1:03:53-1:03:54] You are the strangest criminals	Voice: Pandji's voice is at high pitch level, evoking a doctrine. The voice comes from the radio.  SFX: Dark and gloomy guitar music
<p><b>Vector shift:</b> (Actor represented by an arm) → Radio (frame 1); Pandji → Mudy (frame 2)</p> <p><b>Shot size and angle:</b> Close shot → Medium-long shot. The first frame is shot in close angle and at oblique, eye-level angle. Meanwhile, the second frame is shot in medium-long shot and at oblique, high angle.</p> <p><b>Gaze:</b> No direct gaze at audience</p>			<p><b>Back-translation:</b></p> <p><b>Pandji:</b> <i>Kalian adalah penjahat paling aneh</i></p>	



<p><b>Composition:</b></p> <ul style="list-style-type: none"> <li>- Frame 1: Given (Pandji’s hand) – New (Radio)</li> <li>- Frame 2: Center (Pandji) - Margin (Mungil &amp; Mudy)</li> </ul>	
<p><b>Speech-Text-Image Relation</b></p>	<p>The scene illustrates Pandji, the doctor at Cinta Waras mental hospital, brainwashing the patients. He indoctrinates them, telling them that they are criminals and will have to rob a bank. In this scene, Pandji is trying to brainwash Mudy and Mungil while they are sleeping, indoctrinating them with the idea of being the strangest criminals.</p> <p>Based on the visual analysis, the arm shapes the vector, aiming at the radio. The actor is only represented by the arm. In this case, it can be assumed that what is visible from the actor is only his hand, so the actor remains anonymous. Thus, it can be said that there are only a vector and a goal in the image. The representation of action involving only a goal is called ‘event’ (Kress &amp; van Leeuwen, 2006). When a situation is depicted as an event, what is known is that something happens to the goal, but it remains unknown who does it. Thus, in the first frame, the audience are able to see what is done—that a radio is being turned on, yet they do not know who turns on the radio. In the second frame, the vector shifts from Pandji to Mudy. The vector is realized through Pandji’s gaze. Since the vector is realized through gaze, the narrative process in the second frame is relational, positioning participant who emanates the gaze as the reactor and the participant who is looked at as the phenomenon (see Kress &amp; van Leeuwen, 2006). At the time Pandji looks at Mudy, the narration from the radio says “Kalian adalah penjahat paling aneh,” loosely translated as “You are the strangest criminals.” This can be assumed that Pandji demands Mudy to think that he is the strangest criminal. The process in the second frame thus can be verbalized into the following sentence: “Pandji looks at Mudy, seeing him as the strangest criminal.”</p> <p>This scene is illustrated in two different frame sizes. The first frame uses close-shot, while the second frame uses medium-long shot. The situation in both frames is shot at oblique angle; however, in terms of vertical angle, the situation is shot in different angle. The first frame uses eye-level angle, while the second frame uses medium-high angle. Illustrated in close-shot, in the first frame audience are allowed to closely observe what the represented participants in the scene do (in this case, Pandji is turning on the radio, while Mungil and Mudy are sleeping). The use of close shot also creates a dramatic atmosphere in the scene. The situation is shot at oblique angle which functions to detach audience from the situation depicted (see Kress &amp; van Leeuwen, 2006). In addition to this, the frame uses eye-level angle, representing the equality in power between the represented participants and the audience (Kress &amp; van Leeuwen, 2006). Moreover, the represented participants do not direct any gaze to the audience. The absence of direct gaze to viewer</p>

represents an offer (Kress & van Leeuwen, 2006), which in this case can be understood as an offer to observe the situation depicted in the scene. As the result of the integration of these aspects (frame size, shot angle, and gaze), audience are offered to see or observe what the represented participants are doing in the scene. They are allowed to see the details in the situation (for instance, the appearance of the radio or what other represented participants—Mungil and Mudy—are doing) and experience the situation in more dramatic setting. However, the audience can only observe and are not further involved in the action done by the represented participants; audience are there to know what happens, not to further interact with them. As for the second frame, the audience are positioned further than they were in the first frame. By being further, they are no longer able to observe the situation closely; however, they can see more elements in the image (for instance, Pandji's face, the corner of the room, and so on). When elements in the image are shot in medium shot or middle distance, the elements are represented as “within the viewer's reach, but not actually used” (see Kress & van Leeuwen, 2006). In addition to this, oblique angle depicts elements in the image as something the viewer is not involved with (ibid, p. 136). Moreover, the situation is shot at medium-high angle, which represents the power (the viewer has) over the represented participants in the image (Kress & van Leeuwen, 2006). The represented participants in the second frame do not direct their gaze to the audience. This implies that the represented participants only offer something to the audience (Kress & van Leeuwen, 2006). Overall, the integration of these aspects in the second frame (frame size, shot angle, and gaze) offers the audience to observe the situation in a further distance, yet still within the audience's reach. They are also not involved in the situation and positioned higher than the represented participants. This implies the audience can only see what Pandji is doing to the patients; they are not involved in the action (the brainwashing), either as the agent or as the patients. The audience is positioned as having more power than the represented participants because they are not brainwashed and, unlike the hospital patients, cannot be brainwashed by Pandji.

As for layout composition, the scene is illustrated through two layout composition. For instance, the first frame is composed in given – new layout composition. According to Kress and van Leeuwen (2006), when an element is positioned at the left, the element becomes a *given* information and, in contrast, when an element is positioned at the right, the element becomes a *new* information. In the first frame, Pandji (represented by his arm) is the given information while the radio is the new information. It can be assumed that Pandji, as a doctor, represents an information that has been known by audience. In other words, the audience are already familiar with doctors; they already know what the doctors do in real life. In contrast, the radio, despite already being common object in audience's lives, represents something that the audience may not know. The audience know Pandji, but what is played on the radio remains new information which should be focused on. In other words, a doctor turning on the radio and playing something for the patients is an

information that should gain more attention. In the second frame, center – margin layout composition is used. When an element is placed at the center of the image, that element becomes the center or the core information and, in contrast, elements bordering the center become the margins (see Kress & van Leeuwen, 2006). This composition positions Pandji as the center and the patients—Mudy and Mungil—as the margins. Since Pandji is the core information, the audience are directed to focus on him. By focusing on Pandji, audience are able to observe more details in him (for instance, the direction of his gaze, or what is he doing). Meanwhile, despite being the margins, Mudy and Mungil also carry information and meanings; however, the meanings they carry depend on Pandji. Without Pandji, Mudy and Mungil represent only mental hospital patients who are sleeping. However, since Pandji is there and the audience know that Pandji is trying to perform a brainwash, Mudy and Mungil represent mental hospital patients who are about to be brainwashed.

Regarding the relationship between the verbal and the non-verbal modes, it is found that the narration adds new meanings to the moving images. The meanings conveyed through the moving images are different to the ones conveyed through the narration, even though to some extent both narration and moving images corresponds to each other. The correspondence can be observed from the narration line and the gaze of the character. As the narration says “Kalian adalah penjahat paling aneh” (loosely translated as “You are the weirdest criminals”), the character directs a gaze at the other character. Moreover, the subtitles serve as a translation for the dialogue (as seen on the second frame). What is conveyed in the subtitles corresponds to the dialogue. In the second scene, Pandji’s gaze to Mudy corresponds to his narration on radio (“You are the strangest criminals”). In addition to this, the sound produced from the radio and dark background music used in the scene create mysterious ambiance which intensifies the situation that Pandji is brainwashing his patients, indoctrinating them with the idea of being the strangest criminals. When the verbal element further defines the meanings in non-verbal elements, and vice versa, the relationship between the verbal and the non-verbal mode is extension (see Barthes, 1977, in Kress & van Leeuwen, 2006). Since the verbal and the non-verbal modes convey the different, yet complementary meanings, it can be said that the narration (and hence, the subtitles) adds news meanings to the moving images and sounds, making the meanings in the scene complete.

Scene 11: Shooting Candil				
Time	Moving images	Indonesian dialogue	English subtitles	Sound and sound effect (sfx)
1:11:58 – 1:12:00		<b>Ari:</b> <i>Sudah siap? Satu, dua, tiga!</i> <b>Candil:</b> <i>Waaa! Aduh! Pak!</i>	[1:11:58-1:11:59] Are you ready? // One [1:11:59-1:12:00] Two // Three.	Voice: Ari counts down to shoot Candil. Ari's voice is at high pitch level.  SFX: gunshot sound, James Bond-ish orchestral music in background
1:12:01 – 1:12:03		[Candil screams] [Desk plate titled "Manager"]	[no subtitles]	Voice: Candil screams  SFX: gunshot sound, glass breaking
<b>Vector shift:</b> Ari → Candil  <b>Shot size and angle:</b> Both frames are shot in medium shot. However, the first frame is shot at oblique, medium-low angle. Meanwhile, the second frame is shot at oblique, high angle.  <b>Gaze:</b> No direct gaze at audience			<b>Back-translation:</b>  <b>Ari:</b> <i>Apa kamu siap? Satu, dua, tiga.</i>	

<p><b>Composition:</b></p> <ul style="list-style-type: none"> <li>- Frame 1: Center (Ari Kriting) – Margin (blank space)</li> <li>- Frame 2: Center (Candil) – Margin (Door jambs)</li> </ul>	
<p><b>Speech-Text-Image Relation</b></p>	<p>The scene illustrates Ari Kriting shooting a hostage (portrayed by Candil) as a way to threaten the police officers. This scene contains humour in which Ari Kriting counts down to zero (as if he is counting down on New Year’s eve) before shooting Candil. Moreover, Candil hysterically pleads for being released. Ari eventually shoots Candil regardless Candil’s hysterical and childish reaction.</p> <p>Based on the visual analysis, the vector shifts from Ari to Candil. This scene is constituted of two sequential frames. The first frame portrays Ari and the second frame portrays Candil. The vector is realized through the rifle that Ari uses to shoot Candil, going from left to right where Candil is positioned at. Since Ari imposes an action on Candil, the narrative process in this frame is action process, positioning Ari as an actor, the participant “from which the vector emanates, or itself, in whole or in part, forms the vector” (Kress &amp; van Leeuwen, 2006, p. 63) and Candil as the goal, the object the vector aims at (see Kress &amp; van Leeuwen, 2006). This action process can be verbalized into the following sentence: “Ari shoots Candil (using a rifle)”</p> <p>The scene is illustrated in medium shot, using two different shot angles for each frames. The first frame uses oblique, medium-low angle, while the second frame uses oblique, medium-high angle. According to Kress and van Leeuwen (2006) elements shot in medium shot or middle distance are represented as “within the viewer’s reach, but not actually used.” In terms of horizontal angle, both frames are shot at oblique angle, which represents detachment (Kress &amp; van Leeuwen, 2006). However, in terms of vertical angle, the first frame uses medium-low angle. When an element or represented participant in the image is shot at low angle, the relationship between the represented participant and the viewer is portrayed as one in which the represented participant has power over the viewer (see Kress &amp; van Leeuwen, 2006). In contrast, the second frame uses medium-high angle, which implies that the viewer has power over the represented participant (see Kress &amp; van Leeuwen, 2006). Moreover, represented participant in each frame does not direct a gaze to the audience. When represented participant in the image does not show a direct gaze to the viewer, the image offers something—the participant as an information to the viewer (see Kress &amp; van Leeuwen, 2006). Thus, in the first frame, the image offers audience to see what the represented participant is doing. The audience are positioned at the distance which is close enough to the participant, but are not able to interact further with the participant; they can just observe the situation. Moreover, the represented participant is portrayed as having power over the audience. However,</p>

since the audience do not further interact with the participant (and the represented participant does not show direct eye-contact with the audience), it can be assumed that the power is not imposed on the audience, yet the audience are able to know that the participant has more power than them. Meanwhile, the second frame offers audience to see what happens to the represented participant. The audience are positioned at the distance which is close enough to see the situation, yet still they are not involved in the situation. Being not involved in the situation, the audience can only observe the situation and do not directly experience the shooting as the hostage. In addition, the represented participant is portrayed as having less power than the audience. This might be due to the fact that the represented participant is shot and the audience are not involved in the shooting as the hostage because they just observe the shooting.



The scene employs center – margin layout composition which composes two frames constituting the scene as a whole. Using center – margin composition, the element placed at the center of the image becomes the center or the core information and, in contrast, the element bordering the center become the margins (see Kress & van Leeuwen, 2006). In the first frame, Ari becomes the core information since he is positioned at the center, bordered by blank space in his left and right side (to be more specific, two window jambs appearing on the image can be considered the margins). Moreover, Ari appears salient in the frame. Being the salient, core information, Ari is expected to gain attention from the audience. This way, the audience can focus on what Ari is doing. In the second frame, Candil becomes the center, bordered by two door jambs serving as the margins. Candil, just like Ari, is expected to gain attention from the audience since he is the core information. As the result, in the second frame audience can focus more on Candil and what happens to him.

Regarding the relationship between the verbal and the non-verbal modes, the scene illustrates the correspondence between the dialogue and the moving images, in which the subtitles serve as the translation for the dialogue, as seen on the first frame. The non-verbal modes corresponds to both dialogue and subtitles on the screen. Specifically, the dialogue (which contains a countdown to one) and the subtitles (which serves as the translation of the countdown) is in accord with the character's gesture, voice intonation, and facial expression. According to Barthes (1977, in Kress & van Leeuwen, 2006), when verbal element restates the same meanings as in non-verbal elements, and vice versa, the relationship between the verbal and the non-verbal mode is elaboration. This means that the subtitles and what is illustrated in the scene convey the same meaning, only in different ways. However, there is one dialogue line (uttered by Candil) which is not translated. At the time the line is uttered, the camera does not turn at Candil, thus making audience have no idea what kind of expression or gesture Candil show. Even though the line is uttered in a certain intonation (Candil pleads for freedom in such a childish, hysterical way), non-Indonesian audience might have no idea what Candil says because the line “Waa! Aduh! Pak!” is left untranslated. The absence of the translation of Candil's line is



	paradoxical. The absence allows audience to focus more on Ari's dialogue line and action. On the other hand, the audience miss a message carried in Candil's line because, regardless Candil's intonation which may signify childish fear, the audience still need the translation as a guide to understand the meanings in the whole scene.
--	--



Scene 12: The robbers escaping				
Time	Moving images	Indonesian dialogue	English subtitles	Sound and sound effect (sfx)
1:21:27 – 1:21:29		[no dialogue]	[no subtitles]	Voice: [no dialogue]  SFX: car engine, water swoosh, James Bond-ish orchestral music in background
1:21:30 – 1:21:32		[no dialogue]	[no subtitles]	Voice: [no dialogue]  SFX: gunshot sound, James Bond-ish orchestral music in background
<p><b>Vector shift:</b> Car's direction shapes the vector (no goal) (frame 1); Police officers → the car (frame 2)</p> <p><b>Shot size and angle:</b> Medium-close shot → Medium-long shot. The first frame is shot in medium close shot and at frontal, eye-level angle. However, the second frame is shot in medium long shot and at frontal, eye-level angle.</p> <p><b>Gaze:</b> Represented participants in both frames show direct gaze at the audience</p>			<p><b>Back-translation:</b> [no subtitles]</p>	

<p><b>Composition:</b></p> <ul style="list-style-type: none"> <li>- Frame 1: Center (Car) – Margin (water splash)</li> <li>- Frame 2: Center (Nirina) - Margin (police officers)</li> </ul>	
<p><b>Speech-Text-Image Relation</b></p>	<p>The scene illustrates the robbers eventually managing to escape from police pursuit. They run away with the money they have robbed. Meanwhile, three police officers try to pursue the robbers. Unfortunately, the car has already gone before the police officers can manage to stop it. They can only shoot the car from afar.</p> <p>Based on the visual analysis, there are two vectors in this scene. The first vector can be seen in the first frame. In the first frame, the vector is realized through the movement of the car. The car is moving to a direction, yet where the car exactly goes to remains unknown. This representation of action involving only an actor is called non-transactional process (see Kress &amp; van Leeuwen, 2006). This implies that as the actor, the car performs an action (moving). However, the car does not impose the action on another objects; it just moves. Thus, the narrative process can be verbalized into the following sentence: “The car is moving (to a certain direction).” However, in the second frame, the vector is realized through the direction of the guns used by the police officers (to be more specific, the bullets shot out of the gun). When analyzed as an individual frame, the second frame has only an actor and a vector. However, since it is known that in the previous frame the robbers run away (on the car), it can be assumed that the police officers are trying to shoot the car. In other words, the vector in the second frame is aimed at the car. Thus, in the second frame the actor is the police officers and the goal is the car. This narrative process can be verbalized into the following sentence: “The police officers is shooting the car.”</p> <p>This situation in this scene is illustrated in two different frame sizes. The first frame, for instance, uses medium-close shot. When an object is shot in medium shot, according to Kress and van Leeuwen (2006), the object is represented as “within the viewer’s reach, but not actually used.” The situation is also shot at frontal, eye-level angle. According to Kress and van Leeuwen (2006), frontal angle implies involvement—what is depicted is the part of represented participants’ and audience’s world. Meanwhile, an element or represented participant in the image which is shot at eye-level angle portrays the equality of power between the represented participant and the viewer (see Kress &amp; van Leeuwen, 2006). The represented participant in this frame (the car) also directs a gaze at the audience. Kress and van Leeuwen (2006) assumes that represented participants who direct a gaze at viewers are usually human or animals, but not always, noting that car headlights can be assumed as eyes looking at the viewer. Referring to this assumption, it can be said that the car directs a gaze at the audience since its headlights frontally directs the light toward the camera. When present, the</p>

gaze (and the gesture) of the participant represents a demand (ibid.). From the integration of these visual aspects (frame size, shot angle, and gaze) in the first frame, it can be assumed that the audience is demanded to be involved in the situation. The audience are also positioned in a close distance and at the equal level to the represented participant. Relating to the plot of the story, what is demanded to the audience is the acknowledgment that the robbers are able to manage to escape from the police pursuit. In addition to this, the close distance between the audience and the represented participant enables the audience to see the power the represented participant has (since the robbers are able to escape and run away with the money). In the second frame, the audience are positioned a little further to the represented participants (the police officers). However, they are still demanded to get involved in the situation. Relating to the plot of the story, what the represented participants demand is for the audience to focus on the participants and see that the police officers are ‘not kidding’ in pursuing the robbers.

In illustrating the situation, the scene uses center – margin layout composition for both frames. When using center – margin composition, the element placed at the center of the image is considered the center or the core information and, in contrast, the element bordering the center is considered the margin (see Kress & van Leeuwen, 2006). In the first frame, the car becomes the center, while the water splash bordering the car becomes the margin. As the center, the car carries the core information that audience should focus on. In this case, the car represents the robbers and their ‘victory’ for being able to escape and carry the money the robbed. The water splash, on the other hand, also carries meanings, yet the meanings depend on the center. Without the center—the car, the water splash would seem nonsense because there is nothing moving through the puddle that creates the splash. Moreover, in the first frame the water splash rather functions to border the focus area so the audience would focus only on the car. In the second frame, Nirina becomes the center, while the other two police officers become the margin. Being the center, Nirina is expected to gain more attention than the other police officers, becoming the focus of the audience. Moreover, the positioning of Nirina as the center might be due to her important role in the film and, inevitably, her status as a celebrity. Compared to the other two police officers, the audience are more likely to recognize Nirina. Thus, by directing the audience to focus on the core information, Nirina becomes more salient.

Regarding the relationship between the verbal and the non-verbal modes, there is no dialogue line in this scene. Moreover, the scene also does not show any subtitles functioning as the translation or transcription of any information. Since the scene is devoid of dialogue and subtitles, the meanings are distributed to visual and audio mode. The visual mode functions to illustrate what happens in the scene or what are represented participants doing in the scene, meanwhile the audio mode functions to strengthen the message (for instance, gunshot sound as the represented participants shoot) and intensify the ambiance (for instance, James Bond-ish orchestral music). This implies that there is no relation between

the verbal and the non-verbal elements in this scene since both dialogue and subtitles are absent. As the result, to understand the meanings in the scene, audience can only grasp or guess the meanings through the moving images.

Scene 13: Arresting the robbers				
Time	Moving images	Indonesian dialogue	English subtitles	Sound and sound effect (sfx)
1:30:49 – 1:30:51		<b>Cak Lontong:</b> Dari mana kalian tahu?	[1:30:50-1:30:51] <i>How did you guys know...?</i>	Voice: Cak Lontong's voice is at high pitch level  SFX: James Bond-ish orchestral music in background
1:30:52 – 1:30:54		<b>Cak Lontong:</b> Eh, kamu berkhianat, ya? Hey, kamu berdua berkhianat, ya?!  <b>Nikita:</b> (frame 1) <b>Nikita:</b> (frame 2)	[1:30:52-1:30:54] <i>Hey, did you betray me? Did you?</i>	Voice: Cak Lontong's voice is at high pitch level  SFX: James Bond-ish orchestral music in background
<b>Vector shift:</b> Agung Hercules & Cak Lontong → Nikita (frame 1); Cak Lontong → Nikita (frame 2)			<b>Back-translation:</b> <b>Cak Lontong:</b> <i>Bagaimana kalian tahu?</i>	

<p><b>Shot size and angle:</b> The first frame is shot in medium-long shot, while the second frame is shot in close-shot, depicting Cak Lontong in a close distance. In terms of horizontal angle, both frames are shot at frontal angle. However, in terms of vertical angle, the first frame is shot in medium-low angle and, in contrast, the second frame is shot at eye-level angle.</p> <p><b>Gaze:</b> No direct gaze at audience</p> <p><b>Composition:</b></p> <ul style="list-style-type: none"> <li>- Frame 1: Center (Nikita) – Margin (Agung Hercules &amp; Cak Lontong)</li> <li>- Frame 2: Given (Police officer) – New (Cak Lontong)</li> </ul>	<p><i>Hey! Kamu mengkhianati saya, ya?</i></p>
<p><b>Speech-Text-Image Relation</b></p>	<p>The scene illustrates the robbers (now portrayed by Nikita Mirzani and Agung Hercules) and the crime kingpin (portrayed by Cak Lontong) arrested by the police officers. After the robbers manage to escape from the bank, they meet Cak Lontong, a crime kingpin to give the money the robbed. However, the police officers manage to track the robbers' whereabouts and eventually arrest them. Cak Lontong thinks that Nikita and Agung betray him for telling the police about the robbery and their whereabouts.</p> <p>Based on the visual analysis, the vector shifts from Agung Hercules and Cak Lontong to Nikita Mirzani, realized through their gaze. In the first frame, both Agung Hercules and Cak Lontong direct their gaze at Nikita. However, the second frame only depicts Cak Lontong in a closer distance. Since Cak Lontong's position in the second frame remains same as his position in the first frame, and he faces the same direction as he does in the first frame, it can be assumed that Cak Lontong still directs his gaze at Nikita. Since the vector is realized through gaze, the narrative process in the second frame is relational, positioning participant who emanates the gaze as the reactor and the participant who is looked at as the phenomenon (see Kress &amp; van Leeuwen, 2006). Thus, in this scene, Cak Lontong and Agung Hercules become the reactor, while Nikita becomes the phenomenon. Since Cak Lontong thinks that he has been betrayed by the robbers, this situation can be verbalized in the following sentence: "Feeling betrayed, Cak Lontong glares at Nikita."</p> <p>The scene is illustrated in two different frame sizes. The first frame, for instance, uses medium-long shot. Meanwhile, the second frame uses close shot. According to Kress and van Leeuwen (2006) elements shot in medium shot or middle distance are represented as "within the viewer's reach, but not actually used." In addition to this, close shot (or close up) represents intimate distance between the represented participants in the image and the viewer (Kress &amp; van Leeuwen,</p>



2006). In terms of horizontal angle, both frames use frontal angle. According to Kress and van Leeuwen (2006), frontal angle implies involvement—what is depicted is the part of represented participants’ and audience’s world. Moreover, in terms of vertical angle, the first frame uses medium-low angle, while the second frame uses eye-level angle. An element or represented participant in the image which is shot at low angle portrays the power the element has over the viewer (see Kress & van Leeuwen, 2006). In contrast, an element or represented participant in the image which is shot at eye-level angle represents the equality of power between the element and the viewer (Kress & van Leeuwen, 2006). Furthermore, any represented participants in the scene do not direct a gaze to the audience. This implies that the represented participants are only offered as an information for the audience (see Kress & van Leeuwen, 2006). It can be assumed that from the integration of these visual aspects (frame size, shot angle, and gaze), the audience are offered to get involved in the situation depicted in the scene, as if they were there in the real location. However, the involvement is only in form of observation; the audience are not demanded to know or interact further with the represented participants. In the first frame audience are positioned further from the represented participants (the robbers and the police officers), yet they are allowed them to observe the situation from a wider perspective. However, the audience is positioned as having less power than the represented participants, making the robbers seem more powerful than the audience. In contrast, audience are positioned closer to the represented participant in the second frame, enabling them to observe more details in the represented participant (for instance, Cak Lontong’s facial expression) within an intimate (or in this case, close) distance. In this frame, the represented participant (Cak Lontong) is portrayed as in the equal power relation to the audience. This might be due to the fact that the robbers and the crime kingpin have been arrested and, therefore, they are no longer more powerful.

This scene employs two different layout compositions. The first frame is illustrated through center – margin layout composition, while the second frame is illustrated through in given – new layout composition. In center – margin composition, the element placed at the center of the image becomes the center or the core information and, in contrast, the element bordering the center become the margins (see Kress & van Leeuwen, 2006). In addition to this, in given – new composition, the element placed at the left becomes the given information, while the element placed at the right becomes the new information (Kress & van Leeuwen, 2006). In the first frame, Nikita is positioned at the center of the frame, thus becoming the center or the core information. Meanwhile, Agung Hercules and Cak Lontong is positioned bordering the center, thus becoming the margins. As the center or core information, Nikita gains more attention from the audience. Indeed, the three criminals commit crime. However, the reason why Nikita should be more focused on might be due to the fact that earlier in the beginning of the film, Nikita is portrayed as one of the bank tellers, yet she is actually one of the robbers. In other words, the audience are directed to focus on a character who is actually a robber, but

disguises herself as a bank teller. Moreover, in the second frame, the police officer becomes the given information and Cak Lontong becomes the new information. From this positioning, it can be assumed that the police officer, as the given information, is a character whom audience are already familiar. Audience know what police officers in real life do and police officer arresting criminals is not something new for the audience because it is one of their jobs. On the other hand, Cak Lontong represents new information—a character whom audience do not encounter regularly. In real life, crime kingpin is a character who does not appear regularly. This character is even somewhat unwanted, for crime kingpins commit crime and might do harm to other people.

Regarding the relationship between the verbal and the non-verbal modes, the subtitles serve as the translation for the dialogue in the scene. What is conveyed through both dialogue and subtitles corresponds to what is depicted visually. In the second frame, for instance, the suspicion indicated in the dialogue (“Eh, kamu berkhianat, ya? Hey, kamu berdua berkhianat, ya?”) and the subtitle (“Did you betray me? Did you?”) is visualized through Cak Lontong’s facial expression and voice intonation. This implies that the verbal and the non-verbal elements constituting the scene express the same meaning, only in different ways. According to Barthes (1977, in Kress & van Leeuwen, 2006), when verbal element restates the same meanings as in non-verbal elements, and vice versa, the relationship between the verbal and the non-verbal mode is elaboration. Thus, since the relationship between the verbal and the non-verbal mode is elaboration, what is conveyed through the subtitles is elaborated visually and acoustically.



Scene 14: The real identities				
Time	Moving images	Indonesian dialogue	English subtitles	Sound and sound effect (sfx)
1:32:11 – 1:32:13		[no dialogue]	[no subtitles]	Voice: [no dialogue]  SFX: mid-tempo hard rock music in background
1:32:14 – 1:32:16		[no dialogue]	[no subtitles]	Voice: [no dialogue]  SFX: mid-tempo hard rock music in background
<p><b>Vector shift:</b> Ernest walking, emanating a vector (no goal) (frame 1); Upper lines of two cars shaping the vector → Fico (frame 2)</p> <p><b>Shot size and angle:</b> Medium-long shot → Medium shot. The first frame is shot in medium-long shot and at oblique, eye-level angle. However, the second frame is shot in medium shot and at oblique, low angle.</p> <p><b>Gaze:</b> Ernest directs a gaze to the audience (frame 1)</p>		<p><b>Back-translation:</b> [no subtitles]</p>		





<p><b>Composition:</b></p> <ul style="list-style-type: none"> <li>- Frame 1: Given (Ernest) – New (Car)</li> <li>- Frame 2: Center (Fico) - Margin (Cars)</li> </ul>	
<p><b>Speech-Text-Image Relation</b></p>	<p>The scene portrays the real identity of the robbers. They are actually secret agents who try to investigate a bank robbery plan involving a doctor from Cinta Waras mental hospital, Pandji, and his accomplices. The agents are hypnotized by Indro, their boss, indoctrinated with the idea that they have mental illness. They are taken into the mental hospital, only to be brainwashed with the idea that they have to robber a bank. This way, Indro has access to investigate Pandji’s plan and thwart the robbery with the help of Candil (who is previously seen shot by the robbers). This scene shows the robbers’ real identity as secret agents, wearing exquisite outfits and riding their fancy cars and motorcycles.</p> <p>Based on the visual analysis, there is a vector shift in each frame constituting the scene. In the first frame, the vector is realized through Ernest’s step. Ernest is seen walking to a certain direction and he does not impose an action on any other element; Ernest just walks. This representation of action is a non-transactional process since there is no goal or any elements to aim at (see Kress &amp; van Leeuwen, 2006). The narrative process in the first frame involves only an actor—Ernest who emanates a vector by walking to a certain direction. This process can be verbalized into the following sentence: “Ernest is walking (to a certain direction).” Similarly, in the second frame, there are two vectors which are realized through the upper line of two cars shown in the left and the right part of the frame. Both vectors aim at Fico who is positioned at the center of the frame. Again, the representation of action in the second frame is an ‘event’, since the narrative process involves only a goal—Fico. However, who emanates the vector remains unknown. In general, the vectors in the second frame function to help directing audience to focus on the goal.</p> <p>The scene is illustrated in two different frame sizes, as well as shot angles. The situation in the first frame, for instance, is shot in medium-long shot and at oblique, eye-level angle. According to Kress and van Leeuwen (2006, p. 127-128) elements shot in medium shot or middle distance are represented as “within the viewer’s reach, but not actually used.” Moreover, medium shot represents a distance between the represented participant and the viewer, which is social (Kress &amp; van Leeuwen, 2006). In terms of horizontal angle, the situation is presented at oblique angle. An image presented at oblique angle represents detachment of the audience from the represented participant or the situation depicted (Kress &amp; van Leeuwen, 2006). Meanwhile, in terms of vertical angle, the frame uses eye-level angle, representing the equality in power between the represented participants and the audience (Kress &amp; van Leeuwen, 2006). Moreover, the represented participant—in this case, Ernest—directs a gaze to the audience. Direct gaze toward the viewer, according to Kress and</p>

van Leeuwen (Kress & van Leeuwen, 2006), indicates that the represented participant demands something from the viewer. From the integration of these visual aspects (frame size, shot angle, and gaze) in the first frame, it can be assumed that Ernest, as the represented participant demands something from the audience; he wants the audience to focus on him. Despite being in the equal relationship to the audience, the audience are detached from the situation, implying that the situation depicted is not the part of audience's life. In other words, Ernest demands the audience to recognize him as a secret agent who wears exquisite outfit and drives nice car. However, on the other hand, Ernest tells the audience that he is not the part of the audience's life, implying that the audience are not secret agent like him. In the second frame, the situation is depicted in medium shot. This time, the audience are able to observe the represented participant (in this case, Fico) from a closer distance, yet Fico is still unreachable. The situation in the second frame is also presented at oblique angle, implying that what is seen in the frame is not the part of the audience world and that audience are not involved in the situation. However, Fico is depicted at medium-low angle, positioning him in a relationship in which he has power over the audience. When an element in the image is presented at low angle, the element is said to have power over the viewer (see Kress & van Leeuwen, 2006). Moreover, Fico does not direct a gaze to the audience, implying that Fico is only offered as an information to the audience (Kress & van Leeuwen, 2006). From the integration of these visual aspects in the second frame, it can be assumed that the audience are offered to see or observe Fico as a secret agent who is more powerful than them. However, the audience are not involved in the situation, implying that the secret agent world as depicted in the scene is not the audience's world.

The scene is illustrated in two different layout composition. The first frame, for instance, uses given – new layout composition. In given – new composition, the element placed at the left becomes the given information, while the element placed at the right becomes the new information (Kress & van Leeuwen, 2006). In the first frame, Ernest becomes the given information, while his car becomes the new information. Being the given information, Ernest represents and information which has been already known by audience. In this case, Ernest represents a character whom audience are already familiar. It is assumed that in real life, it is common to see people who wear more exquisite outfit or have fancy appearance like Ernest. On the other hand, fancy car is assumed to be something the audience do not encounter or see regularly. It is not an object all people can use every day, for fancy car is considered expensive object. In addition to this, in the frame Ernest's car appears more salient than Ernest himself. Due to its size, the car fills half of the frame, making it the most salient object or element in the frame. Thus, being the new, salient information, the car gains more attention than Ernest. The second frame uses center – margin layout composition. In center – margin composition, the element placed at the center of the image becomes the center or the core information and, in contrast, the element bordering the center become the margins (see Kress & van Leeuwen, 2006). Thus, in the second frame, Fico

	<p>becomes the core information and two cars positioned at his left and right becomes the margins. Fico carries the main messages, while the cars carry additional information which depend on Fico. As the core information, Fico is considered a portrayal of a secret agent who ‘has it all’ (power and money). Fico’s power is represented by the vertical angle from which Fico is presented, while the money is represented by Fico’s motorcycle. Two cars bordering Fico’s left and right side convey additional message regarding what a secret agent has (for instance, a secret agent has fancy cars). Without Fico as the core information, what is conveyed by the cars is only the fact that they are fancy cars.</p> <p>Regarding the relationship between the verbal and the non-verbal modes, the scene does not have dialogue lines. Moreover, there is no subtitle line for this scene. Since the scene is devoid of subtitles and dialogue lines, any messages in this scene are rendered through audio and visual mode only. The visual mode functions to portray the life and identity of the secret agents, while the audio mode (in this case, the background music) functions to intensify the meanings and create an ambiance. This implies that there is no relationship between the verbal and the non-verbal elements in the scene since both dialogue and subtitles in the scene are absent. As the result, to understand the meanings in the scene, audience can only grasp and guess the meanings from the moving images.</p>
--	--

Scene 15: The end of the hypnosis				
Time	Moving images	Indonesian dialogue	English subtitles	Sound and sound effect (sfx)
1:34:57 – 1:34:59		<b>Indro:</b> <i>Satu, dua, tiga</i> <b>[Narration overlap]:</b> <i>Dan pada saat yang tepat nanti..</i>	[1:34:58-1:34:59] And at the right time	Voice: Indro counts from one to three to end the hypnosis, speaking at high pitch level. The narration overlaps his dialogue. SFX: Punk rock music in background
1:35:01 – 1:35:02		<b>[Narration overlap]:</b> ... <i>kalian akan saya bangunkan kembali.</i>	[1:35:00-1:35:02] I will wake you up from your hypnosis.	Voice: Indro reads the narration at medium pitch level. SFX: Punk rock music in background
<b>Vector shift:</b> Indro → Secret agents (ex-robbers)  <b>Shot size and angle:</b> Close up → Medium-close shot. The first frame is shot in close shot and at frontal, eye-level angle. However, the second frame is shot in medium-close shot and at frontal, eye-level angle.  <b>Gaze:</b> No direct gaze at audience			<b>Back-translation:</b>  <b>Indro:</b> <i>Pada saat yang tepat saya akan membangunkan kalian dari hipnotis.</i>	

<p><b>Composition:</b></p> <ul style="list-style-type: none"> <li>- Frame 1: Given (the agents) – New (Indro)</li> <li>- Frame 2: Center (Indro) - Margin (the agents)</li> </ul>	
<p><b>Speech-Text-Image Relation</b></p>	<p>The scene illustrates Indro, the boss of the secret agents, trying to wake the agents from their hypnosis. After Indro and the agents manage to thwart the bank robbery, Indro ends the hypnosis, bringing the agents back to their consciousness and identity as secret agent.</p> <p>Based on the visual analysis, the vector in this scene shifts from Indro to the agents, as seen on the second frame. The vector is realized through Indro’s gaze at the agents. Since the vector is realized through gaze or eyeline, the narrative process is relational, positioning the participant who does the looking as the reactor and the participant who is looked at as phenomenon (see Kress &amp; van Leeuwen, 2006). In this scene, Indro is looking at the secret agents while counting down to wake them up from their hypnosis. Thus, this situation can be verbalized into the following sentence: “While counting down, Indro is looking at his agents.”</p> <p>The first frame portrays Indro in close shot, while in the second frame, Indro is presented in medium shot. In terms of shot angle, Indro is presented at frontal, eye-level angle in both frames. In addition to this, the represented participants (in this scene, Indro and the secret agents) do not direct a gaze to the audience. According to Kress and van Leeuwen (2006), close shot (or close up) represents intimate distance between the represented participants in the image and the viewer. Meanwhile, medium shot represents the distance between the represented participants and the viewer, which is social (see Kress &amp; van Leeuwen, 2006). Regarding the shot angle, frontal angle implies involvement of the viewer in the situation depicted—what is depicted is the part of represented participants’ and audience’s world (see Kress &amp; van Leeuwen, 2006). In addition to this, when an element or represented participant in the image is shot at eye-level angle, the image represents the equality of power between the element and the viewer (Kress &amp; van Leeuwen, 2006). Moreover, the absence of gaze represents an offer, positioning the represented participant in the image as an information for the viewer (Kress &amp; van Leeuwen, 2006). From the integration of these visual aspects (frame size, shot angle, and gaze), it can be assumed that in this scene, the audience are offered to get involved in the situation by observing what the represented participants (in this case, Indro and the agents) do. In the first frame, audience are allowed to interact with Indro in a close distance. In the second frame, audience are positioned further to Indro, yet they are still able to interact with him and observe the situation. However, the interaction is limited to observation only. This implies that the audience</p>

are not demanded to do something or believe in something; they are just there to observe the hypnosis.

The scene uses two different layout composition. The first frame is composed in given – new composition, while the second frame is composed in center – margin composition. In given – new composition, the element placed at the left becomes the given information, while the element placed at the right becomes the new information (Kress & van Leeuwen, 2006). In the first frame, the agent (appearing blurred) becomes the given information and Indro becomes the new information since he is positioned at the right part of the frame. Being the new information, Indro is expected to gain more attention from the audience and become their focus. Meanwhile, in center – margin composition, the element placed at the center of the image becomes the center or the core information and, in contrast, the element bordering the center become the margins (see Kress & van Leeuwen, 2006). In the second frame, Indro becomes the center or the core information, while the agents become the margins. As the core information, Indro portrays a main character who controls the hypnosis and becomes the focus of the audience. On the other hand, the agents (margins) also carry meanings; however, their meanings would depend on the center. Without Indro, the agents only represent people who are gathered in a dark room. Since Indro becomes the center, the agents then represent hypnotized people who are under control of Indro.

Regarding the relationship between the verbal and the non-verbal modes, the subtitles serve as the translation of the Indro's narration. However, Indro's dialogue line is left untranslated; only his narration is translated. Nevertheless, it is important to note that the narration and the subtitles function to add new meanings to the moving images scene the moving images, when observed alone, may not be easily understandable due to the possibility of meanings. The addition of narration and subtitles specifies the meanings in the moving images. According to Barthes (1977, in Kress & van Leeuwen, 2006), when a mode adds new or different meanings, the relationship between the verbal and the non-verbal mode is extension. Thus, in this scene the relationship between the verbal and the non-verbal mode is extension. The visual mode provides the illustration to of the situation, at the same time. Meanwhile, the subtitles functions to restate the meanings in the dialogue and adds new meanings to the moving images so the message in the scene becomes complete. Furthermore, the audio mode (the background music) functions to intensify the ambiance.



## Theatrical Poster of *Comic 8*



ABDURRAHMAN KARIM, 2016

*The Analysis of Amateur Subtitling of Comic 8 from a Multimodal Perspective.*  
Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu