How to Decode Gestures and Facial Expressions in Thai TV Drama for Audio Description

Kulnaree Sueroj

University of York, UK

Abstract: The Thai Audio Description (AD) Policy for television was officially published in 2016, but its practical implementation has been slow. A salient limitation is the lack of knowledge and research feeding into the creation of AD guidelines. Therefore, the main objective of this study is to focus on the AD strategies for television drama by studying how to communicate the meaning of gestures and facial expressions of characters to visually impaired audiences. Most of the themes in the television drama relate to human emotions, and people express their emotions through gestures and facial expressions. The difficulties of describing gestures and facial expressions audially are due to the complexities of cluster of physical actions. The concept of creating a character, the functional discourse of AD and the concept of gestures and facial expressions are applied to the textual analysis of the sample clips and their scripts. The methods used include three steps to decode gestures and facial expressions in order to select the key visuals to be described in the AD scripts. The first step is that, the entire scene must be divided into small parts, taking emotional changes into account. Secondly, to identify the functions of the audio description and finally to analyse the sequences of physical actions of the gestures and facial expressions, which represent the intent of the communication. These results suggest that the audio describer should not use a single strategy for all cases and the selection strategy of AD depends on four key conditions which were: 1. the sound gap length; 2. the appropriate position for the insertion of the AD; 3. the number of words used in AD writing must not affect the optimal speed of the description; and 4. the genre of the TV drama. Further research should study on how Thais people with visual impairment perceive gestures and facial expressions.

Keywords: audio description, gestures, facial expressions, emotions, television drama, sight loss, disability, accessibility

Introduction

A total of 191,264 Thai people were registered as being visually impaired with the Department of Empowerment of Persons with Disabilities in September 2020 (Mattawanukoon, 2020). All disabilities should have the same rights to access media as those without disabilities. From 2013 to 2020, audio description (AD)—a service to support people with sight loss to access visual information—has been available in Thailand on various platforms including television, films, plays, and museums (Karuchit, 2017; Sueroj, 2018). Television is a media which boasts high advancement of audio description, including in Thailand. This mass media provides accessible services in terms of priority, which contributes to policy implementation since a significant number of people consume various free television content (Leung, 2018). In 2016, the National Broadcasting and Telecommunications Commission (NBTC), as a Thai media regulator, published an official law supporting television access services including audio description. This is in line with the announcement made by the National Broadcasting and Telecommunications Commission regarding support and protection of the rights of people with disabilities to access television programmes (The National Broadcasting and Telecommunications Commission, 2016). Although the Thai policy on audio description has been enforced, in practice the proportion of audio description in Thailand has not increased significantly (Sueroj, 2018). The salient challenge is the limitation of knowledge. Most AD guidelines in Thailand are based on general basic principles and are not effectively applied for all television programmes (Karuchit, 2017; Sueroj, 2018). The present study tackles this issue by developing additional AD guidelines for specific genres. The television drama genre was selected in the study for two reasons. First, the number of television dramas with AD is likely to increase after the Thai media
regulator expanded the policy which previously required informative programmes such as documentaries to provide AD to include all programmes in 2020 (The National Broadcasting and Telecommunications Commission, 2020). Second, television dramas have always been the most popular genre in Thailand (The National Broadcasting and Telecommunications Commission, 2020). Most themes in television dramas relate to human emotions, which people express through their gestures and facial expressions (Lewis, 2012). These non-verbal communications must be provided through audio description since such expressions affect comprehension of the storyline for visually impaired audiences. In particular, the difficulties of describing gestures and facial expressions audially are due to the complexities of cluster of physical actions (Borg, 2013; Lewis, 2012). These reasons led to the study on Audio Describing Gestures and Facial Expressions in Thai television dramas. The content of this article is derived from the analysis of the gaps between theories and practice in the validation testing of the tool in this doctoral study. The following sections are presented with a literature review, research methodology, findings, conclusion, and discussion.

**Literature Review**

This study concentrates on dynamic audio description for television. Audio description (AD) for television is a dynamic audio description, which is a spoken commentary added to audio-visual media for visually impaired audience to enable a comprehension of visual elements which cannot otherwise be accessed by people with sight loss (Fryer, 2016). A discussion of the challenges of describing gestures and facial expressions and the advantages of related concepts for describing gestures and facial expressions are given in the following section.

<table>
<thead>
<tr>
<th>Concepts and Theories</th>
<th>Challenges</th>
<th>Benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Functional discourse of audio description (AD)</td>
<td>Time constraint is a crucial problem for television audio description (AD) since AD scriptwriters must provide AD synchronised with the original material. AD scriptwriters therefore need to select certain key visuals that affect the themes or storyline to describe.</td>
<td>There are three types of AD functions: (1) Discourse-supporting gesture and facial expressions. Gestures and facial expressions reinforce what is being said, and can also repeat what is being said. For example, if a boy is asked, &quot;How old are you?&quot; he could reply with four and raise four fingers at the same time (Mazur, 2014). (2) Discourse-filling gestures and facial expressions. Gestures and facial expressions complete what is being said, then to understand the speech, we need to see the gesture. For example, if a boy is asked, &quot;How old are you?&quot; he may not say anything but simply show four fingers (Mazur, 2014). (3) Discourse-conflicting gestures and facial expressions. This characteristic of this discourse involves gestures and facial expressions that contradict or do not support what is being said. For example, if someone says that they are going to the</td>
</tr>
</tbody>
</table>
It would be helpful to support the selection of key elements that need to be described by a limited sound gap. This means that certain parts of the scene require AD as a supporting function, so that AD can be omitted in a limited sound gap since visually impaired audiences understand the storyline from conversation. However, some parts of the scene require AD as a filling function and conflicting function. They cannot be excluded because there is no sound to help the interpretation of the visually impaired.

| 2. The concept of building character | As human emotions are very sophisticated, they are complicated to describe in AD. Moreover, many physical movements are often unconscious responses, so they are difficult to indicate and select for AD. | There are three steps of building character:

1. Identifying inner feelings.
2. Analysing sequences of physical actions in order to break the entire scene into fragments.
3. Selecting physical actions which are necessary to carry out the emotional purpose (Hodge, 2010; Stanislavski, 1989). This concept would be useful to simplify the complexities of gestures and facial expressions from the whole scene into small parts. These will support AD scriptwriters to identify appropriate gestures that reflect the characters’ true intentions. This approach is used to guide the visually impaired to interpret the messages in the same way as visually able audiences. |

| 3. The concept of Bit–Beat | Some scenes are continuous actions and include a lot of muscle movements to express feelings, so it is usually difficult to spot the turning point of emotions. | This concept would benefit AD scriptwriters in analysing the sequences of physical actions in order to break the entire scene into fragments by concentrating on "Bit"—the turning point of emotions—and "Beat"—the turning point of the level of emotions, including the order of appropriate sequences of emotional expressions. This makes it easier for AD scriptwriters to control meaning by describing gestures and facial expressions in order to communicate more accurately (Hodge, 2010; Stanislavski, 1989). |

| Benefits for writing AD script |

| 4. AD strategies for Gestures and facial expressions | There are five strategies for describing |
In addition, several reception studies (audience studies) describing gestures and facial expressions in films and television dramas show similar findings on the feasibility of AD's subjective style, which allows AD scriptwriters to interpret and convey the emotions of characters as verbal descriptions in AD format. The results of Chmiel and Mazur (2012), Leung (2018), and Vercauteren and Orero (2013) on audio describing gestures and facial expressions from reception research are similar. The participants preferred the use of emotive adjectives when dealing with emotions such as anger or fear, which is a subjective style of AD (Leung, 2018). These results explain how facial expressions are extremely difficult to describe objectively because the human face is able to express highly sophisticated emotions through the use of many muscles, so a single facial movement may indicate multi-faceted emotions (Chmiel and Mazur, 2012). Consequently, even if objective audio descriptions are accurate, the cluster of physical actions consisting of various muscles would be very difficult for visually impaired audiences to grasp the correct emotions (Vercauteren and Orero 2013). In the Thai context, there is no specific research that focuses on describing gestures and facial expressions, although there is a single study on audio description for Thai television drama. The mix-method research on audio description for TV drama series and perception of visually impaired audiences was published in 2017. One of the research objectives was to study the interpretation of visual codes in the audio description of TV drama series with AD's non-interpretative visual code and AD's interpretive visual code. The quantitative data collection was a survey that utilised a questionnaire to collect data from two groups. Each group consisted of 20 blind and partially blind participants which were tested with AD for TV drama in different versions. Qualitative data was gathered by
focus group. The participants were congenital blind, acquired blindness, and low vision. The results showed no statistical significance in the comparison of perception between AD's non-interpretative visual code and AD's interpretive visual code among all types of visually impaired people (Janevatchararuk, 2017).

**Research Methodology**

The present study uses a qualitative research methodology. The objective is to find effective AD strategies to describe gestures and facial expressions in Thai television dramas. The research question has the goal to achieve how to communicate the meaning of gestures and facial expressions of characters in Thai television dramas to visually impaired audiences. All data is collected through a textual analysis of the Thai TV dramas and their scripts or novels. In addition, the purposive sampling which is a type of non-probability sampling is used to determine the criteria for the selection of the sample clips (Neuman, 2003). The four criteria to choose Thai TV dramas’ sample clips are as follows:

1. Thai contemporary TV dramas in the digital age focusing on the period between 2018-2020. Digital transmission has had an impact on the styles of the Thai TV dramas, both in terms of storytelling and production (Kaewthet et al., 2018). It was noted that the modern themes of Thai contemporary reflect the way of life of people who had previously gone beyond the influence of mainstream interests. A new storyline was written using the out-of-sequence story writing in which a flash forward to a crisis is depicted at the beginning, then the rest of the story is presented in chronological order. Moreover, digital cameras were used and angles more closed (Kaewthet et al., 2018). The changes in storytelling and production have had an influence on AD strategies, so it is important to concentrate on Thai contemporary TV dramas in the digital age.

2. Thai television dramas broadcasted in primetime slot. Thai TV dramas broadcast in prime time could be a trendsetter. The peak time of Thai TV dramas is around 7:00 p.m. to 10:30 p.m. It is expected to draw a wide audience (Thairath, 2020). Whenever some drama storylines become popular, they will be created in the same style by other producers. Compared to international AD research on television, prime time programmes are also worth studying as a trendsetter (Chmiel and Mazur, 2012).

3. Sample clips of Thai TV dramas that represent the main challenges of describing gestures and facial expressions of characters in Thai TV dramas. There are five main challenges: 1. there is a small sound gap to insert AD; 2. there are too few conversational or musical clues that support the unsighted audience’s interpretation; 3. the characters often and suddenly change mood; 4. the expressions of characters are to conceal their real intentions; and 5. the different styles of acting in different genres. Mehrabian (1981) identified three elements of communication that are body language, voice, and words. About 55% of the meaning of any message comes from visual body language, including gestures, postures, and facial expressions. About 38% of the meaning is vocal, derived from the non-verbal element of speech such as tone, pitch, speed, and pace, and 7% of the meaning comes from the actual words, like facts and data (Borg, 2013; Mehrabian, 1981). AD scriptwriters must find an appropriate sound gap to insert AD to complete 55% of the meaning conveyed in gestures and facial expressions of the characters, which does not overlap with 45% of the meaning of the original sound (Borg, 2013).

4. The clips that the researcher can access the serial scripts of sample clips or their original novels. The actors are working on a script, so they know the subtext. These scripts can reveal the real intent of characters in TV dramas, so this explains why the study of serial scripts is relevant.

In this article, the researcher demonstrates an analysis of a highlight clip from the Thai TV drama My Husband in Law (2020) which was published on the official YouTube channel of channel 3.
The Findings

Data was collected and analysed through textual analysis based on three key principles, Stanislavski’s concept of creating a character (1989), the functional discourse of AD (Mazur, 2014), and the concept of gestures and facial expressions (Borg 2013; Lewis 2012). Two main issues are derived from the findings: 1. the process of analysis of gestures and facial expressions; and 2. the techniques used in AD scriptwriting to describe gestures and facial expressions.

1. The process of analysis of gestures and facial expressions.

The methods used include three steps to decode gestures and facial expressions in order to select the key visuals to describe in an AD script.

Step 1: The entire scene must be divided into small parts, taking emotional changes into account. This step is useful to convert a proper understanding of gestures and facial expressions into AD forms and to transmit these non-verbal communications to visually impaired audiences. This is because it is difficult to locate and select key visuals to convey the key message to be described in AD format. As a result, AD scriptwriters must break the whole scene into small sections by concentrating on “Bit”, the turning point of emotions, and “Beat”, the turning points of level of emotions, of characters in each scene. The following section will provide an example of using “Bit” and “Beat” to break the whole scene into small parts.

The sample scenes were selected from a Thai TV drama called My Husband in Law. Its genre is a secret love story that integrates drama and comedy. The synopsis is that a multi-talented Muey (female protagonist) has been living with her crush, Thien (male protagonist), for over seven years. During that time, her crush deepens into love, but Thien is utterly indifferent to her feelings. Unfortunately, this smart, rich, and good-looking Thien faces a big problem when he has slept with the wife of a dangerous mobster. Thien’s life is in danger in the aftermath. Thien’s mother forces him to marry Muey to protect him. Muey accepts this condition because it is her only chance to be closer to him. Although Thien reluctantly agrees and marries Muey, he wants to keep the relationship secret (Hypercuz and Queen, 2020).

In the previous scene, Thien knew that Muey had been secretly in love with him, as he had read her diary, but pretended to not know about it. He told his mother that he planned to divorce Muey as soon as possible since he didn’t want Muey and he chose to marry someone else. He knew that Muey unintentionally overheard that conversation. Then the plot goes on to the sample scene, which is 7.47 minutes in length. What follows is a brief description of the acting of the characters and their conversation in the scene:

Thien and Muey hold their umbrellas and walk back home together on a rainy day.

Thien: “Did you hear that?” (his plan to divorce her)

Muey: “Yes, I heard, you said that I was not a woman you wanted to get married to, but no problem. I’m all right.” (She forces a smile)

Muey walks in front of Thien and she slips. Her leg is hurt and Thien offers to carry Muey on his back to get home.

While on Thien’s back, Muey cries hard and pretends to be all right, but she cannot stop her tears.

Muey: “Let’s go and play in the rain.” (She said this because she does not want Thien to know about her crying)
After implementing the principle of Bit and Beat from Stanislavski’s concept, the researcher divided the turning points of the actor’s emotions into three parts: Bit 1 when Thien (actor) and Muey (actress) walk together and Thien feels guilty; Bit 2 after Muey slips, Thien feels caring; Bit 3 after Muey is carried on Thien’s back, Thien feels confusion. The researcher divided the turning points of the actress’ level of emotions into three parts: Beat 1 when Thien (actor) and Muey (actress) walk together, Muey feels sad; Beat 2 after Muey slips, Muey is sadder; Beat 3 after Muey is carried on Thien’s back, she is extremely sad.

Step 2: Identify the functions of the audio description.

After the entire TV drama scene is divided into small parts in the first step, this second step helps AD scriptwriters select the essential parts that need to provide AD. First, AD scriptwriters must observe both camera shots, the acting of the characters, and examine the full script. This is done to explore the true intent of gestures and facial expressions of the characters. Then the concept of the functional discourse of AD is used to analyse the AD functions that represent the significance of AD. There are three AD functions, namely supporting function, filling function, and conflicting function (Mazur and Chmiel, 2012).

Supporting function: AD is a supplement of understanding for visually impaired audiences. It can be omitted because people with sight loss can understand the meaning of visual elements from the original sounds such as conversation or dialogue (Mazur, 2014). For example, when someone asks a woman, “Are you OK?” she said yes, and she nods. Her emotion is really happy. If AD scriptwriters provide AD for describing nodding, AD is a supporting function.

Filling function: AD is provided to complete the meaning of gestures and facial expressions because there is no sound to support the interpretation for people with sight loss. For example, when someone asks, “Are you OK?” B: (does not say anything) she nods. Her emotion is happy. If AD scriptwriters insert AD for describing nodding, AD is a filling function.

Conflicting function: AD can remove misunderstanding of people who have lost their sight. When a character expresses two or three gestures that have opposite meanings, AD scriptwriters must use AD in order to avoid audience confusion (Mazur and Chmiel, 2012). For example, when someone asks, “Are you OK?” C: (does not say anything) she nods and cries. Her intention is sad but pretends to be OK. So, AD in this case is a conflicting function and AD scriptwriters need to provide AD.

It can be clearly seen that when AD scriptwriters can categorise the function of AD they can understand which parts of a TV drama require the insertion of AD.

Step 3: To analyse the sequences of physical actions of gestures and facial expressions, which represent the intent of the communication.

People generally convey emotions through a cluster of physical actions so AD scriptwriters need to observe all the actions of characters and select key actions to be prioritised. There are many facial muscles that convey the emotions in both the upper face, such as the eyes, head, forehead, eyebrows, and nose as well as the lower face such as the cheek, mouth, or lips. For example, when people are happy, their facial expression changes so that the corners of the mouth lift to form a smile. As the eyelids tighten, the cheeks rise, and the outside corners of the eyebrows are pulled down (Lewis, 2012). Meanwhile, when people are sad, their eyelids droop as the inner corners of the eyebrows rise and draw together. The corners of the lips pull down and the lower lip may push up
in a pout (Lewis, 2012). There is a very interesting observation that people can control their actions and communicate their intentions in opposition to them, but their eyes are difficult to control, so the eyes can represent a person’s true emotions (Borg, 2013; Lewis, 2012). Besides identifying the key message to be described, it is important to write a communicative AD script, and the priority of physical action is also important to convey the meaning of gestures and facial expressions.

For instance, when someone asks, “Are you OK?” C: (does not say anything) she nods and cries. Her intention is sad but she pretends to be OK. In this case, AD scriptwriters need to concentrate on her eyes and her head, because these two actions can reflect the sense of sadness while the character pretends to be OK. AD scriptwriters should describe the appearance of the head and eyes respectively because the actress in this scene attempted to convince the actor otherwise by nodding, but she clearly attempts to hold back her tears and conceal her sadness. For example, she nods but her tears well up in her eyes, or she nods but cries. The description style, which provides less or more detail, affects the comprehension and feelings of visually impaired audiences to consume the verbal descriptions. Time constraints are an important condition that impacts the styles of AD scripts.

2. The techniques used in AD scriptwriting to describe gestures and facial expressions.

According to textual analysis of TV drama clips and their scripts, there are six styles of audio describing gestures and facial expressions: 1. to describe the details of physical actions - for example, the eyes widen, and the upper lids rise (Lewis, 2012); 2. to describe the specific expressions - for example, she raises her right hand up; 3. to describe the general expressions - for example, she moves; 4. to indicate the emotions - for example, she is sad; 5. to identify the signs - for example, she shows OK or she gives a big thumbs up; and 6. to combine - for example, she is in fear and her eyes widen, and the upper lids rise (Lewis, 2012). The appropriate selection of AD strategies to describe gestures and facial expressions should be considered under four conditions: 1. the sound gap length; 2. the appropriate position for the insertion of the AD; 3. the number of words used in AD writing must not affect the optimal speed of the description; and 4. the genre of the TV drama.

Conclusion

There are several challenges in describing gestures and facial expressions in Thai television dramas. First and foremost, there are a lot of mixed feelings shown in drama scenes, and they can be interpreted with a variety of meanings, including some facial expressions, which make it difficult to identify. Therefore, the analysis of TV drama scripts is important to understand the real intent of characters. In addition, many characters often change their emotions and express their feelings as opposed to their true intentions. Thus, a literal description might not be able to completely eliminate the confusion of meaning. Furthermore, most Thai TV dramas face the condition of time constraint with a very short sound gap, i.e. 3-4 seconds, but they need to include key message to describe the situation. Therefore, AD scriptwriters also need to apply the AD strategies such as functional discourse to support the selection of AD.

In conclusion, the three-step analysis of the clips supports AD scriptwriting process on describing gestures and facial expressions in four functions: 1. to simplify and clarify the complexity of gestures and facial expressions that reflect the characters’ intention in each scene; 2. to facilitate the process of selecting key visuals, the core message must be expressed in a time constraint condition; 3.to facilitate the effective transmission of meaning of gestures and facial expressions; and 4. to reduce the differences in the interpretation of gestures and facial expressions between sighted AD scriptwriters and unsighted audiences.
Discussion

1. Most Thai guidelines typically provide guidance about what elements should be chosen to describe or what materials should be analysed, but no advice is provided regarding how to analyse them. The findings of the present analysis provide a method for reviewing television dramas prior to AD scriptwriting. According to Thai AD guidelines, it is often recommended to write AD scripts objectively and avoid interpretation. Yet in practice that is rather difficult with time constraints. In particular, describing gestures and facial expressions with AD’s objective style may lead to confusion for some visually impaired viewers who have lost their mental images.

2. The audio describer cannot use a single strategy for all cases, but the selection strategy depends on four main conditions which are: 1. the sound gap length; 2. the appropriate position for the insertion of the AD; 3. the number of words used in AD writing must not affect the optimal speed of the description; and 4. the genre of the TV drama. Leung (2018), Mazur and Chmiel (2012), and Vercauteren and Orero (2013) found that participants prefer emotive descriptions of gestures and facial expressions that are subjective to the AD style. Moreover, people with vision loss may find it difficult to understand the meaning of the objective description and it is a major challenge for them to fully understand its meaning. Further research on the reception of Thai visually impaired audiences should also be conducted to understand how Thai people with vision loss perceive gestures and facial expressions in Thai TV dramas.

3. AD reception research with visually impaired people needs to include the effectiveness assessment of AD strategies. Audio description production relates to three main stakeholders which are original senders, AD scriptwriters, and users. The sender of the audio-visual product encodes the meaning of the story being told through audio-visual elements for sighted viewers. Therefore, sighted viewers must decode the meaning through both audio and visual elements. Meanwhile, unsighted people can only access the content through audio. They may have different ways of understanding the meaning of words compared to sighted audiences. The background of knowledge including experiences impact the decoding of receiver. In the case of audio description, two groups of people have different experiences from different sensory inputs. This affects their interpretation, especially perception of gestures and facial expressions (Chmiel and Mazur, 2012). AD strategies for TV dramas must therefore be evaluated by people with visual impairment because such people are able to construct visual imaginary and they have different ways to process facial expressions compared to sighted people.

Acknowledgement

The article How to decode gestures and facial expressions in Thai TV drama for audio description is a part of doctoral research was funded by Thammasat university. Special thanks to my supervisor, Dr. Mariana Lopez, and to all people involved in my article.

ORCID

Kulnaree Sueroj https://orcid.org/0000-0001-8163-5857

References

How to Decode Gestures and Facial Expressions in Thai TV Drama for Audio Description


