

# CHAPTER I

## INTRODUCTION

### A. Background

Speaking is a skill in oral mode, and it's liked any other skill, is more complicated than it seemed at first and involves more than just words. Many language students consider that speaking ability as a measure of knowing a language. These students define fluency as the ability to communicate with others, more than the ability to read, write, or understand spoken the language. They consider speaking to be the most important skill they can get, and they assess their progress in terms of their achievement in oral communication. Communication is intrinsically multimodal since various semiotic resources are intertwined and all contribute to the meaning-making process in a given situational context so that both verbal and non-verbal cues are crucial to interpret a message fully (Kress and van Leeuwen 1996 and Lemke 1998). However, teacher of English often ignore the importance of speaking in the classroom. Whereas, students need speaking skill to communicate well with using English to find out some information when they faced a foreigner or when they got an assignment from the teacher. Therefore, the needed for speaking and taking out the information from communicating seems to be extremely important for students.

As we know that English in Indonesia taught as a foreign language; hence, there are some problems in teaching speaking — one of the problems that can find the method used by the teachers. Many kinds of teaching method may result in bad or negative effects on students' speaking skill. So many students are hard to get the meaning from the conversation. Another problem is the lack of students' participation in the classroom. For example, when the teacher dominates the activity in the classroom, the students may feel like they do not have enough chance to participate actively. The lack of participation may

result in a lack of motivation. And lack of motivation may influence students' speaking activity.

Based on the title above, the Researcher decides to conduct the observation in Senior High School level with the consideration that there were still many students who were confused to learn English, especially in speaking skill. Many of them are still confused about how to communicate well with using correct grammar, tenses, and also the structure of sentences want they speak.

Several researchers have researched regarding multimodal. First, research by Soujanya Poria, Erik Cambria, Newton Howard, Guang-Bin Huang, Amir Hussain (2014) discusses an effective way for students to learn how to speak with correct structure by seeing visually students can understand by looking at the situation, expression and the ways how to say it.

One of the problems may be that the students feel bored because the techniques used by the teacher are monotonous. Many students felt sleepy when the teacher explained the materials. It may be because the teachers do not use good technique or method in the class. It is necessary for an English teacher to make speaking materials more interesting and motivating so that the students can progress in speaking.

In this research, the researcher wants to make a situation where the students feel to enjoy the learning process so that they will be motivated to speak English. The researcher tries to apply the ways that may be effective to improve the students' speaking ability to solve the problems to make a conversation. The researcher used Multimodal Listening to involve the students in speaking learning in the classroom. Multimodal listening is an effective way for students to learn how to speak with correct structure, by seeing visual students can understand by looking at the situation, expression, and the ways how to say it. The purpose of using Multimodal Listening is to develop students' mindset to become creative and make them able to speak

English well. It makes the students independent. It also keeps the existence of high-level responsibility.

## **B. Research Questions**

About the background of the problem above, the following research questions can identify:

1. How does in-service teacher use multimodal listening to improve her students' speaking?
2. How do EFL students perceive multimodal listening as a media to improve their speaking?

## **C. Purpose**

1. To find out enthusiasm and desire of students to learn and practice what they have learned in the class.
2. To find out the students' willingness to learn English, especially in speaking by Multimodal Listening.

## **D. Research Significances**

1. For the students, the researcher hopes that with Multimodal Listening, the students can develop their English skills, especially in speaking.
2. For the teachers, the benefits obtained are that the teacher can more easily to convey the speaking material more creatively so that the students in the class will not feel bored, and with Multimodal Listening the teacher can make the student more understand about the material.

## **E. Theory Framework**

Multimodal conversational is much more complex than spoken communication agents (Lopez-Cozar& Araki, 2005). Taking into the available modalities interaction can be carried out employing device such a microphone, keyboard, mouse, camera, loudspeaker, display, and glove or haptic hardware (Wahlster, 2005). Multimodal learning has been an important and challenging problem for decades, which aims to bridge the modality gap between heterogeneous representations, such as vision and language.

Many multimodal agents adopt a graphical human, like appearance to provide a more natural and friendly interaction to the user. Depending on the portion of the body shown on screen, they are usually called Talking Heads or Embodied Conversational Agents (ECAs). These are characters that provide auditory and visual feedback, which is particularly useful when the interaction takes place in a noisy environment. Their complexity varies significantly in terms of sophistication and complexity, from simple cartoon-like to complex animated human faces. The characters will be connected to the modules of conversation agent that generate information using the output modalities, for example, speech synthesis, lip movements, facial gesture, and video or image on display (Malaka et al., 2004).

User simulator is a technique that has attracted increasing interest in the last decade for the evaluation of the conversational agents is based on automatic generation of dialogues between the agents and an additional module, called user simulator, which represents user interaction (Zukerman & Litman, 2011; Lopez & Cozar et al, 2003; Schatzmann et al., 2006; Griol et al., 2009). The simulator makes it all possible to generate a large number of dialogues in a very simple way.

In 1970, Ekman et al. carried out extensive studies on facial expressions. Their research showed that universal facial expressions provide sufficient clues to detect emotions. They used anger, sadness, surprise, fear, disgust, and joy as six basic emotion classes. Such basic affective categories are sufficient to describe most of the emotions expressed by facial expressions. However, this list does not include the emotion expressed through facial expression by a person when he or she shows disrespect to someone; thus, a seventh basic emotion, contempt, was introduced by Matsumoto. Ekman et al. developed a facial expression coding system (FACS) to code facial expressions by deconstructing a facial expression into a set of action units (AU). AUs defined via specific facial muscle movements. The ability to perform multimodal fusion is an important prerequisite to the successful implementation of agent

user interaction. One of the primary obstacles to multimodal fusion is the development and specification of a methodology to integrate cognitive and affective information from different sources on different time scales and measurement values. There are two main fusion strategies: feature-level fusion and decision level fusion.

Videos collected from the social media website YouTube. Videos in the dataset were about different topics (for instance, politics, electronics product reviews, etc.). The videos were found using the following keywords: opinion, review, product review, best perfume, toothpaste, war, job, business, cosmetics review, camera review, baby product review, I hate, I like. The final video set had 20 female and 27 male speakers randomly selected from YouTube, with their age ranging approximately from 14-60 years. Although they belonged to different ethnic backgrounds (e.g., Caucasian, African-American, Hispanic, Asian), all speakers expressed themselves in English. The length of the videos varied from 2-5 minutes. All videos were pre-processed to avoid the issues of introductory titles and multiple topics. Many videos on YouTube contained an introductory sequence where a title was shown, sometimes accompanied by a visual animation.

The advantages of using multimodal corpora have been extensively describing in the literature (cf. Ackerley&Cocchetta, 2007; Knight, 2011; Querol-Juli an, 2010). Multimodal corpora are much more limited in size. This is mainly due to the fact that the compilation process is much more complex, as the addition of aural and visual information means that “aligning and transcribing (if at all) different streams of data is naturally more time consuming and technically difficult than when dealing with a single stream” (Knight, 2011: 397). Moreover, their scope is often limited because multimodal corpora, as well as ESP corpora, tend to be domain-specific (Knight, 2011).

## **F. Previous Research**

There are two previous studies concerning the development of students' speaking on classroom through Multimodal Listening. The result of previous studies was useful as a reference for consideration in this present study.

The first previous research, entitled "Multimodal Listening Through Movie Trailers: Towards a Framework for Classroom Implementation" is conducted by Jelena Bobkina and Elena Dominguez Romero (2018). The writers are more focused on the results obtained from the evaluation process revealed about the quality and limitations of multimodal Listening materials. Namely, mode Visual which provides important contextual information and non-linguistic input to guide students through the acquisition process while learning in the class. Many researchers emphasize their assessment as an authentic source of input with various possibilities for implementation in the classroom.

The second previous research, entitled "MAIA (movie analysis in action). A new teaching method in media literacy education" is conducted by Filippo Carlo Ceretti (2015). The writer focuses more on designing interesting learning situations in the classroom. So that teaching and learning activities can be fun, and easily understood by all students. therefore, we can design interesting learning situations so students can learn through audiovisual easily.

## **G. Methodology**

### **1. Research Design**

This research was conducted to find information about how effective Multimodal Listening in developing students' speaking in English learning activities. This study uses qualitative methods and a case study approach because it provides in-depth information. According to Creswell (2012), the research design is a special procedure consisting of data collection, data analysis, and report writing involved in the research process.

## 2. Research Site and Participant

Multimodal studies conducted at MAN 1 Bandung, West Java Province, Indonesia. Researchers consider the research site useful and effective because this institution is familiar to researchers. The familiarity between the researcher and the research site will make the research process more natural.

Study participants selected by purposive sampling technique Creswell (2012), in qualitative research, the researchers identified participants and sites in sampling-based on people and places that could best help them understand their main phenomena.

The researchers choose individuals and sites to learn or understand the main phenomena, and they can also choose samples or individuals representing the entire population. This study involved students of XI grade MAN as participants.

## 3. Research Instruments

Research instruments necessary to know the purpose of research. According to Creswell (2012), interviews and observations designated as data collection procedures in qualitative research. They choose the instrument of this research study.

### a. Observation

Observations were chosen to be research instruments. Creswell (2012) states that observation is the process of collecting information directly by observing people and placing them in the research location. Creswell also adds to the actual behavior and behavior found in the treatment process.

The research will be conduct on XI grade students at MAN 1 Bandung. In research, researchers act as examiners who want to do an experiment with students and record every activity carried out by students. By observation, the researcher will know to what extent students can understand the material provided.

b. Interview

Interviews are needed to validate the data that has taken. Semi-structured interviews with open questions are used in research to get in-depth information. Interviews were used in this study to examine the accuracy and find out whether students of XI grade MAN 1 Bandung could develop their speaking skills after the teacher used the multimodal listening to delivered the material. The researcher recorded interviews with audio recordings and notes to store information from the people interviewed.

4. Data Collecting Method

The researcher used two instruments in collecting data; observation, an interview.

a. Observation

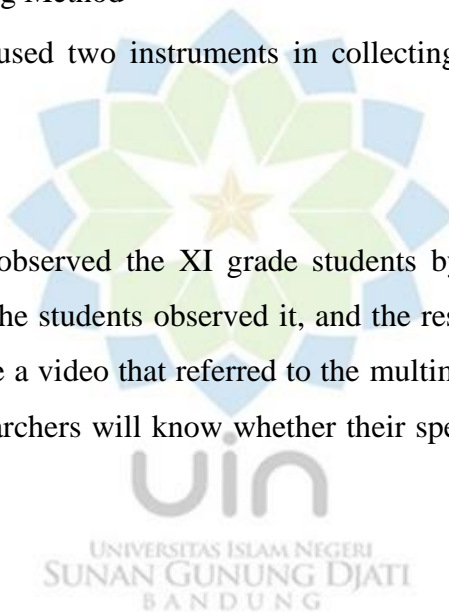
The researcher observed the XI grade students by giving a multimodal listening, and then the students observed it, and the researcher would instruct the students to make a video that referred to the multimodal listening that had given. So here researchers will know whether their speaking skills developed or not.

b. Interview

Researchers used semi-structured interviews to clarify specific research questions, why students choose to watch short films to hone their talking skills.

5. Analyzing Data

To analyze data, researchers used the Creswell theory (2012). The theory mentions several steps in analyzing data qualitatively.





### 1. Preparing and organizing the data

The researchers arrange data from observations and interviews. Observation videos, students work videos, and interview results transcribed for analysis.

### 2. Exploring and coding the data

Researchers are re-reading the data to get a general sense. Then the data is encoded based on the needs of the research question

### 3. Coding to build description and themes

Data coding is used to improve the description of MAN 1 Bandung class XI students in using multimodal listening in developing their speaking skills.

#### a. Observation

Data transcription observation is used to answer two research questions about how can students develop their speaking skill in the classroom by used multimodal listening.

#### b. Interview

Data transcription observation is used to answer two research questions whether by used the multimodal listening will the students' willingness to learn English, especially in speaking increase.

#### 4. Representing and reporting qualitative findings

The research findings from observations and interviews published in the narrative discussion to indicate whether students of class XI MAN 1 Bandung can develop their speaking skills after the teacher delivered the material by used multi modal listening, and the reason why it is advantageous to develop their speaking skills.

#### 5. Interpreting the findings

After representing and reporting the results, the researcher explained the meaning of the research. The interpretation included advancing personal view, comparison between conclusion and literature, suggesting limitation, and future research.

#### 6. Validating the accuracy of the findings

To validate the accuracy of the finding triangulation used. According to Creswell (2012), validation intends to have participants, external reviewers, or the data sources provide the evidence accuracy of the information in a qualitative report.