

CHAPTER I

INTRODUCTION

This chapter provides an initial overview of the research that will be conducted which includes the background of the study, research questions, purposes of the study, scope of the study, and significance of the study.

1.1. Background of the Study

Learning can be carried through many modes of communications (e.g. text, images, audios, videos). In recent years, the combination of the aforementioned modes has increased in learning materials as it is useful to communicate meaning and stimulate learning. The use of various modes of communication to create meaning is called multimodality (Kress, 2010). Not only is this approach receiving significant attention because it offers variety in learning (Healy, 2008), it also encourages student engagement because the use of multimodality in learning is more interesting and brings out excitement (James, 2014). There are several types of modes, e.g., linguistic, visual, aural, spatial, and gestural (New London Group, 1996). Among those aforementioned types of modes, linguistic mode (written or spoken mode, also widely known as verbal mode) and visual mode have been in the area of interest as they are widely used in learning materials such as educational videos.

The use of multimodality in English learning has been increasing over the years and it has become a distinctive highlight of modern teaching (Bezemer & Kress, 2016; Chapelle, 2016; Swan, 2018). The rising number shows that the use of multimodality in English learning has an effective impact on the learners. It has been reported that it can aid learners, build engagement, promote literacy in modern communication, and make

expressions of information more efficient in learning (Chaudhary & Mohan, 2019; Firtriana & Wirza, 2021; Sidabutar, 2021). Given the numerous researches on the development and benefit in multimodality, it shows how multimodality is valuable in learning.

Specifying to Indonesian context, according to a report by UNICEF, UNESCO, and the World Bank in June 2022, Indonesia experienced a massive learning crisis (e.g. reading inability, inability to understand simple written test, underperformance in literacy and numeracy) after the COVID-19 pandemic. Thus, in order to encourage learning recovery, a new curriculum is implemented. Schools in Indonesia has started implementing a new curriculum arranged by the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia (*Kemendikbudristek RI*) namely *Kurikulum Merdeka*. In consort with the new curriculum, teachers claimed that learning was not fully dependent on textbooks, but also a variety of learning materials. One of the mentioned varieties of materials is utilizing video in learning.

Videos are easily accessible in this modern times everywhere, especially since the convenience of an online video-sharing platform (i.e. YouTube) emerged. Teachers and students can both learn from the utilization of videos in the classroom, because they reveal more than just information. There are a lot of studies that has shown positive impacts of the usage of videos in facilitating learning. Videos can minimize the common problem of low motivation in learning due to fear and anxiety (Mahardika & Suwastini, 2023), improve the students' basic language learning skills, vocabulary, pronunciation, and their creative/critical thinking (Ariantini, et al., 2021), and accommodate students' various learning styles i.e. visual, auditory, and kinesthetic (Kaso, et al., 2023). Additionally, these videos can be produced by anyone for different purposes. Most importantly, many studies

point out that student-made videos are improving engagement and learning in the classroom. (Stanley & Zhang, 2018; Anas, 2018; Campbell, Heller, & DeMara, 2019)

Mykytiuk et al. (2025) conducted a study on the application of student-generated digital visual content for learning English as a second language (ESL) by undergraduate students enrolled in the course Introduction to Legal English to investigate the educational potential. In the study, the students were instructed to create visual content, supported by structured use of artificial intelligence (AI) to see whether it could improve second language learning outcomes. It was evident from the questionnaires that the students need to have a good linguistic competence, conceptual clarity, and creativity to receive a good AI visual content that they wanted. This shows that there was a strong connection between verbal narration and visual representation to convey students' understanding regarding the topic.

The use of creating videos as a medium to demonstrate students' comprehension has been around for a while (Bussard, 2016). A study conducted by Abdelhadi & Belaid (2025) investigated the use of student-produced videos as authentic assessment practice. The effectiveness of the use of these videos incorporates numerous areas of development (i.e. linguistic development, language skills, cognitive benefits, psychological engagement, creativity and digital literacy, student engagement & collaboration, and career & professional skills). These studies show that not only does the use of student-made videos are beneficial for identifying students' deficiencies in learning, but it is also evident to display students' performance and represent their comprehension. The use of videos has tremendously helped teachers, students, and even researchers to both consume and communicate information, which has shaped human's interaction, pattern recognition, and interpretation in various environment (Fazeli, Sabetti, & Ferrari, 2023).

In the context of university students who deals with real-life demonstration of their future role in their field, examining student-produced videos as a part of their assessment has been enforced as a way to enhance and evaluate their knowledge and skill acquisition (Lam & Yunus, 2023; Argudo-Serrano, et al., 2024). In addition to evaluating their skill and knowledge, these students will have an actual data that they rewatch and retract multiple times (Nagro & Cornelius, 2013) rather than trying to recall their personal feelings, memories, and thoughts of retellings of a lesson (Robinson & Kelley, 2007). The way these student-produced videos are being used to assess, analyse, and promote further improvements in teaching indicates that they certainly represent the creators' comprehension of the lesson delivered in the videos (Dalton & Smith, 2015).

Though some studies have exhibited the increasing use of videos and how they can improve learning, exploring the elements in the video and how they interact is crucial, yet complex (Jewitt & Price, 2012). Consequently, there is still a huge room for improvement on the clarity of how teachers and students can identify the elements of videos that contribute to improving learning and representing comprehension. Research by Fazeli, Sabetti, and Ferrari (2023) indicated the increasing number of videos in various format and sources, yet there is a lack of guidelines to analyze the multimodality of the video data. The void in lacking guidelines has been the major drive of their research to propose Visual-Verbal Video Analysis (VVVA) method.

VVVA method, adapted from Multimodal Theory by Kress (2010) and Visual Grounded Theory by Konecki (2011), provide systematic approach into dissecting the video data. The method employs six steps used for analysing video data: (1) collecting, organizing, and reviewing data, (2) transcribing verbal data, (3) choosing units of analysis, (4) extracting and coding data, (5) Organizing, describing, and interpreting extracted data,

and (6) reporting findings. This method brings identification of connection between verbal and visual aspect in videos into play. This connection between verbal and visual aspect reveals that there is cognitive process needed to produce those videos in order to promote learning.

Verbal and visual aspects of multimodality in learning are very crucial as they streamline communication and information expression (Sidabutar, 2021). Multimodality relies heavily on verbal and visual cues, such as texts and sounds, because students are learning to deconstruct how verbal and visual representations, in both viewing and producing videos, actively shape meaning (Chaudhary & Mohan, 2019). A study by Lim, et al. (2022) shows that visual resources do not function merely for an aesthetic purpose, but actively collaborate with verbal text to create meaning and help express complex concepts.

Research by Mayer (2005) reveals how connection between verbal and visual aspects takes place. The study reveals that meaningful learning depends on the ability to carry out three important cognitive processes (selecting, organizing, and integrating) of those verbal and visual aspects in a memory store (Mayer & Fiorella, 2022). The core issue is that Mayer's Cognitive Theory of Multimedia Learning (CTML) was originally designed to study how people learn from multimedia, whereas this research looks at what happens when students create multimedia. Though we cannot see into their brains while watching a video, the videos can be analysed as an externalized artifact of their cognitive processing. Cognitive Theory of Multimedia Learning (CTML) is sufficient in framing this study as it can encompass how student-created video is a mirror of their cognitive architecture.

When students consume media, these three steps happen internally. When students create media, they must execute these exact steps externally. To produce a functional

video, students must decide which visual elements to select, how to sequence them linearly, and how to synchronize them with spoken/written text. Therefore, the final video artifact is not just a creative project; it is the physical footprint of the student's Selecting, Organizing, and Integrating (SOI) processes. In which, these active SOI processes can be reverse-engineered or tracked by using CTML. The core assumption in CTML is dual processing (verbal and visual channels). When analysing a student-produced video, CTML provides a clear lens to evaluate whether the student understands how these two channels interact. How a student structures these dual channels reveals whether their own mental model of the subject matter is integrated or disjointed.

Based on the mentioned occurrences of student-produced videos as an assessment in educational settings and a lack of clarity on how those self-made videos can display their own comprehension, there is a need to dive deeper into how student-made videos as externalized artefact of students' internal cognitive processes are observed. This present study aims to overcome the gap by using Cognitive Theory of Multimedia Learning by Mayer (2005) to shed light on how the verbal and visual representations represent students' cognitive processes and how they carry out learning that represents students' comprehension.

1.2. Research Questions

Based on the background of the study above, the formulated research questions are:

1. What type of verbal and visual representations are identified in the student-made videos?
2. What kind of cognitive processes are identified in the student-made videos based on the verbal and visual representations?

3. How do the cognitive processes carry out learning that represents the students' comprehension?

1.3. Purpose of the Study

Based on the research questions above, the purposes of this study are:

1. To identify types of representations in the student-made videos.
2. To recognize the types of cognitive processes in the student-made videos based on the verbal and visual representations.
3. To analyze how cognitive processes carry out learning that showcases students' comprehension.

1.4. Scope of the Study

This study is focused on the verbal and visual representations involved student-made videos to recognize the present cognitive processes which represent the students' comprehension. The study treated the student-made videos as representative artefacts of externalized cognitive processes and the study was not intended to appraise students' active thinking processes.

1.5. Significance of the Study

The findings in this study are expected to give a contribution for video creators, viewers, users, and other researchers. This study can add further insight about how elements in a video can interact to represent learning comprehension and create meaning, which gives a holistic approach as to how video creators convey information and audiences perceive and interpret content.

The findings of this study can practically help students show how their learning is delivered through their own created videos and can help teachers view the student-made videos to assess and provide feedback. This study can serve as additional points to investigate the elements used in the video, and bring forth further understanding about student-created videos.

