BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA
FACULTAD DE LENGUAS

An Analysis of Students' Digital Multimodal Text Production in the EFL Classroom

A Thesis Submitted to the Faculty of Languages

for the Degree of

Maestría en la Enseñanza del Inglés (MEI)

Presented by:

Blanca Yael Cardona Sierra

Thesis director:

Dra. Leticia Araceli Salas Serrano

Puebla, Pue. **November, 2021**

An Analysis of Students' Digital Multimodal Text Production in the EFL Classroom

This thesis has been read by the members of the committee of

Blanca Yael Cardona Sierra

And is considered worthy of approval in partial fulfillment of the requirement for the degree of

Maestría en la Enseñanza del Inglés (MEI)

Master in English Teaching

Dra. Leticia Araceli Salas Serrano
Thesis Director

Dr. Michael Thomas Witten
Committee Member

Dra. Teresa Aurora Castineira Benítez
Committee Member

Dra. Elizabeth Flores Salgado
Committee Member

Abstract

The covid-19 pandemic has led to a rapid spread and development of digital technology and an inevitable increase of digital multimodal text production in the English as a Foreign Language (EFL) classroom. Derived from the constant changes in the students' lifestyles, teachers need to change how they teach and not continue perceiving language as the sole mode of meaning-making (Kress, 2000). The latter cannot only be attributed to reading and writing standards but rather to the incorporation of multiliteracies, mainly multimodal literacy. The production of videos allows learners to construct meaning by combining multiple modes. Miller (2007) claimed that digital video composing is a "quintessential multimodal literacy that allows orchestration of visual, aural, kinetic, and verbal modes electronically" (p.66). Therefore, digital multimodal text composing can highly contribute to developing the skills required in 21st-century education such as critical thinking, creativity, collaboration and communication. The current study aims to explore EFL students' compositions of digital multimodal texts and analyze the semiotic resources students employed to construct meaning in their digital multimodal texts. In addition, this research seeks to examine students' perceptions of Digital Multimodal Composing (DMC) in order to explore the impact of video production on their learning process engagement. This study adopted an action research approach as its methodological framework, regarded as a qualitative approach to inquiry. Hence, the focus of this study was to initiate changes in a participative community to improve education practice and make students aware of their role when learning a foreign language. The data-gathering instruments included a focus group, a questionnaire, and a multimodal discourse analysis of the students' artifacts

Key words: Digital multimodal composing, multimodal literacy, semiotic resources.

Acknowledgments

First and foremost, I would like to thank God for all his blessings in my life. For the strength and faith, he gave me to accomplish this work.

I would also like to express my deepest admiration and gratitude to my thesis advisor, Dra. Leticia Araceli Salas Serrano whose experience and invaluable contributions guided me throughout this process. I am extremely grateful for her great deal of support, encouragement, motivation, and confidence to carry out this work as well as her assistance through each stage of this process. She consistently allowed this paper to be my own work but steered me in the right direction whenever she thought I needed it.

My gratitude extends to the thesis committee for spending time reading this work. I would like to acknowledge Dr. Michael Thomas Witten for inspiring my interest in multimodality since the first day of classes and also for his helpful comments and suggestions to improve my research. I am also grateful to Dra. Elizabeth Flores Salgado for her time, knowledge, and valuable comments. Additionally, I would like to express gratitude to Dra. Teresa Aurora Castineira Benítez for her insightful comments and suggestions. Thank you all for accepting to read my thesis providing me with the tools that I needed to improve and complete my work.

Last but not least, I would also like to thank my husband and my parents for their support and encouragement through all this process. They have been my inspiration and my greatest motivation in this journey. This accomplishment would not have been possible without them. Thank you all.

Dedication

I dedicate my thesis work to my husband and the love of my life, Misael, who has always believed in me even more than I believe in myself. Thank you for your support and encouragement during this process. I am truly grateful for having you in my life. I want you to know that you inspire me to be the best version of myself every day. Thank you, dear husband, for standing by my side and always bringing a smile to my face.

This work is also dedicated to my beloved parents, Mary and Edgar, who have always loved me unconditionally and taught me to work hard for the things I want. Thank you, God, for giving me, such wonderful parents and family. Thanks mom and dad for your infinite love and support. Thank you for always being there for me and encouraging me to give my best in everything I do.

I am thankful for your support throughout this process. I cannot express in words what you all mean to me. I love you with all my heart and soul.

Table of contents

Chapter III: Methodology

Chapter IV: Data analysis

Chapter V: Conclusion

List of tables, graphs and figures

Chapter I: Introduction

1.0 Introduction

The current study analyzes students' composition of digital multimodal texts in order to identify the semiotic resources learners employed to make meaning when producing multimodal texts. According to Van Leeuwen (2005), semiotic resources or modes are those actions and artifacts we use to communicate. These modes include facial expressions, gestures, language, image, music, oral/written language, speech, layout, videos, symbols, charts, colors, and more. Modes express meaning individually, but they can also merge into each other multimodally to orchestrate meaning. Multimodality, then, refers to the manipulation of different semiotic resources that work together to make meaning. In the words of Kress and Van Leeuwen (2001), multimodality is "the use of several semiotic modes in the design of a semiotic product or event, together with the particular way in which these modes are combined" (p.20). Therefore, multimodality can be an effective educational resource for EFL learners due to the integration of multiple modes. Multimodal practices in the EFL classroom can highly contribute to the development of language skills and digital literacy as well. Video production allows learners to construct meaning by combining multiple modes such as written and oral language, gesture, visual, sound, and movement. Miller (2007) claimed that digital video composing is a "quintessential multimodal literacy that allows orchestration of visual, aural, kinetic, and verbal modes electronically" (p.66). Thus, teachers and learners must be aware of the semiotic and digital resources they are exposed to in their daily lives and how they can take advantage of them. Digital video production as a multimodal literacy practice may broaden learners' and teachers' notions of school literacy from only reading and writing print to also composing visual and auditory texts (Miller, 2007). Thus,

incorporating multimodal literacy practices in the EFL classroom can enhance learners' language and digital skills.

Students are totally engaged in various multimodal texts due to their active participation in social media and increasing exposure to computer-mediated communication through the use of different electronic devices such as cell phones, computers, tablets, among others. The covid-19 pandemic has led to a rapid spread and development of digital technology and an inevitable increase of digital multimodal text production in the EFL classroom. Derived from the constant changes in the students' lifestyles, teachers need to change how they teach and not continue perceiving language as the sole mode of meaning-making (Kress,2000). As Prensky (2001), pointed out "our students have changed radically. Today's students are no longer the people our educational system was designed to teach." (p.1). Acknowledging the previously mentioned point, it is urgent to make changes to the way literacy is conceived. Meaning-making cannot only be attributed to reading and writing standards but also to incorporating multiliteracies, mainly multimodal literacy. Consequently, such pedagogies are necessary to prepare today's students to succeed in the 21st century. In order to achieve this goal, substantial changes are demanded in education regarding the implementation of multiple modes as a resource of knowledge and communication.

Even when the current generations have adopted technology as an essential aspect of their daily lives and know how to use it for communicating in social networks, this does not necessarily mean that they can use technology as a learning resource (Díaz-Barriga, 2020). Thus, learners need to incorporate digital and multimodal literacies into their everyday lives to such an extent that they can have the knowledge and the ability to use a wide range of technological tools for educational purposes, especially for learning a foreign language. EFL learners might start taking control of their learning process when producing digital multimodal texts, either working individually or

collaboratively. Therefore, the production of digital multimodal texts can highly contribute to the development of the skills required in 21st-century education such as creativity, critical thinking, collaboration and communication. This study aims to contribute to multimodal literacies by analyzing the semiotic resources students used to create digital multimodal texts as well as exploring their perceptions toward multimodal practices.

1.1 Significance of the study

This study will be significant in the way that it can highlight the importance of multimodality in the English language classroom and how the various semiotic resources implemented in the design and production of videos contribute to constructing meaning and allow learners to participate actively in their learning process. Besides, the study will provide English language teachers insights into the importance of exploring innovative and creative learning opportunities for students to become involved in creating multimodal artifacts. It may also generate useful implications concerning how language teachers can use Digital Multimodal Composing (DMC) in order to maximize the benefits of technology and new literacies for learning.

New literacies involve the use of digital technologies, multiple literacies, and modes beyond print literacy. Today's learners require the ability not only to read and write but also to know how to use the different digital technological tools effectively. We live in a changing world that relies ever more on digital technologies; thus, it is imperative to prepare students for the challenges that the 21st century may present. Learners no longer deal with print texts when reading and writing. Instead, they interact with multiple modes such as images, videos, audio, and different media such as mobile phones, computers, or tablets, just to mention a few; all of them offer unique ways to create and convey meaning (Spires et al., 2012). Therefore, it is necessary to incorporate these

new literacies and digital technologies into the curriculum and help students become proficient in the new literacies of 21st-century technologies.

Additionally, the production of videos enhances learners' creativity and promotes different values such as respect and tolerance to different ideas when working collaboratively. Multimodal practices foster students' active participation in their learning process and instantly allow them to become the principal actors in producing those artifacts. The great variety of technological resources demands critical media literacy skills from students. Thus, there is a great need to become more literate in multimodal material to aid teachers in engaging students in their learning process and gradually incorporating the production of digital multimodal texts in their daily pedagogical practice. Furthermore, the current study results will positively encourage English language teachers to not limit their teaching activities to traditional literacy (Forester and Meyer, 2015) and provide them with more resources to construct meaning and implement the multiple literacies that learners require to be successful.

1.2 Theoretical context of the research

This section describes how the current study draws on similar studies that scholars in this area have conducted in the past. It is imperative to point out that previous studies have proved the importance of developing our students' multiliteracies through digital multimodal composing practices (Illmi, Retnaningdyah & Munir, 2020; Husbye & Zanden, 2015; Miller, 2011; Cope & Kalantzis, 2009). In Singapore, Liang and Lim (2020) reported that it is necessary to keep working on developing instructional practices for multiliteracies in the classroom as well as training teachers in the design of multiliteracies lessons and pedagogical frameworks to guide their teaching of digital multimodal composing skills. In Taiwan, Yeh (2018) focused on how constructing multimodal products benefits students' multiliteracies. The results obtained in this

study showed that language teachers should create more opportunities for learners to become involved in creating multimodal artifacts such as digital videos. In another study carried out in Malaysia, Ganapathy (2016), found that using multimodal approaches in meaning-making promotes teaching and learning experiences that are inherently multimodal to acquire literacy skills necessary for today's world without being restricted to one mode of design. The findings from this study affirm that students' perception of multimodal approach lessons in an ESL classroom is highly engaging, self-directed, and learner-centered and promotes meaning-making with minimal guidance from the teachers. Hafner (2013) proposed the production of multimodal ensembles within an English undergraduate course at a university in Hong Kong. In this study, digital literacy practices are incorporated in the form of digital video projects. The findings suggest that the production of this digital multimodal text provided students with opportunities to engage in language learning and practice the 21st-century skill of orchestrating semiotic resources in various modes in order to make meaning through multimodal ensembles. Moreover, Jiang (2017) examined the affordances of DMC for EFL learning in a Chinese university. The findings showed that DMC is an essential learning activity that should be explored and integrated into EFL classrooms as a legitimate literacy practice since it facilitates EFL learning.

Although various studies have been carried out internationally to examine the use of multiliteracies and digital multimodal text production in the EFL classroom, few studies have been conducted locally. Since this study is different in settings and research methods, it aims to fill the gap by exploring the students' perceptions of digital multimodal composing and analyzing the multimodal resources they employed during the video production to enhance their multiliteracies.

1.3 Research setting

The research was conducted in a public high school "Centro de Estudios Tecnológicos Industriales y de Servicios" (CETis 17) which belongs to Dirección General de Educación Tecnológica Industrial (DGTI) and it is located in San Martin Texmelucan, Puebla. The participants were students enrolled in the fifth semester and were taking English classes five hours per week at the moment of the study. Overall, the participants' age range was from 17 to 18 years old. They possessed an A2 English level of proficiency in the English language according to the Common European Framework of Reference for languages (CEFR). This study implemented digital multimodal text composing to engage students in producing multimodal digital videos and explore their perceptions of the video creation process.

1.4 Aims of the research

The present study seeks to explore EFL students' composition of digital multimodal texts and analyze the semiotic resources they used to construct meaning. Additionally, this research aims to examine students' perceptions of digital multimodal composing in order to determine the impact of video production on their learning process engagement.

1.5 Research questions

Four research questions have been formulated in order to guide the study:

RQ1: What semiotic resources did the participants use in the digital multimodal text production?

RQ2: How did the different semiotic resources employed in the videos created by participants contribute to convey meaning?

RQ3: What are the benefits of implementing digital multimodal text production in public high school students?

RQ4: What are the public high school learners' perceptions of implementing multimodal videos in the EFL classroom?

1.6 Overview

Having presented the general topic and a discussion of the problem that this paper addresses, the work will move forward in the following manner. Chapter II presents a critical review of the literature that is relevant to this study. It is meant to provide a theoretical base to serve as a foundation for the subsequent analysis and conclusions to draw on. Chapter III outlines the methodological framework that the study employs to select, gather, and analyze data. It describes the research design, context, participants, data collection, and procedure used for this research. In Chapter IV, the analysis and results of the study are presented. Relevant results are highlighted to provide the reader with insights into the importance of exploring innovative and creative learning opportunities for students to get involved in the creation of multimodal artifacts such as videos. Finally, Chapter V provides a discussion of the findings and the implications that this research had on its particular context.

1.7 Conclusion

The current chapter provided a general overview of the use of multimodality in the EFL classroom to develop digital literacy and engage students in their English language learning process. Today's learners need to develop new skills that help them overcome the challenges that the 21st century may present. This chapter also introduced the significance of the study, emphasizing the

importance of multimodality in the English language classroom and how the various semiotic resources implemented in the design and production of videos contribute to constructing meaning. Besides, the current chapter outlined the aims of the study and provided the reader with a general idea about the theoretical context and research setting of this study.

Chapter II: Literature Review

2.0 Introduction

This chapter provides a critical review of the literature that is relevant to this study. It is meant to provide a theoretical base upon which the study is built to inform the reader of the pertinent theoretical issues underpinning this research and support the analyses and findings. Thus, this chapter provides a discussion around the concepts of multimodality, semiotic resources, and multimodal discourse analysis. It then describes Kress and Van Leeuwen's grammar of visual design. Finally, the literature review addresses digital multimodal composing, multiliteracies, multimodal literacy, and digital literacies.

2.1 Multimodality

Language has been considered the exclusive mode of communication and the principal resource for meaning-making. However, communication is not merely monomodally achieved but multimodally constituted by several modal resources (Canel, 2019). Modal resources are also called semiotic resources. Halliday (1978) coined this term and claimed that language is not just a set of rules but a resource for making meaning. Semiotic resources are those actions and artifacts we use to communicate (Van Leeuwen, 2005). Examples of modes include facial expressions, gestures, language, image, music, oral/written language, speech, layout, videos, symbols, charts, tables, and graphs. Bezemer and Kress (2008) defined mode as "a socially and culturally shaped resource for making meaning" (p. 171). Therefore, meaning is conveyed by various semiotic resources that depend on the particular context and interest of the sign-maker. Modes express meaning individually, but they can also merge into each other multimodally to orchestrate meaning. Multimodality, then, refers to the manipulation of different semiotic resources that work

together to make meaning. In the words of Kress and van Leeuwen (2001), multimodality is "the use of several semiotic modes in the design of a semiotic product or event, together with the particular way in which these modes are combined" (p.20). The emerging technological resources have triggered the constantly evolving of semiotic resources and how they shape and represent meaning. The concept of multimodality offers valuable insights into the linguistics field and allows us to understand that all communication is multimodal (Darvin, 2015). This fundamental tenet emphasizes that multiple and several modes complement each other to convey a more meaningful message.

Consequently, multimodality can be an effective tool in the EFL classroom, which provides new ways of talking about language and engaging students to learn it and practice it (Van Leeuwen,2015). Darvin (2015) asserted that multimodality plays a paramount role in language education and research. Besides, Álvarez (2016) claimed that "Second language studies have been impacted by 'the multimodal turn'" and that "multimodality is a response to the challenges linguistic description is facing in light of the changes in the way texts are designed, produced, and disseminated" (p. 100). Multimodality offers valuable tools that learners in the EFL classroom can explore. The following section describes this issue in detail.

2.2 Multimodality in the EFL Classroom

Multimodality has emerged to decentralize language as the sole mode of constructing meaning in the learning and teaching process. Thus, it is crucial to know how the semiotic resources integrate into multimodal artifacts or events (Van Leeuwen, 2005) in the EFL classroom. In this sense, teachers should take the initiative to promote students' multimodal communicative competence, which involves "the knowledge and use of language concerning the visual, gestural, audio and

spatial dimensions of communication, including computer-mediated-communication" (Heberle, 2010, p. 102). Subsequently, teachers and students need to understand how they can use multiple modes beyond language to construct meaning. Grapin (2018) proposed a weak and strong version of multimodality within the EFL classroom.

On the one hand, the weak version considers modes as language scaffolds to support English language development. These scaffolds are removed when students achieve the desired English language proficiency. According to this perspective, multimodal resources strengthen learning without being the main focus (Grapin, 2018). Based on this assumption, multiple modes are "generally considered illustrative supports to the 'real thing'" (Kress, Jewitt, Ogborn, & Tsatsarelis, 2014, p. 51). This perspective perpetuates the view that language is "the standard of precision in meaning"; therefore, all modes are inherently ambiguous (Lemke, 2002, p.321). Regarding the weak version of multimodality, learners, especially beginners can use it as a solid base to start using the language more dynamically and gain knowledge that can serve to develop more specific skills in the future.

On the other hand, the strong version views modes as essential semiotic resources that help learners engage in disciplinary practices (Grapin, 2018). Grapin (2018) endorsed the strong version when he claims that "it is not only necessary but transformative for ELs in the new content standards era, because it allows them to draw from the meaning-making resources at their disposal while engaging in disciplinary practices" (p.36). The implementation of multimodality in the EFL classroom allows English language students from different proficiency levels to take advantage of the multiple semiotic resources and their unique affordances for communicating disciplinary meanings (Grapin, 2018). As Choi and Yi (2015) pointed out, English language teachers should integrate multimodality into their teaching practices by encouraging English language learners to

represent their knowledge in multiples modes beyond the linguistic mode. In this view, Miller (2007) asserted that "knowledge is multimodal, co-constructed, and performed or represented, not absorbed" (p.65). Thus, multimodal learning enhances the skills students need to develop in the 21st-century classroom since a systematic multimodal practice gives as result the integration of various identities, languages, and cultures that are part of a classroom.

2.3 Multimodal Discourse Analysis

Multimodal Discourse Analysis (MDA) is a recent approach that focuses on the idea that meanings are created in texts and interactions in a complex interplay of semiosis across multiple modes (Bathia et al., 2008). Jones (2013) also claimed that MDA "is an approach to discourse which focuses on how meaning is made through the use of multiple modes of communication as opposed to just language" (p.1). MDA can explore different communication modes, such as videos, films, pictures, and advertisements. In this sense, the term mode refers to "the semiotic system with an internal grammaticality, such as speech, color, taste, or the design of images" (Scollon & LeVine, 2004, p. 2). This type of analysis considers how multimodal texts are designed and how the multiple semiotic resources contribute to conveying meaning in a text.

2.4 Semiotic Resources in Multimodality

Since we live in a multimodal age, there is an increasing awareness that meaning is not only constructed by one mode of communication, language. Meaning-making can be realized through a combination of semiotic resources. Semiotic resources, then, refer to:

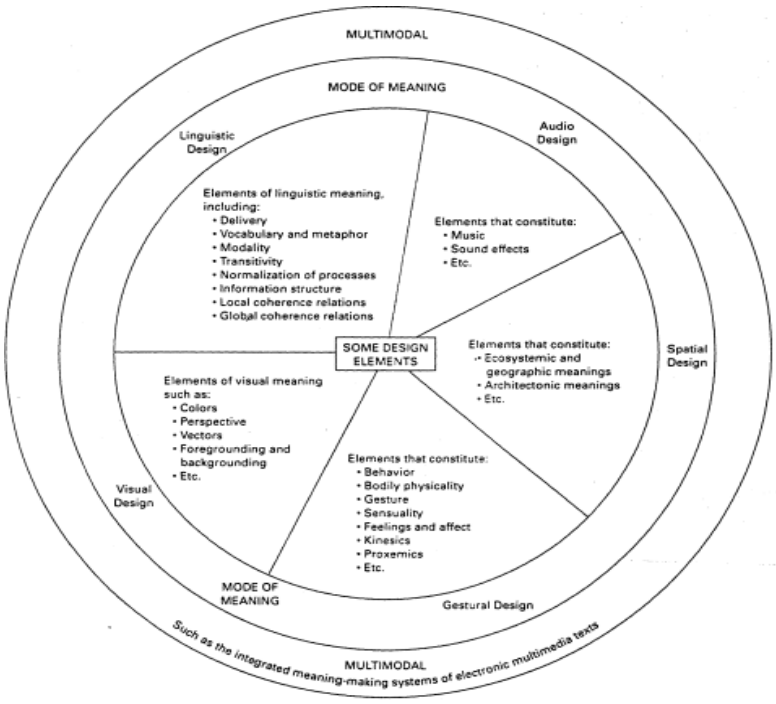
> the actions, materials, and artifacts we use for communicative purposes, whether produced physiologically – for example, with our vocal apparatus, the muscles we use

> to make facial expressions and gestures – or technologically – for example, with pen and ink, or computer hardware and software – together with the ways in which these resources can be organized (Van Leeuween, 2005, p.285).

People decide the modes or semiotic resources to express meaning based on what is available to them according to their particular situation, moment, and time. Therefore, semiotic resources are linked to context and social reality. According to Halliday's theory, Jewitt and Henriksen (2016) "language is a product of social processes; the resources of a language are shaped by the functions it has developed to satisfy the communicative needs of people's lives" (p.146).

The New London Group (1996) used the term "design" to describe the forms of meaning and proposed the design elements, which are divided into six major areas: linguistic (written or oral language), visual (images, layouts), audio (music, sound effects), gestural (body language, facial expressions), spatial (environmental spaces) and multimodal design. The latter is the most significant of all six designs since it combines two or more modes in a text, for instance, a video.

Figure 1. Design Elements of Different Modes of Meaning.



Note: taken from The New London Group (p.83, 1996).

Kress and Van Leeuwen expanded on Halliday's work notably developing a grammar of the visual which intends to compare images with language. In their book, Reading images: The grammar of visual design, Kress and van Leeuwen (1996) examined and established visual images as semiotic resources. The following section describes the grammar of visual design in more detail.

2.5 The Grammar of Visual Design

In their book Reading Images: The Grammar of Visual Design, Kress and Van Leeuwen (1996) pointed out that "visual 'grammar' will describe the way in which depicted elements – people, places and things – combine in visual 'statements' of greater or lesser complexity and extension" (p.1). Kress and van Leeuwen's (2006) framework to multimodality is grounded in the Systemic Functional Linguistics (SFL) of Michael Halliday. In this approach, language is perceived as "one of the semiotic systems that constitute a culture" (Halliday, 1978, p.2). It means that language

cannot be interpreted in isolation but within a sociocultural context. In this regard, Kress and Van Leeuwen stated that images can be analyzed according to the three metafunctions in SFL: ideational, interpersonal, and textual. These three metafunctions are not specific to speech or writing per se and can analyze visual semiotic resources. According to Halliday (1978), the three functional components of the semantic system 'metafunctions' are "the modes of meaning that are present in every use of language in every social context" (p.112). Halliday's three metafunctions all the time are being combining in order to produce texts. They are described below:

- Field or ideational/experiential metafunction represents our ideas and experiences about the world.
- Tenor or interpersonal metafunction describes the interactions created by semiotic modes. That is, the projection of social relations in the world and how they are enacted.
- Mode or textual metafunction organizes meaning within coherent and cohesive texts.

In Kress and van Leeuwen's framework to analyze visual images, the three metafunctions are adapted in order to reflect the fact that images "can 'say' (some of) the same things as language – in very different ways" (Kress & Van Leeuwen, 2006, p.50). In Visual Grammar, the three metafunctions are renamed as representational, interactive, and compositional. Therefore, the following sections explain the three metafunctions for better deconstructing and understanding video analyses.

2.5.1 Representational metafunction

The representational metafunction is about "the people, places, and objects within an image—the represented participants (RPs)—and answers the question 'What is the picture about?'" (Harrison, 2003, p.50). There are two types of participants involved in every semiotic act in the

representational metafunction, the represented participants and the interactive participants. The former is the people, places, or things represented in the visual, and the latter are the producers and the viewers of the visual (Kress & Van Leeuwen, 2006). All participants are part of various relationships, structures, and interaction processes that form two types of representational structures: the narrative and the conceptual. Moerdisuroso (2017) pointed out that Narrative Representation Structure (NRS) "presents the ongoing actions and events. Similar to the structure of a sentence that contains the elements of subject-verb-object, a subject in NRS called actor, and objects called goal" (p.87). In the Narrative Representation Structure, "participants are connected by a vector, they are represented as doing something to or for each other" (Kress & Van Leeuwen, 2006, p.59). On the other hand, in the conceptual representation, participants are represented "in terms of their generalized and more or less stable and timeless essence" (Ibid).

Narrative processes are categorized into action processes, reactional processes, speech processes and mental processes, conversion processes according to the kinds of vectors and participants involved. At the same time, the conceptual representations include classificational processes, analytical processes, and symbolic processes. The following table summarizes the basic structures and processes of this metafunction.

Table 1: Basic Structures and Processes of the Representational Metafunction

Structures	Processes
Narrative: Narrative images allow viewers to create a story about the RPs because the images include vectors of motion.	◆ **Action:** The narrative is created by vectors that can be bodies, limbs, tools, weapons, roads, and so forth. ◆ **Reactional:** The narrative is created by eyelines (acting as vectors) between RPs.
Conceptual: Conceptual images do not include vectors. Rather, RPs tend to be grouped together to present viewers with the "concept" of who or what they represent.	◆ **Classificatory:** RPs as "kind of" something or some group (that is, they are members of the same class). Advertisements for beauty products often have classificatory images such as a group of models (for instance, Revlon models). ◆ **Analytical:** RPs are displayed in terms of a "part-whole" structure. The "whole" is a Carrier who possesses "parts" called Attributes. The Supreme Court building in Figure 2 is a Carrier, and its architectural components are its Attributes. A pie chart is an analytical image in which the chart is the Carrier and its segments are Attributes. Diagrams are also analytical processes. ◆ **Symbolic:** RPs are important for what they "mean." A motorbike in an advertisement can, for example, be analytical (that is, asking the viewer to check out its attributes), but it is also symbolic of virility. Abstract shapes such as triangles, squares, and circles also fall in this category.

Note: taken from Harrison (p.51, 2003).

Besides the types of participants, structures, and processes, Kress and Van Leeuwen (2006) associated all these elements with specific circumstances divided into three types: locative circumstances (setting), circumstances of means, and circumstances of accompaniment. Locative circumstances relate other participants to specific participants in terms of foreground and background. Circumstances of means refer to "the tools used in action processes" (Kress & Van Leeuwen, 2006, p. 72). The circumstances of means concern the tools employed to carry out the actions; they are part of a vector and can be objects or even parts of the body such as a finger or hand. The circumstances of accompaniment have to do with people or objects presented during the action but not involved in any way with it, which is to say that they have "no vectorial relation with other participants" (Kress & Van Leeuwen, 2006, p. 75).

2.5.2 Interactive metafunction

According to Kress and Van Leeuwen (2006), the interactive metafunction is concerned with the interaction between the producer, the represented participants, and the viewer of the image; that is how participants interact with each other. Contact, social distance, attitude, and modality are four fundamental aspects of the interactive metafunction.

There are two ways in which gaze establishes contact between participants and viewers, demand and offer. The former is when the represented participants look directly at the viewer. As Kress and Van Leeuwen (2006) suggested, "the participant's gaze (and the gesture, if present) demands something from the viewer, demands that the viewer enter into some kind of imaginary relation with him or her" (p.118). The latter occurs when the represented participant is looking at other sides. In this case, "the viewer is not object, but subject of the look, and the represented participant is the object of the viewer's dispassionate scrutiny. No contact is made" (Kress & Van Leeuwen, 2006, p.119). The facial expressions and gestures play a crucial role when determining whether participants within an image perform demand or offer acts.

The second dimension to the interactive meanings of images is related to the distance between the represented participants and the viewer. The 'size of frame´ is used to determine the social distance through camera shots. Shots are classified into close-up, medium shot, and long shot. The choice of distance can suggest different relations between represented participants and viewers (Kress & Van Leeuwen, 2006, p.124). The close shot involves the participant's head and shoulders, the medium shot shows the participant's body approximately down to the knees, and the long shot shows anything more than that.

Perspective is yet another way in which images bring about relations between represented participants and the viewer. Producing an image also involves selecting an angle, a point of view,

which implies subjective attitudes. Subjective attitudes refer to the angles (horizontal or vertical) at which the participants are portrayed. While a frontal point of view involves the image-producer with the represented participants, the oblique angle indicates detachment and eye level denotes equality, and no power difference is involved (Kress and Van Leeuwen, 2006). Power is generally associated with vertical angles. High angles tend to diminish the individual, and low angles make something look impressive in the sense of superiority. Kress and Van Leeuwen (2006) claimed that:

> If a represented participant is seen from a high angle, then the relation between the interactive participants (the producer of the image, and hence also the viewer) and the represented participants is depicted as one in which the interactive participant has power over the represented participant – the represented participant is seen from the point of view of power. If the represented participant is seen from a low angle, then the relation between the interactive and represented participants is depicted as one in which the represented participant has power over the interactive participant. (p.140)

The following table demonstrates the details of the aspects previously mentioned.

Table 2: Basic features and processes of the interpersonal metafunction

Features	Feature Processes
Image Act and Gaze: The image act involves the eyeline of the RP(s) in relation to the viewer.	◆ Demand: The RP is looking directly at the viewer. A demand generally causes the viewer to feel a strong engagement with the RP. ◆ Offer: The RP is looking outside the picture or at someone or something within the image. In this case, the RP becomes an object of contemplation for the viewer, creating less engagement than that of the *demand*.
Social Distance and Intimacy: Social distance is determined by how close RPs in an image appear to the viewer, thereby resulting in feelings of intimacy or distance.	The viewer can see an RP in six different ways. ◆ Intimate distance: The head and face only ◆ Close personal distance: The head and shoulders ◆ Far personal distance: From the waist up ◆ Close social distance: The whole figure ◆ Far social distance: The whole figure with space around it ◆ Public distance: Torsos of several people
Perspective—The Horizontal Angle and Involvement: This angle refers to the relationship between the position of the RP(s) and the viewer.	◆ The frontal angle: When an RP is presented frontally to the viewer. This angle creates stronger involvement on the part of the viewer as it implies that the RP is "one of us." ◆ The oblique angle: When an RP is presented obliquely to the viewer. This angle creates greater detachment since it implies that the RP is "one of them."
Perspective—The Vertical Angle and Power: There are two possible vertical-angle relationships: 1) that of the RP(s) and the viewer, and 2) that between RPs within an image.	◆ High angle: The RP "looking up" has less power. ◆ Medium angle: The RP "looking horizontally" has equal power. ◆ Low angle: The RP "looking up" has less power.

Note: taken from Harrison (p.53, 2003).

The last dimension is modality, which is defined as "the truth value or credibility of (linguistically realized) statements about the world" (Kress and van Leeuwen, 2006, p.155). Modality deals with true representation of reality and is classified into three categories, high, low, and medium. Kress and Van Leeuwen (2006) argued that images that are closer to reality are considered to have a high modality. On the other hand, texts with a low modality contain less naturalistic aspects regarding color, contextualization, representation, and brightness. Modality is concerned with the following markers (Kress and Van Leeuwen 2006, p.160-63):

1) **Color saturation:** a scale running from full color saturation to the absence of color; that is, to black and white.
2) **Color differentiation,** a scale running from a maximally diversified range of colors to monochrome.

3) **Color modulation**, a scale running from fully modulated color, with, for example, the use of many different shades of red, to plain, unmodulated color.
4) **Contextualization,** a scale running from the absence of background to the most fully articulated and detailed background.
5) **Representation,** a scale running from maximum abstraction to maximum representation of pictorial detail.
6) **Depth,** a scale running from the absence of depth to maximally deep perspective.
7) **Illumination,** a scale running from the fullest representation of the play of light and shade to its absence.
8) **Brightness,** a scale running from a maximum number of different degrees of brightness to just two degrees: black and white, or dark grey and lighter grey, or two brightness values of the same color.

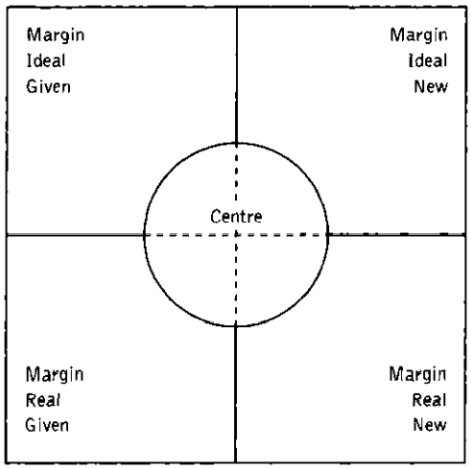
2.5.3 Compositional metafunction

The compositional metafunction deals with how the representational and interpersonal metafunctions integrate into a meaningful whole through three critical systems: information value, salience, and framing (Kress & Van Leeuwen, 2006). The first system, information value, refers to the placement of the various elements within an image. The placement of these elements is attached to different zones of the image, such as left, right, top, bottom, center, and margin.

The elements placed on the left are presented as ‘given’ and are known, commonsense or familiar by the viewers. Conversely, the elements placed on the right zone are presented as ‘new’ since they are unknown. Therefore, the viewer needs to pay special attention. The elements placed on the top side reflect the value of being ideal, whereas the elements placed at the bottom represent

the real. The elements placed at the center of the image provide essential information of the visual space, while the subservient elements placed around the center are considered margins. Kress and Van Leeuwen (2006) exemplified these dimensions through the following diagram:

Figure 2. The dimensions of visual space



Note: taken from Kress and Van Leeuwen (p.197, 2006).

The salience system involves factors such as size, tonal contrast, color contrast, sharpness of focus, and placement in the foreground or background of the represented participants in order to capture the viewer's attention. Kress and Van Leeuwen (2006) argued that "salience can create a hierarchy of importance among the elements, selecting some as more important, more worthy of attention than others" (p.89). Framing, the third key system in composition, is associated with connecting or disconnecting elements in the image through frame lines. The stronger the framing of an element, the more it is presented as a separate unit of information. The more the elements of the spatial composition are connected, the more they are presented as belonging together, as a single unit of information (Kress & Van Leeuwen, 2006). Table 3 illustrates the three systems of the compositional metafunction.

Table 3: Basic systems and elements of the compositional metafunction

System	Elements
Information Value: The placement of RPs allows them to take on different information roles.	◆ Left/Right: RPs on the left side of an image have the value of being "given" knowledge while RPs on the right are "new." ◆ Given = familiar, commonsense ◆ New = an issue, a problem, a solution (Note: This value is based on how we read in Western cultures, that is, from left to right. This does not necessarily apply to cultures in which reading occurs from right to left or in columns.) ◆ Top/Bottom: RPs at the top of an image have the value of being "ideal" while RPs below represent the "real." ◆ Ideal = emotive, imaginary, what might be, often the pictoral elements of an image ◆ Real = factual, informative, down to earth, practical, often textual elements in an image ◆ Center/Margin: RPs in the center provide the nucleus of information to which surrounding elements are subservient.
Salience: Salience refers to the ability of an RP to capture the viewer's attention.	◆ Size: The larger the RP, the greater the salience. ◆ Sharpness of focus: Out-of-focus RPs have less salience. ◆ Tonal contrast: Areas of high tonal contrast have greater salience. ◆ Color contrast: Strongly saturated colors have greater salience than "soft" colors. ◆ Foreground/Background: An RP in the foreground has greater salience than an RP in the background.
Framing: How RPs are framed affects whether they are seen as connected or separate.	◆ **Framelines: The lines within the image that divide RPs or hold them together.** ◆ **Pictorial framing devices: The stronger the lines around the image, the greater the connection.**

Note: taken from Harrison (p.57-58, 2003).

2.6 Multiliteracies

New communicative practices such as emails, mobile phones, social media, internet web pages, videos, and text messages have emerged due to the increase of technology, resulting in new literacies embodied in new social practices. These new ways of communication are focused on different domains besides language, particularly multimodal modes such as images, music, and body movements (Miller, 2007). The New London Group coined the term "multiliteracies" in 1996 to respond to the large number of communication channels derived

from globalization's social changes and the increase of cultural and linguistic diversity (New London Group, 1996). Thus, they suggested a pedagogy of multiliteracies whose primary purpose was to broaden literacy approaches and include not only linguistic practices but also multimodal textual practices within the teaching and learning process combining different modes. Consequently, the traditional literacy restricted to teaching and learning to read and write needed to be supplemented in the pedagogy of multiliteracies by learning how to read and write multimodal texts that integrated other modes beyond language (Cope & Kalantzis, 2009). One of the key elements of multiliteracies is the interconnection of different modes of meaning; the New London group (1996) identified six design elements that are paramount in the meaning-making process: linguistic design, visual design, audio design, gestural design, spatial design, and multimodal design. The latter results in the combination of two or more modes.

2.7 Multimodal literacy

Multiliteracies and multimodality are two terms that have been used interchangeably to talk about the multiplicity of communication modes and linguistic-cultural diversity (Yi, 2014). However, Rowsell and Walsh (2011) explained the relationship between the two as follows:

> Multimodality is the field that takes account of how individuals make meaning with different kinds of modes. Multiliteracies is a pedagogy developed by the New London Group… Multimodality comes first in that it informs how we make meaning, and multiliteracies, as a possible pedagogy, gives us tools for doing so…Multiliteracies as a pedagogy simultaneously accounts for linguistic diversity and the use of multimodalities in communication (p.55-56).

The Millennial generation, youth born in or after 1982 (Howe and Strauss, 2000), are different from previous generations, particularly in their use of Information Communication Technology (ICT). Since the millennial generation is immersed in an online culture, they perceive texts as multimodal resources for meaning-making and not just in terms of printed words, but also in terms of images and music (Miller, 2007). In response to these changes, multimodal literacy practice "reframes pedagogical goals to focus on connecting out-of-school literacies of students through purposeful multimodal design activities in social spaces that engage student lifeworlds and transform classroom learning" (Miller, 2007, p.62).

Multimodal literacy practices significantly impact the teaching and learning process of a foreign language and how students learn and construct knowledge. For instance, learners make videos, digital comics, posters, PowerPoint presentations, and manipulate a wide range of digital tools while orchestrating multiple modes to communicate their ideas effectively. However, English language teachers have not directed their attention to multimodal literacy practices and continue perceiving language as the sole mode of meaning-making (Kress, 2000). Consequently, such pedagogies are necessary to prepare today's students to succeed in the 21st century. Therefore, substantial changes are demanded in education about implementing multiple modes as a resource of knowledge and communication. When students engage into multimodal practice such as the production of videos, they should take into account the multiplicity of modes they are exposed to and how they can be interwoven and orchestrated to communicate meaning effectively (Yi, 2014). Learners need to analyze the role that every mode plays and how it contributes meaning-making. By doing so, students need to improve their multimodal communicative competence. Nelson (2006) emphasized the importance of the multimodal communicative competence as "now more than ever we, our

students and ourselves, need the highest level of understanding of the semiotic workings and affordances of language, as well as of other modes, in order to enact and facilitate powerful personal expression" (2006, p.72).

Accordingly, Yi (2014) considered that it is essential that English language learners develop a multimodal communicative competence to be better engaged in multiple literacies for academic, personal, and social purposes. Multimodal literacy practices are significant to the learner's academic development, as Yi (2014) claimed "multimodal literacy practices can be powerful to engage adolescent ELLs, who were sometimes treated as deficit or delinquent in a traditional sense, in doing good in school" (p.162). Consequently, multimodal literacy practices can actively engage students with the process of learning a foreign language.

2.8 Digital Literacies

Even though young learners are considered "digital natives" because they have been exposed to technology almost all their life, this does not necessarily mean they know how to handle technology effectively for academic purposes. Therefore, it is essential to teach them to be critical and make wise decisions when using all those new emerging modes of communication and engage them in learning. People need to develop new abilities and skills regarding all these new digital practices. Jones and Hafner (2012) defined digital literacies as "the practices of communicating, relating, thinking and 'being' associated with digital media" (p.13). Such digital literacies imply mastering technical aspects of digital tools but, more importantly, knowing how to use those tools to do something in the social world involving managing our social relationships and our social identities in all kinds of situations (Jones & Hafner, 2012). According to Hafner (2019), digital literacies can be incorporated into the English language curriculum as follows:

- Engaging students in structured participation in online affinity spaces

- Strategically embedding digital literacies within the curriculum
- Engaging students in digital multimodal composing projects
- Engaging students in telecollaboration or virtual exchange projects

2.9 Digital Multimodal Composing

The increasing use of digital technology has triggered the development of multiliteracies and consequently facilitated new kinds of multimodal forms of representations (Hafner, 2020). Incorporating digital multimodal composing practices, such as video production in the EFL classroom, has allowed learners to include multiple modes as they compose these digital texts. Students bring together images, sound, gestures, oral and written language to convey meaning within video production. This approach engages learners with multimodal forms of communication in digital media in order to maximize their creativity and explore multiple resources of meaning-making. Hafner (2020) claimed that DMC in English Language Teaching means "going beyond traditional writing forms to include other modes made available by digital media" (p.136). Moreover, Jiang (2017) defined DMC as "a textual practice that involves the use of digital tools to produce texts by combining multiple semiotic modes that include, but are not limited to, image, word, and soundtrack" (p.1). Therefore, digital multimodal composing involves learners in producing videos, podcasts, posters, infographics, comics, storytelling, brochures, comic strips in combination with digital technology. Hafner (2020) affirmed that DMC does not mean abandoning language teaching per se but engaging learners with other modes and guiding them to strategically combine those modes to construct meaning.

Outside the classroom, teenage learners spend most of their time making meaning through different texts in their social life, especially social media. English language teachers can take advantage of it and move these multimodal practices to their classrooms. By doing so, learners can engage themselves in their learning process and foster their competence with multimodal and digital literacies. Therefore, digital video composing can be a potent new literacy learning tool that increases student engagement and achievement.

Miller (2010) argued that Digital video composing in the classroom:

> helps teachers move away from print-only literacy; requires orchestration of representational modes—that is, requires multimodal design; engages students in a real-world literacy practice; makes intuitive sense to learners as a social practice that garners attention and viewers; connects to high-status youth media culture; and prompts deep attention to content and to communicating it for a familiar audience (Miller, 2010, p.5).

Digital multimodal composing contributes to developing students 'multiliteracies and fostering the digital skills necessary to succeed in 21st-century education. Additionally, Hafner (2020) pointed out that DMC positively impacts students' engagement with the English language learning tasks and, consequently, better learning.

2.10 Teaching teenagers

Teenage is one of the most critical stages in someone's life because of all of the changes adolescents experience either biologically, cognitively, physically, or psychologically. As Setiyadi (2020) pointed out, teenagers' age range is from twelve to eighteen years old, so they are in an age of transition between childhood and adulthood. Teaching English to teenagers can be viewed as challenging and demanding, although their potential is greater than the

young children. Teenagers might lack motivation and interest in learning, and it takes a lot of time for the teacher to gain their trust and respect. English lessons, as well as other subjects, are not a priority for teenagers. Therefore, teachers have to implement original activities that capture and hold students' attention but also develop their language skills. Different from children, teenagers can have a better understanding of abstract concepts and handle some grammatical rules. Setiyadi (2020) claimed that "teenagers have more increasing capacities for abstraction as a result of intellectual maturation" (p.153). In using a method that introduces grammar explanations, teachers must be aware that grammar mastery is not the ultimate goal of learning the language. In this regard, language competence should be the goal of learning the target language, and the grammar mastery should be considered the basis (Setiyadi, 2020). There are plenty of useful methods that can be integrated into the English language classroom to improve students' language skills and overall understanding of the target language.

As Setiyadi (2020) suggested, the Grammar Translation Method, the Natural Approach, and the Communicative Approach can work optimally for adults and teenagers because both adults and teenagers may have superior cognitive abilities, and these approaches can provide them with enough opportunities to use the target language by introducing imaginative situations or more abstract concepts that would be difficult to assimilate for young learners. Furthermore, social interactions are fundamental for normal development during the teenage years; on the contrary, isolation can cause dysfunctional behaviors during adulthood. Thus, teachers need to foster a positive language-learning environment in the classroom that contributes to developing these social interactions. It is also crucial to implement activities that promote group work and collaborative learning and foster a learner-

centered approach, in which the teacher assumes the role of facilitator instead of lecturer (Gentile & Orué, 2012). Such strategies promote student engagement and language learning.

2.11 Learning styles in teenagers

Every individual decides the best way to learn a foreign language depending on their cognitive abilities and socioemotional skills. The way adolescent learners acquire and process new information is different in every individual, considering different variables such as type of school, place of living, education level of parents, and socioeconomic status (Kinjari and Gopal, 2020). Learning styles are those cognitive, affective, and physiological traits that guide how learners perceive, interact with, and respond to the learning environment (Lesiak, 2015). Skehan (1991) defined learning styles as "a general predisposition, voluntary or not, toward processing information in a particular way" (p. 288). Several people may find that they have one dominant learning style, while others utilize different learning styles according to different situations. No matter what the condition is, individuals must recognize their learning styles and take advantage of them to better understand their learning process and consequently improve their learning strategies (Kinjari & Gopal, 2020). It can be beneficial for teachers to identify adolescent students' learning style preferences to design effective teaching strategies. According to some researchers (Reid, 1987 and Brown, 2006), field Independent/dependent style, left- and right-brain styles, ambiguity tolerance, and visual/auditory/kinesthetic styles are of great relevance to English language teaching.

The first type of learning style is divided into field-independent and field-dependent style. The former involves the ability to perceive a particular item avoiding any other distracting item, for instance, reading a book in a noisy train station. Brown (2006) argued that a field-independent style in the English language classroom involves an in-depth analysis of details and mastering

exercises such as drills and other focused activities. The individuals who possess this type of learning style show an outstanding performance in deductive lessons. On the other hand, students with a field-dependent style tend to be more successful with inductive lessons. Left-and Right-Brain dominance is thought to be equally important in foreign language learning. According to Brown (2006), students with left-brain dominance prefer a deductive style of teaching. They may be considered as analytic readers who rely on language in thinking and remembering. These students happily respond to verbal instruction and prefer talking and writing. However, learners with right-brain dominance find an inductive style of teaching more effective. They are intuitive learners who prefer drawing and manipulating objects. These students rely on images and respond to illustrated instructions.

A third type of learning style mentioned by Brown (2006) is Ambiguity Tolerance, which is the degree of tolerance people have to manage different situations and ideas opposite to their own beliefs. When learning a foreign language, people may often make mistakes in grammar or pronunciation. In this case, it depends on the person tolerates or not such situations. Brown (2006) suggested that "successful language learning necessitates tolerance of such ambiguities, at least for interim periods or stages, during which time ambiguous items are given a chance to become resolved" (p.127).

The Visual-Auditory-Kinesthetic learning style covers three techniques of sensory learning: visual (sight), auditory (sound), and kinesthetic (touch or motion). Kastner and Stangl (2011) argued that visual learners are more appealing to reading and writing tasks. They prefer processing information with charts, drawings, or other visual modes (Brown, 2006). Auditory learners process information more effectively when they listen to lectures, have a conversation (Kastner & Stangl, 2011). They learn through repeating the material out loud rather than writing

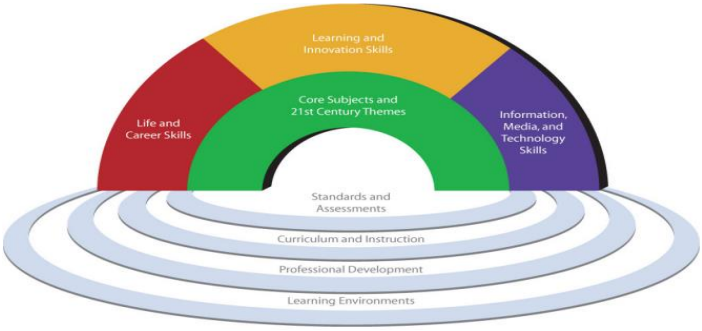
it down (Lesiak, 2015). Kinesthetic learners prefer experiential learning that involves body movement (Brown, 2006). They like to play roles, use gestures during the speech, and experiment with hands-on training (Kastner & Stangl, 2011).

2.12 21st Century skills

Since life in the 21st century has become more globalized and multicultural due to the increased use of new digital technologies, learners need to develop new skills to be successful not only in education but also in their personal and professional lives. Students need to know how to apply all these skills effectively; therefore, there is a great demand to adapt the 21st-century skills to the curriculum. Different authors have grouped these skills into several categories. For instance, Binkley et al. (2011) grouped the 21st Century skills into four broad categories: 1) ways of thinking, 2) ways of working, 3) tools for working, and 4) living in the world. Within these categories, they identified ten skills.

On the other hand, Partnership for 21st Century Skills (P21) provided a more comprehensible framework for learning in the 21st century. The framework is divided into four categories: 1) key subjects and 21st themes, 2) life and career skills, learning and innovation skills and information media, and technology. The framework is presented in Figure 3 below.

Figure 3. Framework for 21st Century



Note: taken from The Partnership for 21st Century Skills (p.1, 2009).

As the framework shows, students have to master some core subjects (English, reading or language, world languages, arts, mathematics, economics, science, geography, history government, and civics) and 21st-century interdisciplinary themes (global awareness, financial, economic, business and entrepreneurial literacy, civic literacy, health literacy, and environmental literacy). Moreover, learning and innovation skills include creativity and innovation, critical thinking and problem solving, communication and collaboration.

According to P21 (2009), in the 21st-century, students are expected to think creatively, work cooperatively, implement innovations, reason effectively, use systems thinking, make judgments and decisions, solve problems, communicate clearly, and collaborate with others. Since we live in a world mediated by digital technology, learners are expected to access and evaluate information, use and manage information, analyze media, create media products, and apply technology effectively. In order to meet these objectives, citizens, workers, and students need to develop information literacy, media literacy, and ICT literacy. The last category is life and career skills, divided into flexibility and adaptability, initiative and self-direction, social and cross-cultural skills, productivity and accountability, and leadership and responsibility. Having acquired these skills, learners are more able to adapt to change, be flexible, manage goals and time, work independently, be self-directed learners, interact effectively with others, work efficiently in diverse teams, manage projects, produce results, guide and lead others, and be responsible to others. Digital multimodal composing, mainly the production of videos, positively contributes to the enhancement of these skills.

Education needs to undergo a radical transformation to enable students to acquire and develop the skills they need to be successful in work and life. However, every school in every country requires to adapt this model according to their context.

2.13 Conclusion

The recent chapter summarized the most important concepts about multimodality, semiotic resources, multiliteracies, and digital literacies concerning digital multimodal composing. Furthermore, Kress and van Leeuwen's grammar of visual design framework was described to understand better the analysis carried out to meet the objectives of the current research. The following chapter presents the methodology in general for this study.

Chapter III: Methodology

3.0 Introduction

This chapter describes the methodology used to gather data for this research. It provides the reader with an explanation of the research design, research context, participants, data collection, and procedures involved in this study. First, the research design and context are described. Then, the participants of the study are presented. After that, the description of the instruments used to collect data as well as the analysis procedures are addressed. Finally, the last section provides a brief conclusion for this chapter.

3.1 Research design

This study adopted an action research approach as its methodological framework, regarded as a qualitative approach to inquiry. Creswell (2008) points out that the focus of qualitative research is on participants' perceptions and experiences; therefore, one of the objectives of this study is to analyze students' perceptions toward the creation of multimodal content, videos. Creswell (2008) claims that "qualitative research is a means for exploring and understanding the meaning individuals or groups ascribe to a social or human problem" (p.4). Hence, a qualitative approach entails an interpretative process that intends to understand specific phenomena regarding the meanings people bring to them.

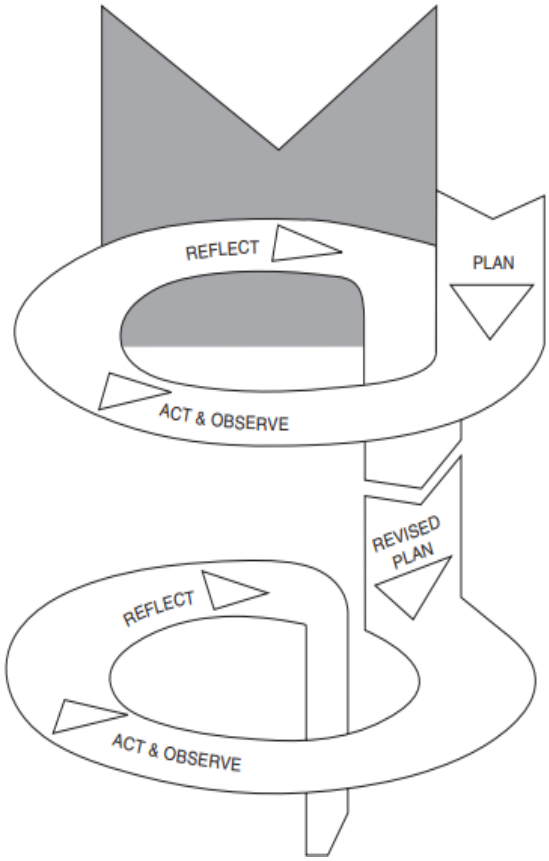
On the other hand, Richards and Farrell (2005) pointed out that action research refers to "teacher-conducted classroom research that seeks to clarify and resolve practical teaching issues and problems" (p.171). Tripp (2005) mentioned that educational action research promotes teachers' development as researchers to enhance their teaching practice and, consequently, their students' learning. Reason and Bradbury define action research as follows.

> Action research is a family of practices of living inquiry that aims, in a great variety of ways, to link practice and ideas in the service of human flourishing. It is not so much a methodology as an orientation to inquiry that seeks to create participative communities of inquiry in which qualities of engagement, curiosity and question posing are brought to bear on significant practical issues (p.1, 2008).

Hence, the focus of this study was to initiate changes in a participative community to improve education practice and make students aware of their role when learning a foreign language. In this sense, learners start getting involved in their learning process by producing videos that make them perceive the English class as something out of the ordinary. Students practice what they have learned in class while developing digital literacy when they interact with different technological tools.

Teachers can use action research as a tool of improvement by identifying issues or problems they face, designing meaningful interventions to solve them, gathering, analyzing, and reflecting on the data to know the impact of these changes (Pardede, 2018). The action research is conducted during regular classroom teaching, so it does not interfere with the class learning schedule, as is the case of this study. According to Koshy et al. (2011), action research is "a method used for improving practice. It involves action, evaluation, and critical reflection and –based on the evidence gathered –changes in practice are then implemented" (p. 2). Action research follows a spiral cycle consisting of four major stages: planning, acting, observing, and reflecting (Kemmis & Robin McTaggart, 2005).

Figure 4. The Action Research Cycle



Note: taken from Kemmis and McTaggart (p.278, 2005).

According to Kemmis and McTaggart (2005), the process might not be as neat as this spiral cycle since the stages can overlap, and initial plans become archaic in the light of learning from experience. The key to success is that participants have a strong sense of development in their practice by achieving a change that marks a meaningful difference in the initial problem.

3.1.1 Planning the research

Based on the foundations of qualitative action research, the study was guided by the following stages.

Table 4: The stages of the study

Stage 1. Planning	Since an action research project aims to improve a specific practice, the first step consisted in identifying and limiting the problem. In this study, there was a need to involve students and engage them in their learning process. Therefore, the researcher gathered information about this issue and reviewed related literature in order to develop a research plan.
Stage 2. Acting	Once the problem was identified and an action plan was designed, the researcher implemented multimodal tasks (videos) to involve students in creating multimodal content and at the same time got them engaged in their learning process.
Stage 3. Observing	At the end of the implementation, the data were collected, organized and then analyzed.
Stage 4. Evaluating and reflecting	Based on the data analysis, the researcher determined whether the intervention resulted in the intended change of improvement and identified the problems that arose during the action implementation.

Note: own creation

Besides the actual teaching in the current situation, Covid-19 pandemic, the scope of English language teaching can go beyond the traditional focus on reading and writing to produce multimodal texts. Therefore, the current study also used a Multimodal Discourse Analysis (MDA) to analyze the students' digital video projects. This methodological approach provided tools for analyzing and describing the semiotic resources learners used to construct meaning through videos to determine their role in creating digital videos. As O'Halloran (2011) mentioned, MDA is "concerned with theory and analysis of semiotic resources and the semantic expansions which occur as semiotic choices combine in multimodal phenomena" (P.2).

In terms of analysis, MDA also follows three dimensions of meaning-making based directly on Halliday's metafunctions. The ideational metafunction represents how social experience is construed, the interpersonal metafunction enacts social relations, and the textual metafunction pertains to how information is organized to achieve textual cohesion and coherence.

Therefore, MDA explored how the different modes interact with the three metafunctions to create meaning. The videos' analysis was complemented with the information gathered from a focus group and a questionnaire whose primary purpose was to elicit the students' perceptions toward the production of videos and explore the impact of digital multimodal composing on the students' engagement in their learning process.

3.2 Research context

This research project was conducted in a public high school, "Centro de Estudios Tecnológicos, Industriales y de Servicios No. 17 (CETIS No.17)", located in San Martín Texmelucan, Puebla. The school belongs to the "Dirección General de Educación Tecnólogica Industrial (DGETI)" which is a unit attached to the "Subsecretaria de Educación Media Superior." The school offers technical education, and its primary purpose is to incorporate the students into the productive sector through internships and professional service in different companies. The institution is oriented to the learning and development of technological and humanistic knowledge. It aims to train technical high school graduates who develop, strengthen, and preserve a technological culture and an industrial infrastructure contributing to the country's economy and social needs. The school has around 1674 students divided into two shifts (morning and afternoon) and five technical careers (programming, industrial mechanics, computer software and hardware maintenance, logistics, and clinical laboratory). According to the information provided by the academic coordinator of the institution, there are eight English language teachers who have classes with a

minimum of 25 and a maximum of 54 students per class. Based on the coordination data, 60% of the teachers hold a Bachelor's degree in Modern languages and English language teaching.

The English program is divided into five courses that last four months and are designed according to the Common European Framework of Reference. Regarding the schedule, students from the first to the fourth semester have English classes three hours per week compared to students from the fifth semester who have a five-hour class per week. It is important to highlight the fact that students do not have English classes in their last semester. In addition, a school policy prevents them from buying textbooks; to solve this situation, the instructors have to look for alternative resources that help learners develop their language skills. Since English language teachers do not follow a specific methodology, they can design any strategy they consider pertinent or adapt the most effective method according to the students' needs and contexts. Thus, working with digital video projects can be a great tool that promotes student-centered learning and allows them to play a more active role while engaged in their learning process (Samaranayake, 2020). The creation of multimodal content allows learners to participate actively in their learning process as well as help them develop language skills and become digitally literate.

3.3 The Study participants

The participants of this study were 20 students of fifth grade from a public high school, "Centro de Estudios Tecnológicos Industrial y de Servicios N° 17" in San Martin Texmelucan, Puebla. There were 12 female and 8 male participants whose ages ranged from seventeen to eighteen years old. All the participants were enrolled in the English as a foreign language course level V, with an A2 English level within the Common European Framework. The participants belong to a low-middle socioeconomic status, and their parents' main economic activities are part of the

commercial and agriculture sectors. The majority of students come from rural communities even though the school is located in an urban area. In order to know the feasibility of the study, at the beginning of the course, the participants answered a questionnaire (google forms) in which they mentioned they had access to technology and other resources that allowed them to create digital multimodal content.

Before starting the data collection process, consent forms were sent to the students and their parents since some participants were minors. The students and parents signed the consent forms and gave permission to work with their videos. Those consent forms provided information about the research project to be carried out, the privacy of the participants, the no reward and voluntary participation in the research, the participant's right to avoid answering questions, and their possibility to withdraw from the process at any moment. In addition, those forms stated and guaranteed the use of the information provided only for research purposes as well as the protection and anonymity of the participants' identities

3.4 Justification of the use of multimodality in the study

Since the main objectives of this study were to analyze the semiotic resources, students used to construct meaning through videos and their perceptions toward this multimodal task in order to explore the impact of videos on the students' learning process engagement, it was necessary to do an implementation incorporating technology and multimodality in the 5th-semester course planning. This implementation started in September 2020 and concluded in February 2021. The course was divided into five units whose goals were to create different multimodal tasks, mainly videos. The purpose of these videos was to expose learners to a multimodal environment that contributed to developing and improving language skills. In this way, students could get involved in their learning process by having direct and active participation. Based on the students' context

and needs, adding multimodal elements in the English class to attract the students' attention was considered.

3.5 Data collection and data analysis

This study was conducted during the fifth semester of high school, from September 2020 until February 2021. Data was collected for 17 weeks, during which the researcher met for 80 class sessions with the students. This study considered the following techniques for data gathering: a focus group, a questionnaire, and a multimodal discourse analysis of the students' artifacts.

At the end of the course, an online focus group was used to gain an in-depth understanding of the students' perceptions toward the creation of videos in the English class. Peacock et al. (2009) defined an online focus group as "a selected group of individuals who have volunteered to participate in a moderated, structured, online discussion in order to explore a particular topic for the purpose of research" (p.119). The objective of the online focus group was to gather information about public high school students' perceptions, opinions, and experiences toward the creation of multimodal content in their English class. A group of thirteen students was interviewed following a specific set of questions. In the first part of the interview, students expressed their thoughts and feelings about the emergency remote teaching due to the COVID-19 pandemic, as well as their experiences with online learning. The second part dealt with questions directly related to their perceptions and opinions toward the implementation of videos and how the creation of multimodal content impacted their learning engagement. The last part of the interview consisted of questions regarding the students' opinions about the course and the adjustments that can be made to improve it. The participants connected to the meeting via ZOOM video-conferencing, and the discussion lasted 50 minutes. In order to analyze the students' responses, the interview was audio-recorded

and then transcribed. For data analysis, each participant received a code (e.g., S1, S2, S3) to identify and classify the information provided in a specific category.

A questionnaire was administered at the end of the course in order to obtain more precise information about the learners' perceptions and opinions toward the implementation. Questionnaires are defined as any written instruments that provide participants with a series of questions or statements that they have to respond to, either selecting from existing possibilities or writing out their answers via paper or online (Brown, 2001). The questionnaire was designed in Google forms, and it was conducted in Spanish to avoid any misunderstanding, and in that way, participants completed it without any difficulty. The questionnaire was sent to the participants via WhatsApp.

The third data collection instrument involves the videos students created during the course. Yin (2003) also addressed the students' work samples as artifacts, which are "a technological device, a tool or instrument, a work of art, or some other physical evidence" (p. 113). These artifacts were collected during the course development, and then a multimodal discourse analysis was conducted. Before starting the data collection process, consent forms were designed and presented to the students. Their anonymity was also discussed at this point, so they felt confident to participate in this study.

3.6 Planning and implementing the Multimodal tasks

Due to the nature of the study, three multimodal tasks were carried out in order to develop students' language skills and got them involved in their learning process as they had the opportunity to work in the creation of videos mediated by different technological tools and multiple resources to create meaning beyond language. At the beginning of the course, the researcher provided the participants

with different tools that helped them perform the tasks more effectively. The production of multimodal content in the English class was a gradual process that required time and effort since this was the first-time making videos for some learners. In order to receive feedback, the participants shared two videos only with the teacher because they did not feel confident enough to make their videos public at this point. However, the last video was not only shared with the teacher but also with the whole class; at this stage of the process, students had already gained confidence and experience producing multimodal content. The three multimodal tasks are described as follows.

Table 5: The multimodal tasks

Task	**Purpose**	**Procedure**	
Video 1: Digital Story telling	To tell a personal story.	Students created digital storytelling about their past.	Four stages: • Preproduction • Production • Postproduction • Distribution.
Video 2: Digital role-play	To enhance learners' communicative skills and exposed them to an authentic situation where they had to use the language with a purpose.	Students performed a role-play called "A visit to the doctor." They made a video in which they had to act out a medical consultation. Three steps: preparation, presentation, and feedback.	
Video 3: Elections commercial	To make students reflect on their current country's socio-political issues and how they could improve or solve them if they were the president of their country.	Students made a video commercial to present their proposals of what they would do if they were elected president of their country.	

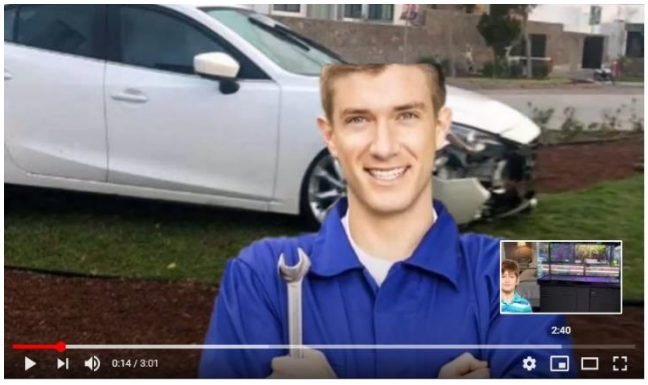
Note: own creation

Video 1: Digital Storytelling

In the first multimodal task, students created digital storytelling about their past. This task aimed to involve students in creating content, rather than just being passive participants; in this sense, they could get engaged in their learning process. Students worked individually since the purpose of the task was to tell a personal story. In order to carry out this project, learners went through a set of processes consisting of four phases: preproduction, production, postproduction, and distribution. Students developed a general idea of their story and how it would progress during the first phase. Then, learners started writing a transcript and creating a storyboard; in this step, students began to make decisions about images, video, and sound and how they had to organize them to illustrate their ideas and emotions. In the second phase of the process, learners gathered and created images, audio, and video, considering that everything they chose would impact their digital storytelling. Students used this time to record themselves reading their scripts and working on pronunciation and fluency aspects.

The postproduction phase involved assembling all resources by blending images, creating transitions between video clips, incorporating music or sound effects. In the last stage, learners uploaded their digital storytelling to their YouTube channel and shared the link with the teacher to receive feedback (See an example in Figure 5).

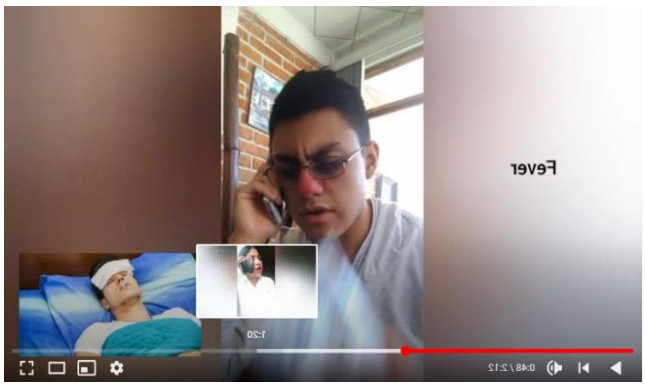
Figure 5. Digital storytelling



Video 2: Digital role play

In this task, students worked in pairs to perform a role-play called "A visit to the doctor." Students made a video in which they had to act out a medical consultation. The objectives of this task were to enhance learners' communicative skills and exposed them to an authentic situation where they had to use the language with a purpose, in this case, told the doctor about their health problems and described their symptoms, so s/he could prescribe medicine/treatment and gave them some advice or suggestions to feel better. In order to meet the objectives, the digital role-play followed three steps: preparation, presentation, and feedback. First, the teacher gave students the guidelines and assigned a partner. It is important to mention that questions and doubts were resolved through a zoom meeting, and during the project, students received continuous feedback from the course instructor. Then, learners carefully planned and wrote the scripts and chose the digital and non-digital elements to create the most suitable setting. They also prepared themselves for acting and got immersed in a real situation. After that, they edited their video and assembled all the digital elements. Finally, they uploaded their video to their YouTube channel and received feedback from the teacher. This multimodal task allowed students to interact with their peers creatively, fostering a sense of community and shared purpose in the learning process. (See an example in Figure 6).

Figure 6. Digital role-play



Video 3: Elections commercial

The last multimodal task aimed to make students reflect on their current government and their country's socio-political issues and how they could improve or solve them if they were the president of their country. Individually students made a video commercial to present their proposals of what they would do if they were elected president of their country. In these proposals, learners had to implement the form and function of the second conditional and specific vocabulary.

The project enhanced the learners' productive skills and promoted information and communication technology (ICT). The project was carried out in four stages: preproduction, production, postproduction, and distribution. First, learners needed to reflect on their government's current issues, and then some proposals were formulated to solve those issues or improve certain socio-political aspects. After that, they started writing a script using a wide range of vocabulary and grammatical functions. Then, storyboarding played an essential role because students began to sketch each section of their video. Once they had the script and storyboard ready, the filming process began. Finally, students put everything together and uploaded their videos to their YouTube channel and shared the link in the course platform so that the whole class could watch them and vote for the best video commercial considering the evaluation instrument, a rubric. Finally, learners received feedback from the teacher and other classmates. (See an example in Figure 7).

Figure 7. Elections commercial

3.7 Conclusion

This chapter presented the information related to the research design, participants, the research context, and the description of the multimodal tasks. Having provided the reader with an overview of the methodological procedures employed to collect and analyze data, the work will now turn to the data analysis section. The three videos previously described will be analyzed according to Kress and Van Leeuwen's framework.

Chapter IV: Data Analysis

4.0 Introduction

This chapter presents the analysis of the videos learners produced in order to identify the semiotic resources learners used to convey meaning. In addition, this chapter seeks to analyze the learners' perceptions toward digital multimodal composing in the EFL classroom. Firstly, an overview of the study is provided. Secondly, a summary of the data collection and data analysis process is presented. Then, a general description of the multimodal texts is given. After that, the analysis and the results are presented. Finally, the conclusions are described.

4.1 Overview of the study

The current study explores EFL students' compositions of digital multimodal texts and analyzes the semiotic resources learners employed to construct meaning in the digital multimodal texts. Additionally, the research seeks to examine students' perceptions of digital multimodal composing to explore the impact of video production on their learning process engagement. This study adopted qualitative action research as its methodological framework consisting of four major stages: planning, acting, observing, and reflecting. During the acting stage, the researcher implemented multimodal tasks (videos) to involve students in creating digital multimodal content and their learning process.

4.1.1 Data collection summary

At the end of the implementation, an online focus group discussion and a questionnaire were conducted to gather information about students' perceptions, opinions, and experiences toward creating multimodal content in their English class and how they improved their learning process engagement. The students' perceptions and opinions contributed to the last stage of the action

research project, evaluation and reflection, to determine whether the intervention resulted in the intended change of improvement and how the problems that arose during the action implementation can be addressed in future research. The second part of the research entails a multimodal analysis of the semiotic resources students employed in producing the three videos and how these semiotic resources were orchestrated to convey meaning.

4.1.2 Data analysis summary

The coding approach was implemented to analyze the participants' comments derived from the focus group discussion and the questionnaire responses to explore the students' perceptions and opinions toward Digital Multimodal Composing in the L2 classroom. Linneberg and Korsgaard (2019) defined coding "as the process of turning raw qualitative data into a communicative and trustworthy 'story'" (p.3). Thus, the coding process consists of labeling and organizing qualitative data to identify different categories or relationships between them. Once the students' perceptions of the implementation were analyzed, the implementation was evaluated, and consequently, the last stage of the action research was concluded. Finally, the second part of the study adopted a multimodal analysis to identify and examine the semiotic resources used in the videos. The framework used to examine how various semiotic resources worked together to construe meanings and answered the research questions posed at the beginning of this study was based on Kress and Van Leeuwen's grammar of visual design (2006). Thus, the screenshots of the video were analyzed through three metafunctions of visual text, namely the representational, interactive, and compositional.

4.2 Results of the students' perceptions and opinions

As previously mentioned, a questionnaire and a focus group discussion served as the primary research data source to explore the students' perceptions and opinions toward Digital Multimodal

Composing in the L2 classroom. After collecting the questionnaire responses and transcribing the focus group discussion, the coding approach was used to analyze the information. Following this approach, the data obtained from the focus group discussion was compared with the questionnaires' responses. Twenty learners took part in this study. From the total, thirteen learners participated in both instruments, the focus group discussion and the questionnaire. Three students only participated in the focus group discussion, and four only answered the questionnaire.

4.2.1 Focus Group Analysis

Thirteen students participated in the focus group discussion conducted via ZOOM, eight women and five men. The focus group discussion was carried out in the learners' mother tongue (Spanish) to avoid misunderstandings and obtain more reliable results. Different themes were identified and organized into five categories derived from the coding process: DMC: implications for learning, the usefulness of DMC, feelings & attitudes toward DMC, English language improvement, and challenges.

Table 6: Implications of DMC for learning.

Participants	DMC: Implications for learning
S1	• It is not only a practice for pronunciation but also for the grammatical structure. • It was a triple practice (pronunciation, sentence structure and video editing).
S2	• I learned to use different technological tools. • I learned to be a little more responsible and organize my time. • I learned to distinguish the structures of each verb tense, for example the past simple.
S4	• We try to improve; we try to find out our pronunciation and know how words are expressed • I learned more and I have been learning a lot.
S7	• It helped me to practice the pronunciation much more. • I was able to practice the dialogue and look up the words on my own, making it more didactic for me.

S8	• You learn various things for example, to organize things to do them well and to have the courage to record yourself. • It is important to have a script to not have spelling mistakes later.

The analysis begins with table 6, which shows the learners' comments regarding the implications of digital multimodal composing for learning. As it can be observed, the underlined keywords show that participants could practice pronunciation and grammar while learning how to use different technological tools for the video edition. Moreover, participants pointed out that they learned to be more responsible and organize their time better. The table synthesizes the most common implications of DMC for students' learning.

Table 7: The usefulness of DMC

Participants	Usefulness of DMC
S1	• I consider the video "A visit to the doctor" very useful because we learned how to describe health problems and it can be relevant in our daily life if one day we go abroad and face a similar situation.
S2	• I consider it is very useful to learn new vocabulary and apply it when making the videos.
S3	• I think that making videos helps you a lot to practice the language, for example, in the storytelling I tried to make my dialogue sound a bit more native, substituting words and changing some of them for contractions. • This kind of tasks will help me in the future, when I deal with situations in which I have to speak more formal English.

According to the results showed in table 7, learners considered DMC as a valuable tool to learn and practice relevant vocabulary and a great option to practice the language using different scenarios. The learners also mentioned that these kinds of tasks could help them face similar situations in the future.

Table 8: Feelings & attitudes toward DMC

Participants	Feelings & attitudes toward DMC
S1	• I really like the creation of content. • I like to edit videos
S4	• Working on the production of videos was a very unique experience, a great experience. • I liked working on the video creation.
S5	• I feel satisfied with the work I did. • It was a good experience
S6	• I prefer the process of creating videos, especially record them and edit them. • It is easier to record yourself in front of a camera, because you have more confidence. You know that if you make a mistake, you can record it again or do it again.
S7	• It's easier to stand in front of a camera and be able to repeat it, instead of being in face-to-face classes and only have one chance to do it. • I liked making videos so much.

Table 8 shows the students' feelings and attitudes toward DMC. The keywords "like," "unique experience," "great experience," "feel satisfied," "good experience," and "easier" suggest positive feelings and attitudes toward the production of videos in the L2 classroom. The results show that participants liked working on the video creation and considered it as a positive experience. Learners' comments suggest that creating videos made easier the process of speaking in English. The results also show that learners enjoyed editing the videos as well.

Table 9: English Language improvement

Participants	English language improvement
S3	• My listening improved a lot. I had the opportunity to repeat the audios as many times as I wanted. So, when I got stuck in a sentence or, in a word, I repeated it again and practiced it. • My writing also improved as I had the time and dedication to pay more attention to the language structure.
S8	• I could improve my fluency and my pronunciation.

S10	• I improved both pronunciation and writing. • At the moment of recording the videos, you repeat and repeat and if you make a mistake, you repeat it until you get it well.
S12	• Making videos helped me practice my vocabulary and express myself better in front of a camera.
S13	• The videos helped me a lot to practice the vocabulary. • In face-to-face classes I would not have had the opportunity to repeat it again and again and it would have been more difficult.

Table 9 illustrates learners' perceptions about how DMC helped them improve their English language. In their comments, learners suggest that their listening and writing abilities improved. They also mentioned that they had the opportunity to practice vocabulary and enhance their pronunciation.

Table 10: Challenges.

Participants	Challenges
S2	• One of the difficulties I faced was learning to organize my time in order to be able to record, edit and upload the video. • When you record, you have to organize the sequence of the different scenes of the videos and it was a great challenge for me because I didn't use to organize anything.
S4	• The only thing that was difficult for me was the pronunciation.
S5	• Nervousness • Creativity, because not all of us have the facility to be creative in our projects. • The process was somewhat tedious and the pronunciation was a great challenge.
S9	• It was a bit difficult for me since I am not used to speaking in front of the camera and that makes me so nervous.
S13	• At the beginning, it was a bit difficult for me to edit the video, and also, I was not very creative.

According to table 10, learners faced some challenges when working in the production of videos. Students reported that one of the difficulties they encountered was pronunciation. They also

mentioned that they got nervous when recording the video and did not organize their time effectively. The results show that lack of creativity was another challenge for learners as well as the editing process.

4.2.2 Questionnaire Responses Analysis

The questionnaire was answered by sixteen participants, ten women, and six men. This instrument was made with Google forms. It consisted of eleven open-ended questions and four multiple-choice questions; in total, there were fifteen questions written in Spanish. Since most of the questions were open-ended, coding was the best option to analyze them. The analysis resulted in six categories: Enhancement of creativity, the impact of DMC on learning, language skills improvement, challenges, engagement in the language learning process, and experiences in DMC.

Table 11: Enhancement of creativity

Participants	Enhancement of creativity
S2	• It is a way of expressing my creativity. • I developed my creativity by expressing my ideas.
S3	• I had to find ways to capture the interest of my audience. • A good way to stimulate my imagination.
S4	• I tried to be creative at the moment of making videos.
S5	• I could develop my creativity when making videos.
S9	• I really like working with the production of videos because it helps me enhance my creativity.
S10	• It helps to be more creative. • I developed my creativity when I tried to find out how to make my videos more attractive. • I learned to be more creative.
S11	• Since I wasn't very creative, I had to find a way that my videos didn't look boring.
S13	• It made me discover new applications and new things that contribute to the design and creativity of my videos. At the end, it was easier for me to edit the videos.
S14	• Trying to do something different from your classmates is what motivates you to be creative.

S15	• At the same time that I was trying to learn English, I improved my creativity. • Developing videos gave me the opportunity to show my creativity when writing my script and editing the video.
S16	• I enhanced my creativity in the editing process.
S17	• What I liked the most was being creative because it was entertaining and had the opportunity to enhance my creativity.
S18	• I think that one of the aspects that I developed the most was precisely my creativity.
S19	• During video creation I always try to be creative and put a lot of effort.
S20	• Video production and editing really enhanced my creativity.

Table 11 shows the answers learners gave to a question about creativity. The results report that fifteen participants considered that one of the benefits of DMC was the enhancement of creativity. Learners pointed out that they tried to find different ways to capture their audience's interest by making their videos entertaining, creative, and attractive.

Table 12: Impact of DMC on learning.

Participants	Impact of DMC on learning
S2	• Thanks to this kind of activities our knowledge remains and it is not easy for us to forget. • I learned to use new technological tools. • I can imagine a conversation and practice my pronunciation.
S3	• It allowed me to review my dialogues as many times as I wanted.
S4	• It is a way to practice the language, the pronunciation and even know the meaning of different words. • You can express yourself more.
S6	• As I had the opportunity to repeat and correct my mistakes at the end everything was clearer for me.
S9	• I learned to be more confident and ask about things I didn't understand. • I learned to edit better, I learned to be more careful with videos as they required to have a more formal concept.
S10	• By constantly repeating, it is easier for me to learn.
S11	• It is a way of practicing what you have learned.
S14	• I can express my ideas and learn about new topics while improving my skills in front of a camera.

	• Writing texts per se becomes tedious, but by making videos we explore new learning methods.
S15	• This is a more dynamic way to practice the language. • I improved a little bit more my knowledge of the language.
S16	• It makes easier the way I learn.
S17	• I learned to use new technological tools and developed my language skills.
S18	• I learned to be more organized, creative and to express my ideas better through the videos.
S19	• Using images, audio, music, special effects in my videos made me understand the topic better and also learning English became in something meaningful.
S20	• I learned to express myself better in a foreign language and to be more confident when speaking in English.

Table 12 illustrates the impact of DMC on learning. Participants reported that producing videos in their English class made them repeat, prepare and practice over and over again until they felt confident enough speaking English in front of a camera. DMC allowed learners to express their ideas better while practicing the language more dynamically. Participants also learned to use new technological tools and be more organized. The underlined keywords show how DMC impacted learning and how students could benefit from producing multimodal content.

Table 13: Language skills improvement.

Participants	Language skills improvement (Mostly pronunciation)
S2	• I improved my pronunciation.
S3	• It was a good opportunity to improve my pronunciation. • I could train my ear to distinguish the pronunciation of similar words. • I could improve my listening.
S4	• I dedicated myself to practice and improve my pronunciation. • In each video I tried to pronounce correctly because I know that it is not the same to write a word in English that pronounce it. • Over time I was able to improve my pronunciation.
S9	• I learned the pronunciation of new words and realized that repeating and practicing the words, they easily remain in your head.
S10	• I could practice my reading and pronunciation.
S13	• It helped me to improve my pronunciation a little bit more.

S14	• I improved both my pronunciation and my reading ability.
S15	• I learned to pronounce better.
S16	• I learned the pronunciation of new words.
S17	• The videos helped me enhance my pronunciation a little bit more.
S18	• Creating videos really helped me improve my listening and pronunciation as well as my writing.
S19	• It was a good opportunity to train my ear. Now I can understand better some audios in English.
S20	• I was able to practice my pronunciation much more in order to be more fluent.

Table 13 shows that learners improved their pronunciation which is a sub-skill of speaking. Most of the participants agreed that the production of videos allowed them to train their ears, practice, and improve their pronunciation. Additionally, few learners reported that their listening, reading, and writing skills were improved as well.

Table 14: Challenges

Participants	Challenges
S2	• Fluency and pronunciation.
S3	• Nervousness • Try to record in a quiet place without noise.
S4	• Pronunciation
S5	• Fluency and nervousness.
S6	• Editing always takes me a lot of time, it is not so complicated but laborious.
S9	• Speaking is something hard for me and it makes feel so nervous.
S10	• Time was a big challenge. I needed more time.
S11	• I was not creative. • I was very nervous.
S13	• Pronunciation was difficult for me. • Lack of creativity.
S14	• Editing
S15	• Pronunciation of difficult words.
S16	• Vocabulary • Nervousness
S17	• One of the great challenges I face was my fear of speaking in English.
S18	• My biggest difficulty was editing.

S19	• Time
S20	• I was nervous when speaking English.

In table 14, the challenges learners faced when working with the production of videos can be observed. The most frequent challenges students had to overcome were their fear of speaking in English, pronunciation, and fluency. The second most mentioned challenge was nervousness when learners had to speak in a foreign language. In third place, lack of creativity and editing the videos were other obstacles students encountered in making videos.

Table 15: Engagement in the language learning process

Participants	Engagement in the language learning process
S2	• It encourages me to continue learning the language. • The edition and video production motivates me.
S3	• I looked for activities that helped me with pronunciation and train my ear to identify the sounds of certain words.
S4	• I think you learn a little bit more about the topics when you do it rather than just see it. • I experienced new things in the videos that maybe I had no idea of how to do before, but the best part is that I had fun doing it.
S5	• Before making videos, I realized how important learning a foreign language is.
S6	• I got interested in learning the pronunciation of words.
S9	• It is a good way to learn English and to be more responsible.
S10	• Making videos was a very useful activity that allowed me to express myself and learn.
S11	• I had to pay attention in class in order to understand the topics and be able to make the videos.
S13	• Before making videos, I didn't understand the pronunciation, it was very difficult even in songs, but after the videos, it became a bit easier for me.
S14	• Since I started working with videos, I got interested in improving my pronunciation and the way of speaking in English.
S17	• After working with the production of videos, I got interested in the language.
S18	• I have learned how useful can be to learn English.
S19	• Now I feel that learning English is no longer impossible as I thought before because creating videos helped me a lot.

S20	• I managed to get more interested in the subject because I was very excited to make videos.

The results provided in table 15 show that learners' engagement in the language learning process increased considerably. The participants mentioned that they noticed a difference before and after making videos. After producing videos, some participants expressed that they got interested in keeping learning and practicing the target language. The comments also show that some participants also realized how important learning a foreign language is.

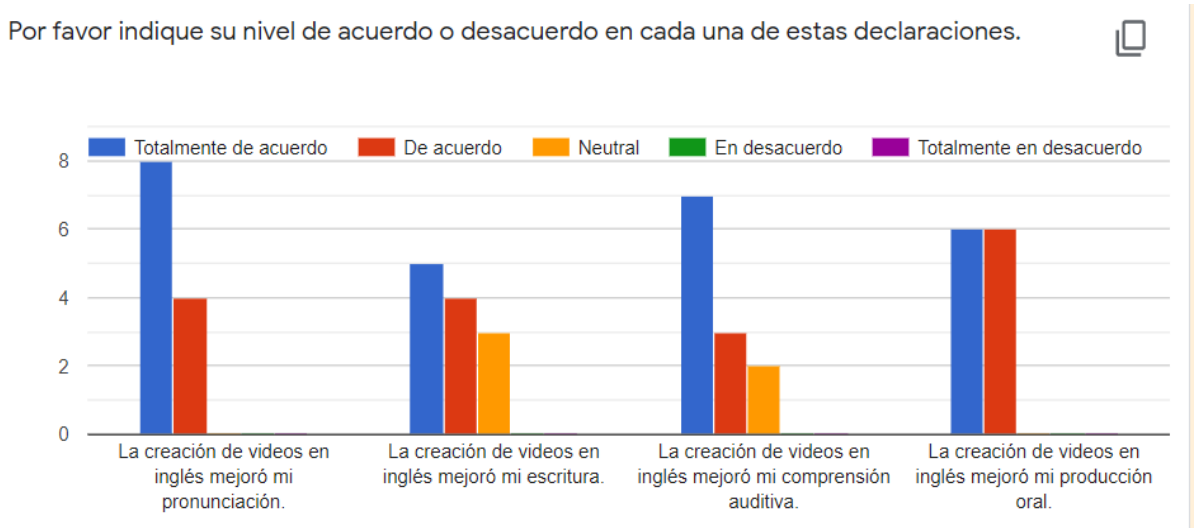
Table 16: Positive experience toward DMC

Participants	Experience in DMC
S2	• It was a good experience because I could use many tools both physical and technological.
S3	• It was a good way to get out of my comfort zone.
S4	• It was a very unique experience. I tried to do my best in making my videos interesting.
S5	• It was very funny to be another person speaking in English.
S6	• It's easier for me to create audiovisual content. • I had fun making videos in English.
S11	• It was fun and it was something I hadn't done before, somewhat it helped me a lot.
S13	• I enjoyed editing the videos. • I liked the process, first I had to write my script and then practice the pronunciation over and over again until I could be fluent. .
S14	• It was a very good experiences, it gives you a twist where everything is repetitive and helps us put into practice everything we have learned, it is highly recommended.
S16	• It was something complicated, but funny. These kinds of activities make the language easier to understand.
S17	• It was a unique experience since I had never done this type of projects and I really loved it.
S18	• I liked working creating this type of content because it helps you learn the language better.
S19	• At the beginning it was complex because I had never done something similar but later, I had fun making the videos.

S20	• I really liked creating videos and I consider it a great experience for learning a foreign language.

The answers presented in table 16 have to do with the participants' experiences producing videos in English. According to the results, all the participants mentioned at least one positive aspect of working on DMC. The adjectives students used to describe their experience were "good," "unique," "easy," "fun," and "great," and the verbs were "enjoy" and "like." As shown in the comments, learners mentioned that their experience making videos allowed them to learn new things related to the target language and the editing process. In the comments, learners also mentioned that it was a complex process initially, but in the end, they had fun making the videos, and they could find a different and new way to practice the language.

Graph 1. Improvement of the language skills



Graph 1 illustrates the learners' level of agreement about language skills improvement. According to the results, eight students strongly agreed, and four of them agreed with the statement, "Making videos in English improved my pronunciation." In the second statement, "Making videos in

English improved my writing," five students totally agreed, four agreed, and three were neutral. Regarding the third statement, "Making videos in English improved my listening," seven participants strongly agreed, three of them agreed, and two were neutral. In the last statement, "Making videos in English improved my speaking," six participants strongly agreed, and six agreed. The results show that most of the participants perceived that the production of videos helped them improve their speaking skills.

4.3 General description of the multimodal tasks

The first video to be analyzed is a digital storytelling in which the participant told a personal anecdote about the day his parents bought him an axolotl. During the narration, the participant described all the things he went through that day. The video lasts 3 minutes and shows a series of pictures that illustrate what he is describing. Unlike the other videos, we can only listen to the student's voice and appreciate the visuals he used in this one. The video can be watched on YouTube at the following link: https://www.youtube.com/watch?v=jE9gOpvP0dQ.

The second data sample to be analyzed is a video called "A visit to the doctor." In the video, two students performed a phone conversation (due to the Covid-19 pandemic, students could not be face to face) in which one of them played the role of a doctor, and the other one took the role of a patient. The video lasts 2 minutes and 12 seconds and can be watched on YouTube at the following link: https://www.youtube.com/watch?v=VxNTgoX92EU.

The third and last video to be analyzed is a commercial titled "Elections." In this video, the participant put forward his proposals of what he would do if he were elected president of his nation and tried to convince the audience to vote for him. The video lasts 2 minutes and can be watched on YouTube at the following link: https://www.youtube.com/watch?v=KlkQlGCenzc.

At the beginning of the course, students were informed they would make three videos to practice what they had learned in their English class. The instructions of the first video were to create a digital story about an event that occurred in the past using pictures, images, drawings, or photographs to illustrate their story and support what they are telling visually. Learners were told not to record themselves since it was, for most of them, a new experience and something completely different from what they had been done in their English classes before. Therefore, some learners were not confident enough to appear in front of a camera speaking in English. Students were asked to record their narration during the digital story production, giving the correct intonation and interpretation to what they were telling. Pronunciation and fluency were two aspects students had to consider when recording their narration.

Unlike the first video, in the second one, students were asked to record a role play in which they had to dress up as a doctor and patient, respectively. However, they could use other kinds of visual support such as images and keywords. Students were also asked to record themselves in the last video, but now they had to portray a presidential candidate talking about his proposal to improve the community. In this video, learners were told they could wear clothing according to their role. In the three videos, the first step was always to draft a script and then polish it based on the teacher's feedback. A requirement for producing the three videos was creativity; learners had to innovate and go beyond conventional.

4.4 Multimodal Discourse Analysis

The framework used to examine how various semiotic resources worked together to construe meanings and answer the research questions posed at the beginning of this study was based on Kress and Van Leeuwen's grammar of visual design (2006). Thus, the videos screenshots were analyzed through three metafunctions of visual text, namely the representational, interactive, and

compositional. The representational metafunction deals with the people, places, and objects within an image. The representational meaning is realized by two processes: the narrative and the conceptual. Kress and Van Leeuwen (2006) argued that narrative patterns are "present unfolding actions and events, processes of change, transitory spatial arrangements," whereas conceptual patterns concern "represent participants in terms of their class, structure or meaning" (p. 59). Kress and Van Leeuwen (2006) referred to two types of participants involved in every semiotic act in the representational metafunction, the represented participants and the interactive participants. The former is the people, places, or things represented in the visual, and the latter are the producers and the viewers of the visual (Kress & Van Leeuwen, 2006). The interactive metafunction is concerned with the interaction between the producer, the represented participants, and the viewer of the image; that is how participants interact with each other. Contact, social distance, attitude, and modality are four fundamental aspects of the interactive metafunction. The compositional metafunction deals with how the representational and interpersonal metafunctions integrate into a meaningful whole through three critical systems: information value, salience, and framing (Kress & Van Leeuwen, 2006).

The first system, information value, refers to the placement of the various elements within an image. The placement of these elements is attached to different zones of the image, such as left, right, top, bottom, center, and margin. The salience system involves factors such as size, tonal contrast, color contrast, sharpness of focus, and placement in the foreground or background of the represented participants in order to capture the viewer's attention. Framing, the third key system in composition, is associated with connecting or disconnecting elements in the image through frame lines. As the three multimodal tasks under examination are mainly composed of visual

images, the construction of the representation, the interactional and the compositional meaning will be explored,

Video 1: Digital storytelling

The content of the first video, "Digital storytelling," is illustrated by edited pictures whose primary purpose is to create a meaningful connection with what the student is telling. The current task shows how the visual and verbal modes interact to construct the integrated meaning. In digital storytelling, the narrator's voice tends to be the primary mode, whereas images and sounds tend to be supportive (Alonso et al., 2013). The narrator's voice provides additional information and specific details such as time, location, background, age, and so on.

Figure 8. Sequence of screenshots

Image 1. Video introduction

Image 2. Family

Image 3. Road map

Image 4. Fish market

Image 5. Fish tanks

The story begins with the introduction of the main character which is the narrator. In the first slide, a picture of a young boy is displayed, and immediately the number 13 appears as a

referent of age within a black background. The black background creates salience as well as the placement of the represented participants (the boy & the number). The contrast between colors allows the viewer to focus on the main character and, at the same time, reinforces the verbal mode. Kress and Van Leeuwen (2006) defined represented participants as "the participants who constitute the subject matter of the communication; that is, the people, places and things (including abstract 'things') represented in and by the speech or writing or image" (p.48). The represented participants in image two are the boy's parents, establishing direct eye contact with the viewer in order to create an imaginary relation. In this case, the represented participants are performing a demand, realizing a visual 'you' (Kress and Van Leeuwen, 2006). Additionally, the represented participant placed in the background indicates the location, depicted by an iconic monument of Mexico City. Images 3, 4, and 5 are regarded as analytical processes which are composed of part-whole relations. Image 3 represents the trajectory described in the story. Image 4 illustrates one of the most important settings, the aquarium, where the main character bought an axolotl, and frame 5 depicts in more detail the different fish tanks found in the aquarium.

Figure 9. Sequence of screenshots 2

Image 6. The boy and the seller

Image 7. Buying the axolotl

Image 8. Asking for permission

Image 6 and 7 contain action processes, either transactional or non-transactional. The most salient participants in frame 6 are the boy, the seller, and the axolotl due to their size, color, position, and contrast against the background. In image 7, the boy's father is the actor, the

participant from which the vector emanates (Kress and Van Leeuwen, 2006), whereas the money and the axolotl are the goals, the objects of the action. On the other hand, image 8 is a reactional process, the woman is the reacter (the person who does the looking), and the man and boy are the phenomena (the participants at whom the reacter is looking).

Figure 10. Sequence of screenshots 3

Image 9. Buying a car

Image 10. Fixing the car

Image 11. Road map

According to Kress and Van Leeuwen (2006), circumstances of means concern the tools used to perform an action. In image 9, the hands are considered the circumstance of means. Circumstances of accompaniment deal with who is present during the action but not involved in any way with it, that is to say, "no vectorial relation with other participants" (Kress & Van Leeuwen, 2006, p. 75). Image 10 is an example of a circumstance of accompaniment due to the lack of vectorial relation between the car and the man. Locative circumstances relate to other participants in the foreground in terms of darkness and lightness, color saturation, detail, or focus. Image 11 shows an example of a locative circumstance, the truck within a map, which represents the location mentioned by the teller.

Figure 11. Sequence of images for visual coherence

The images present different movement effects during the video to emphasize certain crucial elements in the storytelling. It is possible to find the same image but edited in different ways to show a change in the story or a time movement. The images shown in figure 7 are examples of the preservation of the main features and how they guarantee the identification of the topic, even though the addition of certain elements might suggest a change. Regarding the storytelling, the visual mode is an excellent complement to the verbal mode. Combining both modes help the audience better understand the video's content since the election of pictures is relevant to what the student is expressing verbally and at the same time connects with the meaning they are intended to convey.

Video 2: A visit to the doctor

A. Participants and processes.

Figure 12. Sequence of screenshots 4

Image 12. Phone call

Image 13. Doctor

Image 14. Patient

Image 15. Writing a prescription

Image 16. Patient's facial expression

In terms of conceptual representation, image 12 shows a symbolic process, that is, "about what a participant means or is" (Kress & Van Leeuwen, 2006). The represented participant is a phone screen with the word "doctor" on it. This phone call symbol provides the viewer with a general idea of what the video is about and depicts how the communication is established between the doctor and the patient. Furthermore, the most salient represented participants are the doctor and the patient. In image 13, we can observe that the most prominent characteristic of the doctor is the white coat. In contrast, the most identifiable feature of the patient in image 14 is his physical appearance, mainly his red nose, which denotes a type of illness, the flu. In the previous pictures, both the "doctor" and the "patient" are considered actors since the vector emanates from them (Kress and Van Leeuwen, 2006). Images 13 and 14 show action processes; in this case, both participants have a phone conversation. The screenshots also illustrate non-transactional action processes since there is only one participant and no goal.

The reactional process involves the reactor and the phenomena, in which "the vector is formed by an eye line, by the direction of the glance of one or more of the represented participants" (Kress & van Leeuwen, 2006). Images 15 and 16 contain reactional processes; in image 15, the vector

formed by the actor's eye-line is directed to the prescription. On the other hand, in image 16, the physical discomfort and anxiety on the reactor's face indicate illness and fatigue.

B. Contact

In terms of contact, in images 17, 18, and 19, the represented participants do not look directly at the viewer but rather at other sides; therefore, these images offer information. The represented participants become the object of what Kress and Van Leeuwen (2006) called “the viewer’s dispassionate scrutiny” (p. 119), and the viewer becomes an “invisible onlooker” (p. 119). Images 17 and 18 depict a doctor talking to her patient. As it can be observed, the gaze changes direction but never looks directly at the viewer. Image 19 portrays a sick boy talking on the phone. His eyes are directed not at the viewer but at something outside of the image frame. Throughout the video, the offer act is the most prominent due to the nature of the story.

Figure. 13: Sequence of screenshots 5

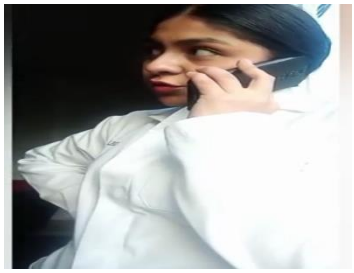


Image 17. Gaze

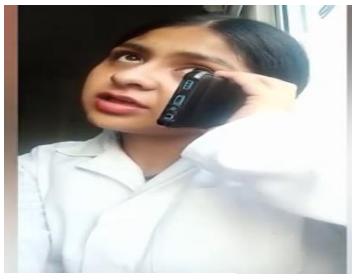


Image 18. Gaze

Image 19. Sick boy

C. Social distance

Physical proximity and intimacy can be determined by social distance. In images, this kind of distance translates to close, medium, and long shots. The various types of shots deal with how close or far to the viewer the represented participant seems. Images 17, 18, and 19 are a sample of the most representative shots shown in the video. Close shots involve the participant's head and

shoulders taking up approximately half of the screen, creating an intimate relationship with the viewers. The close shot captures the participants' gestures and allows the viewer to understand better the emotions and feelings they intend to transmit.

D. Composition

In images 20, 21, and 27, the title placement draws the viewer's attention to the center zone of the slide. As Kress and Van Leeuwen (2006) argued, the elements placed at the center of the image provide essential information of the visual space. The information presented helps the viewer understand the next part of the video and establishes a time movement between the previous scene and the next one. In image 22, the participant who plays the role of a patient is placed at the center of the slide. The word "fever" is located on the upper left side, whereas a picture that illustrates this word's meaning is placed on the lower right side. As Kress and Van Leeuwen (2006) pointed out, the elements placed on the left are presented as 'given' and are known. Conversely, the elements placed on the right zone are presented as 'new' since they are unknown. Additionally, the word "fever" and the picture are considered margins because they are placed around the center. Images 23, 25, and 26 are examples of the previous description and have a "triptych" composition that combines given and new with center and margin elements (Kress & van Leeuwen, 2006).4

Figure 14. Sequence of screenshots 6

Image 20. Linguistic text

Image 21. Medical background

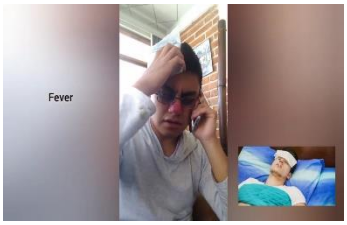


Image 22. Fever

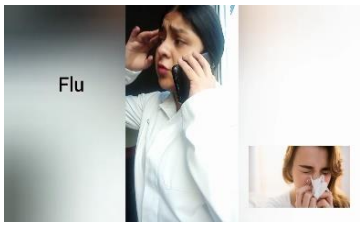


Image 23. Flu

Image 24. Prescription

Image 25. Tablets

Image 26. Check-up

Image 27. A week later

In image 20, the statement “Meanwhile in the doctor’s office” is presented in a linear reading path and written on small rectangular cards, as well as the statement “A week later…” in image 27. The messages are written in capital letters to emphasize the information given, and the contrast of black and white colors stands out and projects power. The messages become a salient element due to the letter font, size, and color contrast. In image 21, the statement “Meanwhile in the doctor’s office” appears again. Nonetheless, this time, the background is illustrated by some images representing medical items to give more prominence to the written message. Images 22, 23, 25, and 26 utilize linguistic and visual meaning-making resources. The linguistic text is written in black font in order to maximize its impact and highlight its importance within the whole image. The represented participant in image 22 is a patient describing his symptoms “I have a headache and fever.” The small image placed on the lower right side represents the main participant and visually illustrates the meaning of the word “fever”—the gestures made by the participant signal physical discomfort and pain. The participant’s hand is placed over his head to visually emphasize the linguistic resource and connect the visual and the linguistic mode. Driskell (2003) reinforced

the importance of gestures in human conversational interaction and argues that gestures enhance listener comprehension and speech production. Moreover, Krauss et al. (1996) claimed that conversational gestures accompany speech and express meaning related to the semantic content of the speech they accompany. Image 23 focuses on the other main participant of the story, the doctor, a young woman wearing a lab coat, a representative feature of the medical field. The small image placed on the lower right side of the frame visually represents the linguistic text “ok, you have the flu.” In image 24, the doctor appears again, but this time the image depicts a medical prescription to reinforce and illustrate the meaning of the verbal mode.

Video 3: Elections

A. Participants and processes.

The representational meaning is achieved through two types of structures: narrative and conceptual. Regarding narrative representations, image 28 involves an action process; the actor is the represented participant performing the action of carrying, whereas the goal is the dog who is the object of the action. Images 29-33 contain reactional processes, with vectors formed by their eye lines. At the same time, these images are considered non-transactional action processes since there is only one actor and no goal. In terms of conceptual representation, images 34 and 35 are symbolic suggestive processes and show the candidate's proposals.

Figure 15. Sequence of screenshots 7

Image 28. Carrying a dog

Image 29. Candidate's introduction

Image 30. Pantries

Image 31. In a farm

Image 32. Demand

Image 33. Vote

Image 34. Proposals

Image 35. Proposals

A. Contact

In terms of contact, in images 28-33, the participant looks directly at the viewer. In this case, the represented participant is performing a demand and consequently establishing an imaginary relationship. As Kress and Van Leeuwen (2006) argued, "when represented participants look at the viewer, vectors, formed by participants' eyelines, connect the participants with the viewer. Contact is established, even if it is only on an imaginary level." In image 32, the represented participant seems to address the viewers with a visual "you," to persuade the audience to do

something, in this context, to have their vote. The candidate is depicted as a strong leader by establishing direct eye contact with the viewer.

B. Social distance

The most common shots in this project are close, and images 281-33 clearly illustrate them. Close shots involve the participant's head and shoulders taking up approximately half of the screen, creating an intimate relationship with the viewers. The close shot captures the participants' gestures and allows the viewer to understand better the emotions and feelings they intend to transmit.

Figure 16. Sequence of screenshots 8

Image 36. Student's name

Image 37. Political party logo

Image 38. Candidate, flag and logo

Image 39. The elderly

Image 40. Teachers' salary

Image 41. Giving pantries

Image 42. Agriculture

C. Composition

The video begins with the student's name written in colorful letters and placed on the middle of the slide (image 36). The name of the political party appeared right after the student's name (Image 37). Both linguistic texts were located at the center of the slide within a white background in order to draw the viewer's attention to this zone and provide them with essential information about the visual space. Since the main character is the student performing the role of a presidential candidate, he was placed in the foreground and the center of the slide, highlighting that he is the primary social actor represented and, therefore, the most salient element of the composition (Image 38). Additionally, the political party logo is located on the right upper side in the ideal-new zone, making it visible and the second most salient feature of the text due to its color, size, and position.

In image 38, the most salient element is the main character, a young boy dressed as a presidential candidate. The boy is wearing a black suit, a red tie, and his political party pin, assembling a formal "personal front." According to Scollon and Scollon (2003), the personal front refers to "a kind of identity kit that one assembles out of the mixed bag of what Goffman calls 'sign equipment,' personal and physical characteristics and objects one might wear or carry" (p.57). Clothing is considered a tool of communication, and in frame 38, clothing denotes power and high status, elements of any presidential candidate.

Moreover, images 39 and 40 show a visual representation of the linguistic text "If I were elected president, I would financially support older adults. If I became president, I would raise the teachers' salary". As the main participant talks, visual elements emerge to support what he is saying to clarify any doubt the audience may have regarding vocabulary. This is a good option to connect the verbal mode with the linguistic one and make easier for learners the acquisition of new words and structures.

Body language plays an important role in nonverbal communication, and posture is an essential component of this type of communication. In frames 38-40, the participant shows a straight posture that might indicate he is focused and paying attention to what he is saying. He also clasps his hands over his lower abdomen, a position which is commonly observed when leaders and politicians stand to pose for photographs and videos. The participant might have chosen this position unconsciously because he might have felt vulnerable or even insecure. However, this position helped him to display confidence and respect. The participant remains in the same posture from minute 0:07 to 0:54, and we cannot observe any hand movement during this time-lapse. The crossed arms posture in image 41 might have various meanings. The first one is that arm crossing means defensiveness which somehow denotes a negative meaning. The second meaning is that the person might be attempting to focus or create an impression of power and control, and the third possible meaning is simply a sign of comfort. Regarding image 33, the thumb up is the most salient gesture accompanied by a big smile. The thumbs-up gesture is perhaps the most common of hand gestures and, in most Western cultures, is the signal for approval or that everything is under control. In this specific context, the thumb-up gesture supports and emphasizes the linguistic text "vote for me, you won't regret it."

Although the main character was wearing formal clothes (a suit and a tie) in previous shots, in images 28, 41, and 42, we can observe that he has radically changed his outfit. On this occasion, he has taken off his jacket, and instead, he has worn a more comfortable shirt, even a cap in frame 41 and a hat in frame 42. His new look makes him appeared as a 'politician' more approachable and closer to the audience than in the previous shots.

The most salient element in frame 41 is the candidate wearing a white t-shirt and a red cap with his political party logo. In order to support the linguistic text "If I were elected president, I

would help the poor" an image about people of lower socioeconomic status receiving support emerged on the left upper side to highlight the candidate's proposal. A farm can be seen as part of the background, whereas the candidate is portrayed wearing a brown linen shirt and a hat in the foreground. An image of a tractor is placed on the right upper side as support for the linguistic text "If I became president, I would help the agriculture sector." Furthermore, in image 28, the candidate is carrying a dog as a visual support of the linguistic text "If I won the election, I would make animal cruelty a serious crime."

4.5 Results of the multimodal analysis

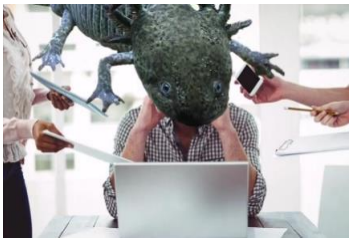
Video 1: Digital storytelling

The following chart shows the results obtained from the multimodal analysis.

Table 17. Multimodal Analysis Video 1

Visual mode	Linguistic mode	Description
	"Hello, my name is Kevin. I was a 13-year-old boy when my parents and I were in Mexico City"	In the first scene, the main character introduces himself. The image of a boy comes out within a black background. When the boy mentions his age, the number 13 suddenly appears. The boy wears a blue and white t-shirt, and the color of the number is white. The colors match, and the black background emphasized the two images.
	"My father bought a car in an insurance carrier"	The picture depicts multiple logos of different insurance companies in order to reinforce and illustrate the verbal information. In the screenshot, we can observe a tall white man wearing a blue shirt and suit. This person represents the boy's father. On the other hand, the car salesman is wearing a black suit and a tie.
	"So, my father had to repair the car" (Hallelujah chorus sound effect→aural mode)	In order to support the idea of repairing, an image of a mechanic emerged. The following screenshot shows a good-looking car accompanied by the hallelujah sound effect, which denotes an effect of relief and achievement because the car had been repaired.

 	"After repairing the car, my father was driving to San Martín but in the trajectory was a fish market"	The student decided to illustrate the trajectory with a road map, placing a car in movement heading to San Martín. The same road map appears when he mentions the fish market, but the car changes the trajectory to the opposite side.
 	"I was excited because I saw a Mexican axolotl for sale and ask my father - will you buy me one, please? I'll take care of him."	The first screenshot depicts the boy, a saleswoman, and an axolotl with an aquarium as a background image. The second screenshot contains the same background picture, but only with the boy and the father.
	Father: "Is it a promise? Well, I am going to check with your mom." Kevin: "We went to the car, when we arrived my father asked her and my mom answered -yes, but Kevin, promise to take care of him-"	The screenshot shows the mom, the dad, and the boy. In this specific scene, the role of the mother when making important decisions is emphasized. Another important fact is the voice intonation with which the student interprets his dad and mom. When he interprets his mom, his voice is softer and sweet, but his voice is gravelly and gruff when he interprets his dad.
	"So, my father and I went to the store and bought the axolotl"	The screenshot depicts the exact moment when the man buys his son the pet. To support the linguistic text, we can observe the man purchasing the axolotl, a wad of bills, and a pair of hands illustrating this action.
	"I was very happy with my new pet, rather illegal pet."	The elements presented in the picture (police officers, patrols) emphasize the meaning of "illegal pet." In addition, the axolotl within a red circle means that it is forbidden to own it as a pet. This image reinforces, even more, the meaning of "illegal" and establishes a connection between both modes.

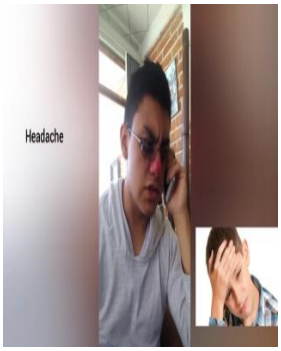
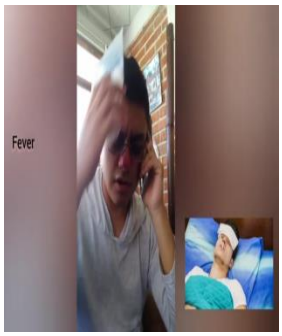
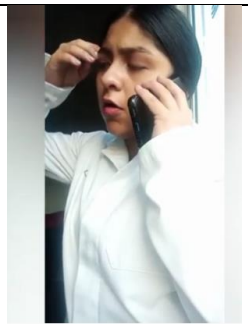
	"But my father had other plans, we went to Puebla city"	The picture illustrates an iconic landmark of Puebla city.
 	"The axolotl died after a month because of stress"	The first picture shows a dead axolotl, and one of the most representative features of dead animals is the tongue out. On the other hand, the second picture illustrates the state of being stressed. In this case, the animal seems to be a human being dealing with labor issues.

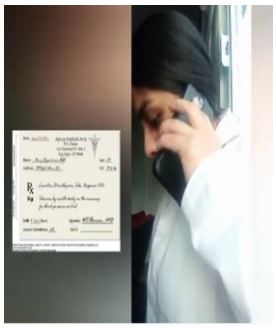
Video 2: A visit to the doctor

The following chart illustrates the results obtained from the multimodal analysis.

Table 18. Multimodal Analysis Video 2

Scene	Visual mode	Linguistic mode	Other modes	Description
Calling the doctor			Aural mode (Ringing sound effect)	The video begins with a phone call. In order to illustrate this action, the student decided to display a screenshot of an actual call.
Asking for symptoms		"How are you? what is the problem?"	Gestural mode	The girl is wearing a white lab coat, the most salient element of a doctor. Due to her facial expression, we can observe she is paying attention to what her patient is saying. At the same time, her body language shows

				she is interested in what the other person is communicating.
Describing symptoms	 	"I have a headache and fever"	Gestural mode Spatial mode	In the first screenshot, the word "headache" appears on the left upper side, the image of the students performing the role of the patient is placed on the center, and the image of a boy touching his head as a visual representation of the linguistic text is placed on the right lower side. Visual modes serve to consolidate and reinforce the meaning of linguistic modes. In the second screenshot, the boy places a cold pack on his forehead to reduce the temperature. The student has painted the tip of his nose with the color red to emphasize his physical discomfort and make his symptoms more real. The participant's body language is another crucial element that contributes to conveying meaning, and his facial expressions show pain.
Giving a diagnosis		"Ok, you have the flu"	Gestural mode	The participant places her hand on her head as a signal of concentration. Her facial expression shows staidness which might imply concern for the patient's health.

Giving a prescription		"I am going to give you a prescription"	Gestural mode Spatial mode	An image of a prescription is placed on the left side, whereas the doctor's image is placed on the right side. The linguistic text "prescription" is supported by the visual mode, and both of them combine to convey meaning.
Prescribing medication		"You should take tablets every three hours for three days"	Gestural mode Spatial mode	The three elements presented in the screenshot are arranged in order of importance. The most salient element is the picture of the doctor talking on the phone. This picture is placed in the center, and two more elements surround it. The linguist text "tablets" is located on the upper left side, the image illustrating the word's meaning is placed on the lower right side. The linguistic and visual modes come together to convey meaning and support each other.
Verifying the given information		"No chocolates?"	Gestural mode	The doctor told the patient, "You should not eat chocolates." This recommendation causes surprise to the patient. Therefore, he astonishingly replies, "-no chocolates?" based on his answer and body language, we can infer that the patient loves chocolates and cannot believe this food is forbidden. There is a connection among visual, linguistic, and gestural modes that converge to convey meaning.

Confirming the information		"Ha-ha-ha, yes, no chocolates and greasy food"	Gestural mode	The participant shows herself as more relaxed and friendly when she notes the tone of surprise in her patient. Her facial expressions depict happiness.
Planning the next appointment		Patient: "When should I call you again?" Doctor: "Your next check-up will be in two weeks"	Gestural mode Spatial mode	The image of the doctor is placed at the center in order to denote importance. The linguistic text "check-up" is located on the upper left side, and the image which illustrates the meaning of this word is placed on the lower right side.

Video 3: Elections

The following chart indicates the results obtained from the multimodal analysis.

Table 19. Multimodal Analysis Video 3

Visual mode	Linguistic mode	Other modes	Description
		Spatial mode	The first element we can observe in the video is the participant's name. The name is presented in a linear reading path, written with colorful gradient lettering. It attracts the viewer's attention because of its placement and contrast with the white background.

		Spatial mode	Right after the name, the logo of the political party is presented. The logo contains the initials “ASG”; however, the student does not mention what the initials stand for. The most salient element of this logo is the combination of colors green, white and red which allude to the Mexican flag and promote patriotism.
	“Hi, good morning. My name is Alfredo Sánchez García. I belong to the political party ‘ASG’”	Spatial mode Gestural mode	The image shows three salient elements, the candidate, the flag, and the political party logo. The participant is dressed as a presidential candidate wearing a black suit, a white shirt, and a red tie. The candidate's formal outfit projects responsibility, commitment, and sobriety. The candidate is also wearing a pin of his political party on the right lapel. This element makes his outfit more outstanding. The Mexican flag is located next to the candidate. The flag symbolizes patriotism, and its strategic location reaffirms the candidate's commitment to the nation. The candidate and the flag are the most salient elements of this image due to their color, size, and position. Regarding the participant's body language, his hands are not visible, but according to the arms position, it can be inferred that he has clasping hands over the lower abdomen. He has neutral facial expressions.

 	"If I were elected president, I would financially support the elderly as well as raise the teachers' salary"	Spatial mode Gestural mode Aural mode	As the candidate talks about his proposals, visual elements emerge in order to support the linguistic mode and clarify any doubt regarding vocabulary. In the first screenshot, the image of a cute elderly couple sitting on a bench is placed in the ideal zone. In the second screenshot, the candidate remains in the same position with the same neutral facial expressions—however, the image on the left changes for one, representing a teacher and a green bill. The location of these pictures allows the audience to visualize the candidate's proposals better. From this moment on, the video contains a piece of soft instrumental music.
 	"If I won the election, I would give food pantries to the people in need"	Spatial mode Gestural mode	The candidate is wearing a more comfortable outfit, a red cap, and a white t-shirt. One of the most salient details of the outfit is the political party logo. The participant has changed the position of his arms. His body language and physical appearance now communicate that he wants to be closer to his audience. Next to the candidate, there is a basket full of food. The content of the image directly connects with the linguistic text. In the second image, there is a group of children carrying these baskets of food. The purpose is to emphasize the linguistic mode as well as illustrate the candidate's proposals.
	"If you chose me as your president, I would build new gyms free of charge"	Spatial mode Gestural mode	The participant is wearing the same outfit and remains in the same position as the previous screenshot. As he talks about his proposal, an image of an outdoor gym is placed on the upper left side in order to reinforce the linguistic mode.

	"If I became president, I would help the agriculture sector"	Spatial mode Gestural mode	The participant's outfit has radically changed. Now, he is wearing a hat and a brown linen shirt with a pin of his logo party. The background looks like a farm. The scenario and the participant's outfit are according to what he is verbally expressing. As he talks about his proposal, an image of a tractor in a field comes out on the upper right side. Additionally, the participant's hands position and body posture are different. It seems that he is sitting, and his arms are by his sides. He adopts a more relaxing and confident posture. The visual mode emphasizes the linguistic mode, and both converge to convey meaning.
	"If I won the election, I would make animal cruelty a serious crime"	Gestural mode	The participant is wearing the same outfit as in the previous screenshot but without the hat. The most salient element in the image is a white dog carried by the participant. The primary participant intends to show his affection for animals in order to gain the people's trust. The participant looks more relaxed and confident. The visual mode supports the linguistic mode, and both combine to convey meaning.
	"Vote for me, you won't regret it"	Gestural mode	The participant appears again wearing a black suit, a red tie, and a pin of his political party. The Mexican flag is also located next to him. However, in this image, we can appreciate a different body language, including posture. His physical appearance and facial gestures transmit confidence and happiness. He is pointing to the viewer to ask for their vote. This action reinforces the linguistic text. His posture seems more relaxed than in previous scenes; he stands upright and places his hands by his sides. Overall, his body language project positivity and self-confidence.

	"Vote for me, you won't regret it"	Gestural mode	The transition of images accompanies the same linguistic text as in the previous scene. Nevertheless, more visual elements are added to this one. For instance, a pair of clapping hands, the word "vote" is presented in a linear reading path, written with colorful lettering and a ballot box. It is important to mention that all these elements are animated images. The participant is holding his hand out with his thumb pointed up to say yes or show approval.
			The video ends with three images, the Mexican flag, the political party logo, and the ballot box.

4.6 Discussion of findings

The findings will be presented in concordance to the research questions stated in the beginning of this research.

RQ1: What semiotic resources did the participants use in the digital multimodal text production?

Kress and Van Leeuwen's grammar of visual design (2006) was the framework used to analyze the semiotic resources students employed to construct meaning based on three metafunctions: the representational, the interactive, and the compositional. The participants of this study followed specific instructions to create the digital multimodal texts; however, they had the freedom to exploit their creativity and used a variety of resources to accomplish the objectives of the tasks. The group of certain semiotic resources is called modes of communication. The New London Group (1996) described modes of communication as resources that permit the design of meanings. The modes that predominate in the first multimodal text, "Digital storytelling," are visual and

linguistic. The latter tends to be the primary mode, whereas the former tends to be supportive. In the first video, the participant decided to include only images to accompany the linguistic mode, and just in the fourth scene included a sound effect. The visual semiotic resources used in this multimodal text were described in chapter four according to the represented participants, processes, circumstances, contact, social distance, salience, and information value. The most salient elements of each scene were determined by their size, color, and placement of the foreground and background. The represented participants provided the viewer with essential information to better understand the story. The student tried to represent himself with an image of a boy who shared physical characteristics such as skin color and complexion. The participant intended to reflect his reality through the use of images. Ganapathy (2016) claimed that visual design refers to "meaning-making based on colors, perspectives, size, shape, vectors, and background" (p.145). In the second and third multimodal texts, the most frequent modes used by the participants were the visual, the linguistic, and the gestural. Since participants recorded themselves, the gestural mode accompanied the linguistic and visual modes to provide more details to the viewer. Students' facial expressions in the second video, "a visit to the doctor," transmitted the emotions and feelings expressed by the participants verbally. Moreover, at the end of the third video, the candidate's hands position reinforced the linguistic mode and expressed the message "vote for me." Pullen and Cole (2010) argued that the gestural mode refers to the way movements are interpreted. It focuses on behavior, posture, facial expression, hand gesture, and body language. In terms of visual design, the student playing the doctor's role wore a white lab coat, the most salient element of a doctor. The "patient" painted the tip of his nose with the color red to illustrate a physical discomfort. Finally, the linguistic mode includes written or spoken words (Pullen &

Cole, 2010). The linguistic mode was used in the oral and written form in order to emphasize and support the visual and the gestural modes. Thus, it is clear that the linguistic and the visual modes complement each other, and images become crucial to support speech. According to the New London Group (2000), when two or more modes work together, it is called multimodal meaning. From the results of the videos made by the participants, most of them involved two to four modes of meaning.

RQ2: How did the different semiotic resources employed in the videos created by participants construe meanings?

The data analysis showed that the three metafunctional dimensions of meaning-making, representational, interactive, and compositional, were found in digital multimodal texts but with some differences in degree.

The findings confirmed that, at the representational level, narrative structures were employed in most of the screenshots. According to Kress and Van Leeuwen (2006), narrative representations occur when participants are connected by vectors and are represented as doing something to or for each other. In narrative structures, both action (transactional and non-transactional) and reactional processes were maintained by represented participants in the three videos.

The interactive meanings of the three videos were realized through contact and social distance. In videos 1 and 2, the analysis of the eye contact of the represented participants with the viewers indicated the represented participants were looking at something or someone within the image (offer). However, in video 3, the represented participant looked directly at the viewers, establishing an imaginary relationship (demand). The second element being analyzed from an interactive perspective was social distance. In all three videos, the close shots were the most

common involving the participant's head and shoulders taking up approximately half of the screen. These shots captured the participants' gestures and let the viewer better understand the emotions and feelings they intend to transmit.

Information value and salience, two compositional elements regarded in Kress and Van Leeuwen's (2006) framework, were found and analyzed in the multimodal texts under consideration. In videos 2 and 3, the title placement draws the viewer's attention to the center zone of the slide. As Kress and Van Leeuwen (2006) argued, the elements placed at the center of the image provide essential information of the visual space. The information presented helps the viewer understand the next part of the video and establishes a time movement between the previous scene and the next one. The represented participants also wore clothes that provided the audience more details about the content of the video. The most salient visual elements in the third multimodal text were the participants' clothing, the Mexican flag, and the political party logo. As the candidate talked about his proposals, visual elements emerged in order to support the linguistic mode and clarify any doubt regarding vocabulary. Although the main character was wearing formal clothes (a suit and a tie) at the beginning of the video, he wore a more comfortable shirt, cap, and hat in subsequent scenes. This new appearance made him a 'politician' more approachable and closer to the audience than in the previous shots.

Since the main character is the student performing the role of a presidential candidate, he was placed in the foreground and the center of the slide, highlighting that he is the primary social actor represented and, therefore, the most salient element of the composition.

RQ3: What are the benefits of implementing digital multimodal text production in public high school students?

Students' perceptions of the benefits of implementing digital multimodal text production in the EFL class were obtained from the focus group interview and the questionnaires conducted after the video projects were completed. The findings suggest that digital multimodal composing promotes students' engagement in the language learning process. Students noticed a difference before and after making videos. After producing videos, some participants got interested in keeping learning and practicing the target language. The findings also reveal that some participants realized the importance of learning a foreign language. Some of the recurring comments from the students are listed below:

> *It encourages me to continue learning the language. (S2)*
> *After making videos, I realized how important learning a foreign language is. (5)*
> *Since I started working with videos, I got interested in improving my pronunciation and the way of speaking in English. (S14)*
> *After working with the production of videos, I got interested in the language. (S17)*
> *I managed to get more interested in the subject because I was very excited to make videos. (S20)*

In parallel, Ganapathy (2016) found that multimodal approaches in meaning-making foster learner's engagement in the teaching and learning of ESL by enhancing their meaning-making abilities with the supplement of ICT as a tool.

Furthermore, most of the student participants' perceptions about implementing multimodal videos in the L2 classroom mainly coincide with the fact that DMC enhanced learners' creativity. According to the answers learners gave to a question about this fact. The findings reveal that fifteen participants considered that one of the benefits of DMC was the enhancement of creativity. Learners pointed out that they tried to find different ways to capture their audience's interest by making their videos entertaining, creative, and attractive.

> *I had to find ways to capture the interest of my audience. (S3)*
> *I tried to be creative at the moment of making videos. (S4)*

I really like working with the production of videos because it helps me enhance my creativity. (S9)
I developed my creativity when I tried to find out how to make my videos more attractive. (S10)
At the same time that I was trying to learn English, I improved my creativity. (S15)

The above excerpts reveal to what extent participants consider that digital multimodal composing contributes to their creative development. Since creativity is considered one of the 21st skills students need to possess to succeed in school and the workplace, DMC can be a platform that allows students to think outside the box and go beyond the conventional.

In addition, the findings of this study reveal that students strongly agreed that videos helped them improve their English language. The majority of the participants believe that DMC allowed them to practice and enhance their pronunciation which is a sub-skill of speaking. On the other hand, few learners reported that their listening, reading, and writing skills were improved.

I could improve my fluency and my pronunciation. (S8)
The videos helped me a lot to practice the vocabulary (S13)
It was a good opportunity to improve my pronunciation. (S3)
The videos helped me enhance my pronunciation a little bit more. (S17)
Creating videos really helped me improve my listening and pronunciation as well as my writing. (S18)

The aforementioned excerpts exemplify how the students benefited from digital multimodal composing and improved their English language skills, mainly pronunciation.

The multimodal video-making process also had a positive impact on learning. Findings suggest that participants could practice pronunciation and grammar while using different technological tools for the video edition. Participants pointed out that they learned to be more responsible and organize their time better. During the video-making process, students had to repeat, prepare and practice their script over and over again until they felt confident enough

speaking English in front of a camera. DMC allowed learners to express their ideas better while practicing the language more dynamically and gaining confidence. Participants also learned to be more organized and maximize their time. Some of the recurring comments from the students are listed below:

> *I learned to use different technological tools. (S2)*
> *As I had the opportunity to repeat and correct my mistakes at the end everything was clearer for me. (S6)*
> *I learned to edit better, I learned to be more careful with videos as they required to have a more formal concept. (S9)*
> *I can express my ideas and learn about new topics while improving my skills in front of a camera. (S14)*
> *I learned to express myself better in a foreign language and to be more confident when speaking in English. (S20)*

The implementation of video making in the English class showed significant benefits to the students. From the focus group interview and questionnaires, students believed that DMC helped them improve their language skills and confidence when speaking in English, even in front of a camera. DMC also contributed to promoting learners' engagement in the learning process of EFL. As the findings reveal, students appear to be positive-minded and motivated to keep learning and practicing the target language.

RQ4: What are the public high school learners' perceptions of implementing multimodal videos in the EFL classroom?

Students' perceptions of implementing multimodal videos in the L2 classroom were examined qualitatively through focus group interviews and questionnaires. The findings of this study reveal that most participants adopted a positive attitude toward the production of videos.

> *I like to edit videos. (S1)*
> *I liked working on the video creation. (S4)*

I prefer the process of creating videos, especially record them and edit them. It is easier to record yourself in front of a camera, because you have more confidence. You know that if you make a mistake, you can record it again or do it again. (S7)

The above excerpts demonstrate that participants liked working on the video creation and enjoyed the editing process. Findings suggest that creating videos made easier the process of speaking in English.

Moreover, the findings reveal that learners faced some challenges when working in the production of videos. Students reported that one of the difficulties they encountered was pronunciation. Some students also pointed out that they had to overcome their fear of speaking English since they got highly nervous when recording the videos. Students encountered difficulties in making videos due to lack of creativity, problems when editing the videos, and time limitation.

The only thing that was difficult for me was the pronunciation (S4)
Editing always takes me a lot of time, it is not so complicated but laborious. (S6)
It was a bit difficult for me since I am not used to speaking in front of the camera and that makes me so nervous. (S9)
Time was a big challenge. I needed more time. (S10)
At the beginning, it was a bit difficult for me to edit the video, and also, I was not very creative. (S13)
One of the great challenges I face was my fear of speaking in English. (S17)
I was nervous when speaking English. (S20)

In accordance with student's perceptions, they reported that making videos was a positive experience. The most recurrent comments are listed below.

It was a good experience because I could use many tools both physical and technological. (S2)
It was a very unique experience. I tried to do my best in making my videos interesting. (S4)
I had fun making videos in English. (S6)
It was a very good experiences, it gives you a twist where everything is repetitive and helps us put into practice everything we have learned, it is highly recommended. (S14)

It was something complicated, but funny. These kinds of activities make the language easier to understand. (S16)
At the beginning it was complex because I had never done something similar but later, I had fun making the videos. (S19)
It was a unique experience since I had never done this type of projects and I really loved it. (S17)

As shown in the comments above, most participants mentioned at least one positive aspect of working on DMC. Learners mentioned that their experience making videos allowed them to learn new things related to the target language and the editing process. In the comments, learners also mentioned that it was a complex process initially, but in the end, they had fun making the videos, and they could find a different and new way to practice the language.

The data above is the result and evidence that using videos can be a good and efficient aid to help students improve their English language skills, enhance creativity and overcome the fear of speaking English. Since DMC focuses more on the learner communicative competence centered in a real communication context, they deploy multiple skills.

4.7 Conclusion

This section presented the analysis of the different modes students used in the three digital multimodal texts and their perceptions toward digital multimodal composing in the EFL class. The multimodal analysis shows that students mainly used visual, linguistic, and gestural modes in the production of videos to maximize communication and capture the viewer's attention. The compositional was the most predominant metafunction. The compositional meaning is realized through the information value and salience (Kress and Van Leeuwen, 2006). According to the questionnaires and focus groups analysis, most of the students believe they enhanced their creativity and improved their pronunciation. In chapter five, the conclusions of this study will be

presented according to the analysis presented above. A more general discussion of the results is presented to provide the reader with a panorama of the implications of the current investigation.

Chapter V: Conclusions

5.0 Introduction

This study aimed to identify the semiotic resources participants used to make meanings when producing multimodal texts and explore their perceptions of the implementation of digital multimodal composing in the EFL classroom. Therefore, this final chapter provides a discussion of the analysis carried out in the previous chapter and presents a summary of key findings followed by the limitations that the current investigation faced. Possible areas for further research are addressed, and finally, the conclusions are drawn.

5.1 Summary of key findings

The modes that predominate in the first multimodal text (digital storytelling) are visual and linguistic. Whereas, in the second (a visit to the doctor) and third (elections) multimodal texts, the most frequent modes used by the participants were the visual, the linguistic, and the gestural. According to the New London Group (2000), when two or more modes work together, it is called multimodal meaning. From the analysis, most of the videos involved two to four modes of meaning.

Moreover, the data analysis shows that the three metafunctional dimensions of meaning-making were found in digital multimodal texts but differed in degree. At the representational meaning, represented participants, narrative and conceptual representation, and circumstances were found. The interactive meanings of the three videos were realized through contact and social distance and in the compositional meaning, elements of salience and information value were also found and analyzed. The findings also suggest that digital multimodal composing promotes the enhancement of creativity, English language improvement, students' engagement in the language learning process, and the positive impact of DMC on learning.

Concerning the students' perceptions of digital multimodal composing, the findings reveal that most participants adopted a positive attitude toward the production of videos. However, participants faced some challenges such as lack of creativity, problems when editing the videos, and time limitations. In accordance with students' perceptions, they reported that making videos was a positive experience.

5.2 Limitations of the study

The current study has two main limitations. The first one deals with time constraints; it would be of great relevance to give students more time to develop their digital multimodal texts. However, the context of the school did not allow it. Secondly, since the study was based on a small sample of participants, the results cannot be generalized. It is recommendable to use a more significant sample of participants in order to obtain more feasible results. Moreover, as the second part of the study intended to explore students' perception of the implementation of DMC in the English class, the extent to which the results are consistent with the facts is unknown. Finally, it would be more fruitful to emphasize the process students went through and not only focus on the final product.

5.3 Further Research

The study can serve as a point of departure to further research in the implementation of digital multimodal composing in the EFL classroom, considering the benefits this approach has on learners. Other studies can be conducted using similar research procedures to obtain more profound findings to compare the ones of the present study. Secondly, considering this research limitation, other experts in the area can carry out a similar study but follow different procedures to complement this study or explore other related areas. Overall, multimodality in the EFL classroom is a broad topic that can be explored from different perspectives.

5.4 Conclusions

This study outlines an implementation of digital text production in the EFL class, incorporating digital and multimodal literacy practices in the form of video projects. The aim of this study was two-fold. Firstly, it attempted to identify the semiotic resources students used to make meaning. Secondly, it examined students' perceptions of digital multimodal composing to explore the impact of videos on their learning process engagement.

The framework chosen for the multimodal analysis of the three videos was Kress and Van Leeuwen's grammar of visual design (2006). The findings reveal that the most frequent modes used by the participants were the visual, the linguistic, and the gestural. However, the visual proved to be the leading, supporting the linguistic and gestural modes. Incorporating digital multimodal text production in the EFL class can help students understand the meaning-making potential of different modes, particularly the relationship between words and images, in producing multimodal texts.

Furthermore, the multimodal analysis reveals that the three metafunctional dimensions of meaning-making: the representational, interactive, and compositional, were found in digital multimodal texts but with some differences in degree. At the representational level, narrative structures and represented participants were employed in most of the screenshots. The interactive meanings of the three videos were realized through contact and social distance. Regarding the compositional elements, images manifested salience over the linguistic text serving as visual support providing details of what was expressed orally.

The findings regarding students' perceptions of the benefits of implementing digital multimodal text production in the EFL class suggest that digital multimodal composing promotes students' engagement in the language learning process and provides them with opportunities to

enhance their creativity, an essential 21st-century skill. The findings also reveal that the benefits of implementing digital multimodal text production in the EFL class are: engagement in the language learning process, enhancement of creativity, English language improvement, and positive impact of DMC on learning.

In conclusion, today's learners must acquire new literacies and digital technologies to develop the 21st-century skills required to succeed (creativity, innovation, communication, collaboration, information literacy, media literacy, leadership, responsibility, information, media and technology skills). Therefore, the current research promotes multimodal learning experiences to enhance learners' multiliteracies without being restricted to one design mode.

References

Alonso, I., Molina Plaza, S., & Requejo, M. D. P. (2013). Multimodal digital storytelling. *Review of Cognitive Linguistics*, *11*(2), 369–387. https://doi.org/10.1075/rcl.11.2.10alo

Álvarez Valencia, J. A. (2016). Meaning Making and Communication in the Multimodal Age: Ideas for Language Teachers. *Colombian Applied Linguistics Journal*, *18*(1), 98. https://doi.org/10.14483/calj.v18n1.8403

Bezemer, J., & Kress, G. (2008). Writing in Multimodal Texts. *Written Communication*, *25*(2), 166–195. https://doi.org/10.1177/0741088307313177

Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., & Rumble, M. (2011). Defining Twenty-First Century Skills. *Assessment and Teaching of 21st Century Skills*, 17–66. https://doi.org/10.1007/978-94-007-2324-5_2

Brown, D. H. (2006). *Principles of Language Learning and Teaching (5th Edition)* (5th ed.). Pearson Education ESL.

Brown, J. D. (2001). *Using Surveys in Language Programs (Cambridge Language Teaching Library)* (1st ed.). Cambridge University Press.

Choi, J., & Yi, Y. (2015). Teachers' Integration of Multimodality into Classroom Practices for English Language Learners. *TESOL Journal*, *7*(2), 304–327. https://doi.org/10.1002/tesj.204

Cope, B., & Kalantzis, M. (2009). "Multiliteracies": New Literacies, New Learning. *Pedagogies: An International Journal*, *4*(3), 164–195. https://doi.org/10.1080/15544800903076044

Creswell, J. W. (2008). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches, 3rd Edition* (3rd ed.). SAGE Publications, Inc.

Darvin, R. (2015). Representing the Margins: Multimodal Performance as a Tool for Critical Reflection and Pedagogy. *TESOL Quarterly*, *49*(3), 590–600. https://doi.org/10.1002/tesq.235

Díaz-Barriga, A. (2020). La escuela ausente, la necesidad de replantear su significado. In *Educación y pandemía. Una visión académica* (pp. 19–29). iisue. http://www.iisue.unam.mx/nosotros/covid/educacion-y-pandemia

Driskell, J. E., & Radtke, P. H. (2003). The Effect of Gesture on Speech Production and Comprehension. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *45*(3), 445–454. https://doi.org/10.1518/hfes.45.3.445.27258

Ganapathy, M. (2016). The Effects of Using Multimodal Approaches in Meaning-Making of 21st Century Literacy Texts Among ESL Students in a Private School in Malaysia. *Advances in Language and Literary Studies*, *7*(2), 143–155. https://doi.org/10.7575/aiac.alls.v.7n.2p.143

Gentile, P. L., & Orué, A. M. (2012). Getting Teens to Really Work in Class. *English Teaching Forum*, *4*, 16–21.

Grapin, S. (2018). Multimodality in the New Content Standards Era: Implications for English Learners. *TESOL Quarterly*, *53*(1), 30–55. https://doi.org/10.1002/tesq.443

Hafner, C. A. (2013). Embedding Digital Literacies in English Language Teaching: Students' Digital Video Projects as Multimodal Ensembles. *TESOL Quarterly*, *48*(4), 655–685. https://doi.org/10.1002/tesq.138

Hafner, C. A. (2019). Digital Literacies for English Language Learners. *Second Handbook of English Language Teaching*, 1–20. https://doi.org/10.1007/978-3-319-58542-0_46-1

Hafner, C. A. (2020). Digital Multimodal Composing: How to Address Multimodal Communication Forms in ELT. *ENGLISH TEACHING*, *75*(3), 133–146. https://doi.org/10.15858/engtea.75.3.202009.133

Halliday, M. A. K. (1978). *Language as Social Semiotic*. Edward Elgar Publishing.

Harrison, C. (2003). Visual Social Semiotics: Understanding How Still Images Make Meaning. *Applied Theory*, *50*(1), 46–60.

Heberle, V. (2010). Multimodal literacy for teenage EFL students. *Cadernos de Letras*, *27*, 101–116.

Howe, N., Strauss, W., & Matson, R. J. (2000). *Millennials Rising: The Next Great Generation* (Highlighting ed.). Vintage.

Husbye, N. E., & Vander Zanden, S. (2015). Composing Film: Multimodality and Production in Elementary Classrooms. *Theory Into Practice*, *54*(2), 109–116. https://doi.org/10.1080/00405841.2015.1010840

Ilmi, Y., Retnaningdyah, P., & Munir, A. (2020). Exploring Digital Multimodal Text in EFL Classroom: Transformed Practice in Multiliteracies Pedagogy. *Linguistic, English Education and Art (LEEA) Journal*, *4*(1), 99–108. https://doi.org/10.31539/leea.v4i1.1416

Jewitt, C., & Henriksen, B. (2016). Social Semiotic Multimodality. *Handbuch Sprache Im Multimodalen Kontext*. Published. https://doi.org/10.1515/9783110296099-007

Jiang, L. (2017). The affordances of digital multimodal composing for EFL learning. *ELT Journal*, *71*(4), 413–422. https://doi.org/10.1093/elt/ccw098

Jones, R. H. (2013). Multimodal Discourse Analysis. *The Encyclopedia of Applied Linguistics*. Published. https://doi.org/10.1002/9781405198431.wbeal0813

Jones, R. H., & Hafner, C. A. (2012). *Understanding Digital Literacies: A Practical Introduction* (1st ed.). Routledge.

Kemmis, S., & McTaggart, R. (2005). Participatory Action Research: Communicative Action and the Public Sphere. In *The Sage handbook of qualitative research* (pp. 271–330). SAGE Publications Ltd.

Kinjari, K., & Ram Gopal, C. N. (2020). Learning style preference among adolescent school students. *The International Journal of Indian Psychology*, *8*(1), 944–955. https://doi.org/10.25215/0801.118

Koshy, E., Koshy, V., & Waterman, H. (2011). What is Action Research? In *Action Research for Improving Educational Practice* (pp. 1–24). SAGE Publications Ltd.

Krauss, R. M., Chen, Y., & Chawla, P. (1996). Nonverbal Behavior and Nonverbal Communication: What do Conversational Hand Gestures Tell Us? *Advances in Experimental Social Psychology*, 389–450. https://doi.org/10.1016/s0065-2601(08)60241-5

Kress, G. (2000). Multimodality: Challenges to Thinking about Language. *TESOL Quarterly*, *34*(2), 337. https://doi.org/10.2307/3587959

Kress, G., & Van Leeuwen, T. (2001). *Multimodal Discourse (Hodder Arnold Publication)* (1st ed.). Bloomsbury Academic.

Kress, G., & Van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). Routledge.

Lesiak, K. (2015). Teaching English to adolescents. *World Scientific News, 7,* 246–260.

LeVine, P., & Scollon, R. (2004). *Discourse and Technology: Multimodal Discourse Analysis (Georgetown University Round Table on Languages and Linguistics)* (1st ed.). Georgetown University Press.

Liang, W. J., & Lim, F. V. (2020). A pedagogical framework for digital multimodal composing in the English Language classroom. *Innovation in Language Learning and Teaching*, *15*(4), 306–320. https://doi.org/10.1080/17501229.2020.1800709

Miller, S. M. (2007). English Teacher Learning for New Times: Digital Video Composing as Multimodal Literacy Practice. *National Council of Teachers of English*, *40*(1), 61–83.

Moerdisuroso, I. (2017). Social Semiotics and Visual Grammar: A Contemporary Approach to Visual Text Research. *International Journal of Creative and Arts Studies*, *1*(1), 80. https://doi.org/10.24821/ijcas.v1i1.1574

Nelson, M. E. (2006). Mode, meaning, and synaesthesia in multimedia L2 writing. *Mode, Meaning, and Synaesthesia in Multimedia L2 Writing*, *10*(2), 56–76.

O'Halloran, K. L. (2011). Multimodal Discourse Analysis. *In K. Hyland and B. Paltridge (Eds) Companion to Discourse*. Published.

Pardede, P. (2018). Seeing Action Research Process in a Practice. *Collegiate Forum*, 282–295.

Partnership for 21st Century Skills. (2010). *21st Century Knowledge and Skills In Educator Preparation*. Pearson.

Peacock, S., Robertson, A., Williams, S., & Clausen, M. G. (2009). The role of learning technologists in supporting e-research. *ALT-J*, *17*(2), 115–129. https://doi.org/10.1080/09687760903033041

Prensky, M. (2001). Digital Natives, Digital Immigrants Part 1. *On the Horizon*, *9*(5), 1–6. https://doi.org/10.1108/10748120110424816

Reason, P., & Bradbury, H. (2008). *The SAGE Handbook of Action Research: Participative Inquiry and Practice* (2nd ed.). SAGE Publications Ltd.

Richards, J. C., & Farrell, T. S. C. (2005). Teacher Research and Action Research. In *Professional Development for Language Teachers* (pp. 255–271). Cambridge University Press. https://doi.org/10.1017/CBO9780511667237

Rowsell, J., & Walsh, M. (2011). Rethinking Literacy Education in New Times: Multimodality, Multiliteracies, & New Literacies. *Brock Education Journal*, *21*(1). https://doi.org/10.26522/brocked.v21i1.236

Setiyadi, A. B. (2020). *Teaching English as a Foreign Language* (2nd ed.). Graha Ilmu.

Spires, H. A., Morris, G., & Zhang, J. (2012). New Literacies and Emerging Technologies: Perspectives from U.S. and Chinese Middle Level Teachers. *RMLE Online*, *35*(10), 1–11. https://doi.org/10.1080/19404476.2012.11462093

The New London Group. (1996). A Pedagogy of Multiliteracies: Designing Social Futures. *Harvard Educational Review*, *66*(1), 60–93. https://doi.org/10.17763/haer.66.1.17370n67v22j160u

Tripp, D. H. (2005). Action research: a methodological introduction. *Educação e Pesquisa*, *31*, 1–21.

Van Leeuwen, T. (2005). *Introducing Social Semiotics*. Routledge.

Van Leeuwen, T. (2015). Multimodality in Education: Some Directions and Some Questions. *TESOL Quarterly*, *49*(3), 582–589. https://doi.org/10.1002/tesq.242

Yeh, H. C. (2018). Exploring the perceived benefits of the process of multimodal video making in developing multiliteracies. *Language Learning & Technology*, *22*(2), 28–37. https://doi.org/10125/4464

Yi, Y. (2014). Possibilities and Challenges of Multimodal Literacy Practices in Teaching and Learning English as an Additional Language. *Language and Linguistics Compass, 8*(4), 158–169. https://doi.org/10.1111/lnc3.12076

Yin, R. K. (2003). Collecting Case Study Evidence: The Principles You Should Follow in Working with Six Sources of Evidence. In *Case Study Research Design and Methods* (pp. 98–219). SAGE Publications.