UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE COMUNICAÇÃO E EXPRESSÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM INGLÊS

MARINHO CRISTIEL BENDER

EDUCATIONAL VIDEOS: ASSESSMENT OF THEIR PEDAGOGICAL FEATURES FOR L2 LEARNING

Florianópolis
2022

MARINHO CRISTIEL BENDER

EDUCATIONAL VIDEOS: ASSESSMENT OF THEIR PEDAGOGICAL FEATURES FOR L2 LEARNING

Tese apresentada ao Programa de Pós-Graduação em Inglês da Universidade Federal de Santa Catarina, como parte dos requisitos para a obtenção do título de Doutor em Inglês. Área de concentração: Linguística Aplicada.

Orientador: Professor Celso Henrique Soufen Tumolo, Dr.

Florianópolis
2022

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Bender, Marinho Cristiel
EDUCATIONAL VIDEOS: ASSESSMENT OF THEIR PEDAGOGICAL FEATURES FOR L2 LEARNING / Marinho Cristiel Bender ; orientador, Celso Henrique Soufen Tumolo, 2022.
227 p.

Tese (doutorado) - Universidade Federal de Santa Catarina, Centro de Comunicação e Expressão, Programa de Pós Graduação em Inglês: Estudos Linguísticos e Literários, Florianópolis, 2022.

Inclui referências.

1. Inglês: Estudos Linguísticos e Literários. 2. EDUCATIONAL VIDEOS. 3. VIDEOS DESIGN. 4. L2 LEARNING. 5. COGNITIVE THEORY OF MULTIMEDIA LEARNING. I. Soufen Tumolo, Celso Henrique. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Inglês: Estudos Linguísticos e Literários. III. Título.

Marinho Cristiel Bender

EDUCATIONAL VIDEOS: ASSESSMENT OF THEIR PEDAGOGICAL FEATURES FOR L2 LEARNING

O presente trabalho em nível de doutorado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Professor Dr. Leonardo da Silva
Instituto Federal de Santa Catarina (IFSC)

Professora Dra. Nayara Nunes Salbego
Instituto Federal de Santa Catarina (IFSC)

Professor Dr. Rodrigo Schaefer
Instituto Federal Catarinense (IFC)

Professor Dr. Bruno de Azevedo
Instituto Federal de Santa Catarina (IFSC) (Suplente)

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de Doutor em Inglês.

Documento assinado digitalmente
Celso Henrique Soufen Tumolo
Data: 03/03/2023 15:14:32-0300
CPF: ***.275.298-**
Verifique as assinaturas em https://v.ufsc.br

Professor Dr. Celso Henrique Soufen Tumolo
Orientador

Documento assinado digitalmente
Celso Henrique Soufen Tumolo
Data: 03/03/2023 15:15:02-0300
CPF: ***.275.298-**
Verifique as assinaturas em https://v.ufsc.br

Coordenação do Programa de Pós-Graduação em Inglês

Florianópolis
2022

ACKNOWLEDGEMENTS

I wish to express my gratitude to all who were involved in my Doctorate studies.

To professor Dr. Celso Henrique Soufen Tumolo, who so gently accepted to be my advisor for the Doctorate and for sharing all his knowledge. Also, for all the valuable discussions we had concerning the difficulties and the doubts I had and all the precious feedback throughout my research study.

To all the professors at PPGI, for the classes I attended and who always provided fruitful insights for my research study.

To all the people at the PPGI secretary, Valdete and the interns, who were so committed with all the bureaucratic processes that a secretary demands.

To professors Dr. Kyria Rebeca Finardi and Dr. Nayara Nunes Salbego, for reading my research proposal and pilot study, and providing me with important feedback concerning my writing and the analysis choices in my qualifying exam.

To professors Dr. Leonardo da Silva, Dr. Nayara Nunes Salbego, Dr. Rodrigo Schaefer and Dr. Bruno de Azevedo for kindly accepting to be part of the evaluation comittee for my defense.

To CAPES for the grant, which was essential for my complete dedication to the research developed.

To my parents, Laci and Mário, who encouraged me to move on with my studies.

To my wife Joice, and my stepdaughter, Ágatha, for always supporting me, always believing in my potential, and always carefully listening to me while I lectured them about all the theories and concepts I learned throughout my research.

The issue is not what learners do with educational materials,
but what educational materials do to their minds.
(Marinho Cristiel Bender)

ABSTRACT

With the development of digital technology, the number of materials aimed at facilitating the learning of an L2 has increased in the past years. From the large amount of digital resources and environments available, videos can be considered of high value in the process of providing learners with the target language input. Researchers have claimed that the use of videos *per se* may not guarantee that the L2 will be learned and that, for effective learning to take place, there is the need to look upon their design regarding cognition. Considering that the design of educational materials should follow how the human cognitive architecture is structured, Mayer (2009), presenting the Cognitive Theory of Multimedia Learning, proposed twelve principles that should guide the design of educational materials that make use of both visual and oral language to promote better learning grounded on the Cognitive Load Theory, extensively discussed by researchers as Sweller (2010) and Kalyuga (2015). Based on the assumption that the design of educational videos can have direct impact on learning either by promoting it, or hindering it, the objective of this Doctoral Dissertation was to analyze ten educational videos produced for the promotion of L2 learning based on three of the twelve principles, namely Signaling, Segmenting and Coherence, with a focus on their design features, and how they could possibly affect the learning process. The ten videos were taken from two sources: the first consists of five videos accompanying three textbooks that were approved by the PNLD (National Program of Didactic Material and Book) program, and the second consists of five videos that accompany the Interchange program textbooks. Although the videos from the two sources present noticeably different features in their design, the analyses of the ten videos provided evidence that the three principles for the design of educational videos may have been part of their design. The first principle, signaling, was examined for the fact that the videos make use of visual or/and oral devices aimed at making specific pieces of information salient to learners in order to direct their attentional resources to them by means of the use of imagery and/or visual effects or with spoken language and/or intonations, respectively. The second principle, segmenting, could also be detected in the videos design due to the fact that the information to be presented to learners is organized in blocks, so to prevent them from engaging in extraneous processing due to overloading their working memory limits. Finally, the third principle, coherence, was also analyzed in the videos design. In spite of being aimed at the while-designing phase, in which information that can be considered unnecessary for not having direct impact on learners' comprehension (and ultimately learning) should be removed, the coherence principle guided the analyses so as to detect which elements in the videos design could hurt learning for being considered irrelevant. Overall, it was possible to conclude that, although the educational videos analyzed differed in their design, the three principles could be seen as aiming at promoting the learning of the L2. The principles were observed in the forms of signaling important pieces of information, presenting the information in blocks and, at times, presenting information that may be irrelevant for not being crucial for comprehension, thus, promote learning. Despite the fact that the analyses remained theoretical, empirical research considering how different video design features can aid or hurt learning is needed.

Keywords: Educational Videos; Videos Design; L2 Learning; Multimedia Learning; Cognitive Theory of Multimedia Learning; Cognitive Load Theory.

RESUMO

Com o desenvolvimento da tecnologia digital, o número de materiais que objetivam facilitar a aprendizagem de uma L2 tem aumentado nos últimos anos. Da ampla quantidade de recursos e ambientes digitais disponíveis, vídeos podem ser considerados de enorme valia no processo de proporcionar insumo da língua alvo aos aprendizes. Pesquisadores alegam que o uso de vídeos *per se* pode não necessariamente garantir que a aprendizagem da L2 aconteça e que, para que a aprendizagem efetiva ocorra, é necessário um olhar sobre o seu design em relação à cognição. Considerando que o design de materiais educacionais deveria se basear em como a arquitetura da cognição humana é estruturada, Mayer (2009), apresentando a Teoria Cognitiva da Aprendizagem Multimídia, propôs doze princípios que deveriam guiar o design de materiais educacionais que fazem uso de imagens e da linguagem verbal oral e escrita para promover melhor aprendizagem com base na Teoria da Carga Cognitiva, amplamente discutida por pesquisadores como Sweller (2010) e Kalyuga (2015). Com base na suposição de que o design de vídeos educacionais pode ter impacto direto na aprendizagem tanto promovendo-a, como dificultando-a, o objetivo desta Tese de Doutorado foi de analisar dez vídeos educacionais produzidos para promover a aprendizagem de L2 com base em três dos doze princípios propostos por Mayer (2009), que são Sinalização, Segmentação e Coerência, com foco nos recursos em seu design, e como eles poderiam possivelmente afetar o processo de aprendizagem. Os dez vídeos foram obtidos de duas fontes: a primeira consiste em cinco vídeos que acompanham três livros didáticos aprovados pelo Programa Nacional do Livro e Material Didático (PNLD), e o segundo consiste de cinco vídeos que acompanham os livros do programa *Interchange*. Embora os vídeos das duas fontes apresentam recursos visivelmente distintos no seu design, as análises dos dez vídeos geraram evidências de que os princípios para o design de vídeos educacionais podem ter sido parte de seu design. O primeiro princípio, sinalização, foi analisado pelo fato de os vídeos fazerem uso de dispositivos visuais ou/e orais com o objetivo de tornar partes específicas da informação salientes aos aprendizes a fim de direcionar seus recursos de atenção por meio do uso de efeitos imagéticos e/ou visuais, ou com linguagem falada e/ou entonação, respectivamente. O segundo princípio, segmentação, também pôde ser identificado no design dos vídeos pelo fato de a informação a ser apresentada aos aprendizes estar organizada em blocos para preveni-los do processamento supérfluo por exceder os limites da memória de trabalho. Por fim, o terceiro princípio, coerência, também foi analisada no design dos vídeos. Embora o seu uso seja para a fase de criação, na qual informações que podem ser consideradas desnecessárias por não impactar diretamente na compreensão dos aprendizes (e consequentemente na aprendizagem) deveriam ser removidas, o princípio da coerência guiou as análises a fim de detectar quais elementos do design dos vídeos poderiam prejudicar a aprendizagem por ser considerados irrelevantes. De maneira geral, foi possível concluir que, embora os vídeos educacionais analisados apresentassem diferences em seu *design*, os três princípios puderam ser observados no que se refere ao propósito de promover a aprendizagem de uma L2. Os princípios foram verificados em forma de sinalização de partes importantes da informação, apresentação da informação em blocos e, por vezes, apresentar informação que pode ser irrelevante por não ser crucial para a compreensão, e assim facilitar a aprendizagem. Apesar de as análises terem se mantido teóricas, pesquisas empíricas considerando como diferentes recursos de design de vídeos podem contribuir ou prejudicar a aprendizagem se fazem necessárias.

Palavras-chave: Vídeos Educacionais; Design de Vídeos; Aprendizagem de L2; Aprendizagem Multimídia; Teoria Cognitiva da Aprendizagem Multimídia; Teoria da Carga Cognitiva.

LIST OF CHARTS

LIST OF FIGURES

LIST OF GRAPHS

LIST OF ABBREVIATIONS

CALL – Computer-Assisted Language Learning

CTML – Cognitive Theory of Multimedia Learning

EFL – English as a Foreign Language

L1 – First Language

L2 – Second Language

LTM – Long-Term Memory

PNLD – *Programa Nacional do Livro e Material Didático* (National Program of Didactic Book and Material)

PPGI – *Programa de Pós-Graduação em Inglês*

Sc – Schemata

SM – Sensory Memory

SSC – Signaling, Segmenting and Coherence

SSW – Signaling, Segmenting and Weeding

UFSC – *Universidade Federal de Santa Catarina* (Federal University of Santa Catarina)

WM – Working Memory

TABLE OF CONTENTS

1 INTRODUCTION

1.1 PRELIMINARIES

Technology has always been part of the second language (L2) classroom to enhance the teaching and learning processes, ranging from the use of the chalk board, textbooks presenting images working as extra aids for the understanding of the topics being studied, the use of the white board, cassette tapes, audio CDs, and DVDs with videos to name some (FINARDI; PORCINO, 2014; TUMOLO, 2014; 2017). All of these artifacts have been part of the second language[1] (L2) classroom and have been utilized by teachers and instructors in order to provide learners with better opportunities for the development of the language being learned and, as argued by Stanley (2013), should be currently part of the class environment.

In order to make information more rapidly accessible, as discussed by Sacerdote (2010), technology, once restricted to analogical or physical artifacts, has evolved to the digital format and has become part of people's lives, as can be seen in the use of computers, tablets and smart phones for communication, entertainment and information (CHINNERY, 2022). In the case of inserting digital resources in the context of learning an L2, such practice (and area of research) has been defined as computer-assisted language learning (CALL) in which, although the name suggests learning with the support of computers, the process of learning an L2 can be supported by the use of smart phones[2], tablets or any other digital resources (MARTINS; MOREIRA, 2012). The use of digital resources to provide learners with opportunities to develop an L2 has been accompanied by extensive research on their effectiveness and how they can aid or be an integral part of the context of teaching and learning an L2 (DAVIES; OTTO; RÜSCHOFF, 2013; MOTTERAM; SLAOUTI; ONAT-STELMA, 2013; THOMAS; REINDERS; WARSCHAUER, 2013; CHAPELLE; SAURO, 2017).

According to Blake (2017), the emergence of different digital resources and environments afforded by the Internet such as blogs, wikis and websites where users can have vast access to content of varied subjects and topics (*e.g.* YouTube) offered valuable materials for the teaching and learning processes (WATKINS; WILKINS, 2011; ALMURASHI, 2016;

[1] Although there has been extensive discussion on the terms Second Language (L2) Acquisition, Second Language (L2) Learning, English as a Foreign Language (EFL) and English as an Additional Language (EAL), in this Doctoral Dissertation, the term L2 learning will be used in order to refer to the learning of English as a foreign language in formal contexts, that is, in schools.

[2] The emergence of MALL (Mobile-Assisted Language Learning) as a subarea of CALL encompasses the use of smart phones in the process of learning an L2 (CHINNERY, 2006).

SHARIFF; SHAH, 2019; SHOUFAN, 2019). According to Chapelle and Sauro (2017), technology-supported environments can contribute to the process of learning and knowledge building once learners can have access to different types of materials, texts, genres and educational and social activities they may engage in, either in class or out of the class. From the countless digital resources that have appeared and been implemented in the L2 classroom aiming to provide learners with different opportunities to experience the language in varied contexts and with a wide array of objectives, one of the most significant resources can be videos.

The use of videos in the teaching and learning process has been investigated and discussed over the past decades (LONERGAN, 1984; FERREIRA; JÚNIOR, 1986; GREENFILED, 1988; MORAN, 1995; TUMOLO, 2012; 2017) concerning how they can be implemented in the L2 classroom to aid learners in the process of sense-making not only of the language itself, but also of the information they are presented with. In addition, a more critical view towards videos has been taken as to how they can aid the way information is processed, promoting learning (PLASS; CHUN; MAYER; LEUTNER, 1998; PLASS *et al.*, 2014).

Although videos in the L2 classroom were initially a peripheral source used mostly for broadcasting information, their importance in facilitating language learning rapidly began to be noticed (HAMBROOK, 1986). One example is Greenfield (1988), who brought a critical view concerning the use of images and words in the educational context. The author provided discussions regarding the benefits of providing learners with information from different sources, that is, when learners are provided with information through the electronic communication means, such as the television, cinema, and even video game, they are aided in developing their reasoning abilities and their cognition is improved (GREENFIELD, 1988). For that reason, for the author, audiovisual materials could be considered of high relevance in the educational context.

Wang (2015) explains that when incorporating videos in the L2 classroom,

> [t]he first and fundamental goal is to facilitate the development of EFL learners' language skills and in particular their listening and speaking skills. Video materials provide visual aids for the listening materials that EFL learners are exposed to. When they are practicing their listening skills, learners often find it difficult to hold their attention long enough once they are exposed to long conversations or passages without visual aids. Videos can provide much more information for listeners and can keep their attention focused on the aural material (WANG, 2015, p. 23-24).

The inclusion of videos in the L2 classroom provides learners with materials that make use of visual and aural information to present not only linguistic information, but also the

context in which the interaction takes place (TUMOLO, 2017). When watching videos, learners can see and hear complete conversational situations in which they can observe information such as the number of participants, their genders, ages and also facial expressions (LONERGAN, 1984), which, would not be possible to observe otherwise. By having access to such contextual clues, learners can have valuable input to aid them in the process of inference making, thus, aiding comprehension. In this case, not only can videos be considered a source for providing learners with meaningful input, but they can also support language processing.

In this sense, as explained by Tumolo (2017), the use of videos in the context of learning an L2 can offer a variety of input to enhance the learning of the target language, such as presentation of new topics, linguistic exposure and input, as well as formal instruction to language uses and explanations. According to the author, using videos in the educational context can be closely associated to the content-based approach, in which the language is not the goal of the study or of the learning process, but the means for knowledge-building processes resulting in the integration of the content being studied.

Another important factor that classifies videos as an important resource in the process of knowledge construction is that they can play a crucial cognitive role. When processing information, learners engage in the construction of mental representations of the information being received, that is, of the information being seen and heard. For the fact that videos are designed in a fashion presenting information by means of using images and sounds (spoken words) (MAYER, 2009), the construction of mental representations necessary for comprehension is more likely to be benefitted due to the support of the visual and aural input. In the effective processing of information, the three types of memory are at the service of learning: the first one is the sensory memory, in which the information enters learners' minds, in the case of videos, through the eyes and the ears. The second is working memory, which is responsible for the storage and the processing of the new input and the construction of mental representations of the information (MAYER, 2009; IBRAHIM *et al.*, 2012). The third is the long-term memory, responsible for storing the processed information as strings of knowledge.

Using videos in the L2 classroom, however, does not guarantee that learning will take place. Once, on the one hand, there seems to be an unlimited number of videos aimed at learning an L2, on the other, not every video can be argued to aid learners in the processing of information in the same way. Depending on the design used for their productions, videos can have positive effects. For that, their design must be cognitively suitable to aid comprehension and thus, learning. To be cognitively suitable to aid comprehension and thus, learning, videos

must be designed attending to the cognitive loads implied in the processing of the information it conveys.

Three cognitive loads have been discussed: the intrinsic, the generative and the extraneous (AYRES; PAAS, 2007; MAYER, 2009; MAYER; LEE; PEEBLES, 2014).

The intrinsic load is associated to the level of intricateness and complexity of the information presented in the video. The generative load is associated to information processing that actually leads to learning, which allows for new strings of knowledge to be generated. The extraneous load concerns the information that may be considered extra, superfluous, or unnecessary and irrelevant, thus hurting learning for exceeding learners' processing capacity (the three cognitive loads will be explained in detail in chapter 2).

Based on the three types of cognitive loads involved in the process of learning, Mayer (2009) has proposed a set of twelve principles that should guide the design of any material that makes use of the combination of images and written and oral words in order to promote learning (a discussion on this will be presented in chapter 2). For the fact that the author uses the term multimedia in a comprehensive way, in which any kind of educational materials that make use of images and words can be considered multimedia materials (MAYER, 2009), Ibrahim (2012) and Ibrahim *et al.* (2012) propose that three of the twelve principles may be used in the design of educational videos, referring to them as the Signaling, Segmenting and Weeding design (SSW). The first, the Signaling Principle, refers to the signaling of important pieces of information so to direct learners' attentional resources to them. The second, the Segmenting Principle, refers to the segmentation of the information in which learners are presented with the target information in small portions. And the third, the Weeding Principle, is associated with weeding out irrelevant information. The authors claim that

> […] weeding (*i.e.*, removal of non-essential content) can hypothetically prevent the learner from engaging in incidental processing, so that more cognitive resources can be devoted to the processing of essential content. (IBRAHIM *et al.*, 2012, p. 225).

Recognizing that videos can be a potential resource in aiding the process of learning an L2, this Doctoral Dissertation will deal with videos that accompany textbooks for the teaching and learning English as an L2.

1.2 RESEARCH OBJECTIVES

Acknowledging that videos can be considered a powerful source of input for the construction of knowledge, the main objective of this Doctoral Dissertation revolves around the features videos present in their design that can support learning. Considering that there seems to be a need for an effective evaluation of educational videos (GOMES, 2008), the objective of this study is to analyze ten educational videos that were produced with the objective of aiding learners in the process of learning English as an L2 as to whether their design can foster learning (LEE; MAYER, 2018). To this end, this Doctoral Dissertation will analyze videos from two different sources.

The first source refers to videos made available by the PNLD program (*Programa Nacional do Livro e Material Didático*) (BRASIL, 2020), which is a governmental program aimed at providing textbooks and pedagogical materials to public schools produced and distributed by different publishing companies. The second source refers to videos that accompany the Interchange textbooks published by Cambridge University Press, a popular collection of textbooks in Brazil. The Interchange Series is composed by four textbooks (Intro level, Interchange 1, Interchange 2 and Interchange 3), each one with sixteen units, presenting contextualized language features. Each of the units of the four textbooks has a range of extra materials, including a video available online on the Interchange website with free access to teachers and learners, to be used in class or as extra study materials.

Considering this context, the analyses of the ten videos in this Doctoral Dissertation will be grounded mostly on the SSW design proposed by Ibrahim *et al.* (2012). The Weeding principle, however, due to aiming at the removal of unnecessary information, thus, being aimed at the while-designing process, will be replaced by The Coherence Principle, proposed by Mayer (2009) due to the nature of this Doctoral Dissertation, that deals with videos that are already produced and available. Mayer's (2009) nine remaining principles for multimedia materials design are also briefly looked upon in the discussion section (the twelve principles will be explained in detail in Chapter 2). As a whole, by analyzing educational videos, this Doctoral Dissertation aims at investigating how their design can support L2 learning by possibly enhancing the processes involved in the construction of knowledge. Hence, the research questions that have guided this study are:

RQ1 Do the educational videos incorporate the principle of signaling, with features enhancing the L2 input? If so, how?

RQ2 Do the educational videos incorporate the principle of segmenting, facilitating the comprehension of the target language? If so, how?

RQ3 Do the educational videos incorporate the principle of coherence, maintaining relevant information to assist L2 learners? If so, how?

1.3 JUSTIFICATION FOR THE STUDY

This research study can be of high importance in the educational context firstly because it is undeniable that the digital technology has increasingly been part of people's lives for communicating and accessing information in a considerably reachable way (VIEIRA; FINARDI, 2018), thus making any research in the area of CALL relevant. Moreover, the dramatic changes in people's lives that began in 2020 with the COVID-19 pandemic concerning the ways individuals interact and exchange information seems to have brought onto the surface reflections and discussions of critical issues regarding the ways the educational system can be fragile to some extent.

The pandemic demanded the world population to change their lifestyles to a significant confinement style due to the physical distance the spreading of the disease demanded, in which most of the interactions had to be through a computer, tablet or smartphone screen. One of the most remarkable changes that the pandemic has caused can be associated to the educational context and the way learning is organized, demanding from teachers no other alternative but to develop their abilities in the area of teaching and providing learning opportunities through computers. While schools, universities and any other educational institutions were required to migrate the teaching and learning practices to the online mode, for the same reason, they also had to adapt the ways in which the information was distributed and presented for learners to access it.

In order to aid learners in the learning process, teachers and schools had to reinvent their forms of making information available and of encouraging learners to become more autonomous. The face-to-face encounters that regularly took place in schools, in which teachers and learners were together in the same physical location, were replaced by digital activities to be done at home and sent back to teachers via email and virtual meetings. Considering that teachers needed to find (digital) resources to make content and activities available for learners to access online, videos could have played an important role in reaching such objective.

Another important aspect for this research study is the fact that there seems to be a considerable lack of knowledge on the part of language teachers as well as video designers

regarding how learners' cognitive architecture (KALYUGA, 2020; SWELLER; AYRES; KALYUGA, 2011) is structured. Such lack of knowledge can have direct impact on learning, once when the choices concerning materials design is not informed, that is, not based on the knowledge about how information is processed in learners' minds, there is a possibility of making use of materials that hinder learning, instead of fostering it. The rationale for this Doctoral Dissertation is that there is a gap between teachers and designers' deeper understanding of learning from videos and the levels of processes that comprise the construction of knowledge.

Hence, the main objective of this Doctoral Dissertation is to provide an analysis of free of charge online materials that accompany textbooks, and provide teachers, instructors and material designers with insights about the kinds of materials to be chosen and/or developed for learning, and to consider the design features materials can present. The understanding of how learners' information processing system works can provide teachers with the capability of making informed choices when selecting videos, or even when designing their own materials (REED, 2009; BUCHNER, 2018) to be used in their L2 classrooms to best aid learners in the process of knowledge construction so that materials will not be randomly selected because of the appearance or length, for example, but because of the design quality, having the learners' needs always in mind.

1.4 ORGANIZATION OF THE DISSERTATION

This Doctoral Dissertation is organized in five chapters, each one composed of subsections so as to organize the information to be presented and grant a considerable sense of unfolding of the information. The chapters are organized as follows:

Chapter 1 presented the Introduction, with a contextualization of the use of technology for teaching and learning an L2, the objectives of this research study with the four research questions guiding it, and the justification for the relevance of this Doctoral research study in the context of L2 teaching and learning.

Chapter 2 presents two major discussions concerning videos. The first is presented in subsection 2.1, in which the contributions videos can have in the context of learning an L2 as well as the features such resource can offer in order to aid the learning process are discussed. Subsection 2.2 presents discussions concerning how videos can be defined as multimedia constructs based on Mayer's (2009) definitions of multimedia materials and how they can cognitively contribute with the learning process.

Chapter 3 presents in detail the method adopted in this research study, in which further information about the sources from where the videos are taken are presented. Furthermore, the criteria used for the selection and exclusion of videos are explained in detail followed by the explanation of the development and the presentation of the video analysis framework to ground the analyses. The framework was based on three major design principles: Signaling, Segmenting and Coherence.

Chapter 4 is organized in two main sections. 4.1 (the analyses of the videos) and 4.2 (disucssions about the analyses). Section 4.1 is organized in two subsections. The first, 4.1.1, presents the analyses of the five videos selected that accompany the textbooks selected in PNLD program, and the second, 4.1.2, presents the analyses of the five videos selected that accompany the Interchange textbooks. Lastly, section 4.2 presents a discussion concerning the analyses of the ten videos in relation to the three research questions previously presented, 4.2.1 (concerning signaling), 4.2.2 (concerning segmenting), and 4.2.3 (concerning coherence), as well a brefly presentation of Mayer's (2009) nine principles aside the three ones from the analysis framework that could be observed in the videos.

Finally, chapter 5 presents the final remarks with insights that this research study has offered followed by the presentation of the limitations of this study and suggestions for future research that may be conducted to deepen the understanding of how videos can contribute with the process of learning an L2.

2 REVIEW OF LITERATURE

This chapter will present the theoretical background concerning how videos can aid the process of learning an L2 and how their design can facilitate or disturb learning. The chapter is organized in two major sections (2.1 and 2.2).

Section 2.1, Videos in the L2 Teaching and Learning Process, provides a brief discussion concerning how videos can work as a powerful source in the context of L2 teaching and learning. The discussion comprehends videos as providing information in forms of imagery and verbal input, how such input can foster the processes involved in language learning.

Section 2.2, Multimedia and Pedagogy: Videos & Learning, presents discussions concerning videos from a cognitive perspective, how they can affect learning, organized in four subsections.

Subsection 2.2.1, Video as a Multimedia Resource, presents discussions concerning the features that multimedia materials have and how videos can be inserted into such category based on their design. Further discussion is presented concerning how the design of videos can affect learning, either positively or negatively.

Subsection 2.2.2, The Relation Between Educational Videos and Cognition, discusses how the information provided in videos is processed by learners cognitive system based on the Cognitive Theory of Multimedia Learning (CTML).

Subsection 2.2.3, Limitations of Working Memory, presents the implications of the processing limitations learners' working memory can pose on learning and how learning can be affected by the design of materials.

Finally, subsection 2.2.4, Principles for the Design of Multimedia Materials, presents the twelve principles proposed by Mayer (2009) to work as guidelines to be followed for the design of cognitively effective multimedia materials that may foster learning. The principles also ground the framework development for the analyses of the videos (to be explained in chapter 3).

2.1 VIDEOS IN THE L2 TEACHING AND LEARNING PROCESS

In the past years, the way information is distributed has changed. This change can be considered a result of the evolution of three sectors: information technology, media and telecommunications (RAJADELL; GARRIGA-GARZÓN, 2017). Accompanying the evolution of technology, especially in the digital formats, the forms of teaching have also been

changing in order to incorporate the emerging types of sources from which information can be inserted into the classrooms. With the development of new methods which not only focus on teaching but also on learning, videos can be a powerful learning source for having access to information from virtually any part of the world as well as working as a source of exposure to linguistic input, both inside and outside classrooms (BAJRAMI; ISMAILI, 2016).

Discussions concerning videos as a valuable source of input in pedagogical contexts have been conducted over the past decades. The evolution of technology (especially, digital technology) has led to the insertion of videos from different sources aiming at aiding learners in the process of understanding and making sense of information. At first, the discussions revolved around the use of cinema and television in the educational context (FERREIRA; JÚNIOR, 1986). The aspects of educational cinema, that is, videos produced by recording a scene or images with a video camera to be used in the classroom with a pedagogical purpose were discussed by Ferreira and Júnior (1986) in relation to how films could be incorporated in classrooms, and how their features could be explored in order to provide better learning opportunities. According to the authors, this medium (cinema – films) offered a number of special features that could be explored in order to foster learning, such as:

a) Slow motion, where learners' attention could be turned to specific information contained in the video;
b) Fast motion, showing in a short period of time a longer process;
c) Motion detention, that is, stopping (pausing) the video in order to attend to the topic of that frame;
d) Videos with acting, where there is the presence of actors performing the facts to be presented;
e) Cartoons or animations, where the use of drawings or professional cartoons is used to provide information;
f) Sound, where the visual information is combined with commentary or narration of what is being seen;
g) Color, which is considered of importance, depending on what to be shown, as for example to show the texture of an object.

The discussion of the use of videos as an aid to help learners also brings another important issue. Only using videos *per se* cannot guarantee that the contents or information presented will be understood by learners. Videos need to be used with pedagogical concern and preparation so that learners are able to best benefit from them (BULL; BELL, 2010). When learners' needs and the teaching objectives are taken into consideration, the probability of

success in video-based lessons may increase. When learners take the role of active watchers, they engage in a process of making sense of the information by activating their prior knowledge and connecting it to the contents worked in the classes. Moreover, instead of watching the videos passively, learners can watch them as part of their learning journey, becoming aware of the pedagogical purposes embedded in the audiovisual material (HAMMOND; LEE, 2010).

Videos can also be taken as a form of expressing knowledge and feelings. They can entertain and create a combination of the sensorial communication with the emotional engagement of the viewers so that learning can take place meaningfully (MORAN, 1995). Such engagement can also be associated to learners feeling motivated to be able to communicate in the L2, having a better job position, or promotion, or simply for the creation of a pleasant learning atmosphere with materials of their interest (DECI; RYAN, 1985). When learners are presented with the target language input with audiovisual sources, the information can be turned into a more realistic experience, in which they can observe the target language used in contextualized situations and possibly create connections between the video and real-life. These features can afford learners with precious input and work as a scaffold for the development of the language, when implemented, for example, in the L2 classroom.

The process of L2 learning is aided once videos provide learners with images and written and the spoken language[3] in a combined way. When the linguistic input is associated to the visual input, the connection between both can become facilitated, once learners engage in the process of information integration (MAYER, 2009), in which the imagery information complements the verbal information, aiding the learning process. For providing learners with imagery and verbal information, videos can have a highly important role in the context of L2 learning for the fact that they present complete conversational events in which not only can learners hear, but also see the target language in practice (LONERGAN, 1984; KAJDER; YOUNG, 2010). Such argument is grounded on the fact that learning the L2 is supported by the information entering learners' cognitive system by means of not only one sensory mode, but two simultaneously, in this case visual and aural, hence the concept of information integration (MAYER, 2009).

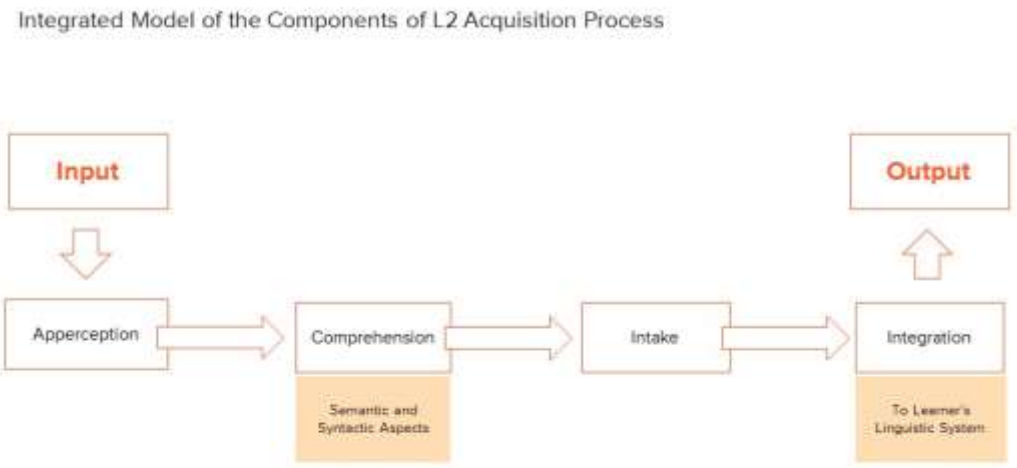
According to Lee and Mayer (2018), the use of different modes of presenting information on videos, images, spoken and written language can promote learning in a more efficient way. Seeing and hearing what is happening can support learners in making associations between the language heard and the visual actions being depicted and features such as the mood

[3] Although images are also known as language, in this Doctoral Dissertation the term *verbal language* will be used when referring to written or oral language.

of participants, their body language and ages, for example, which cannot be observed when simply listening to a conversation or information presented as audio only. It is possible to say, then, that when learners are exposed to videos in which the information is presented by means of visual aids and verbal language, the comprehension process is aided for the target information may complement each other (MAYER, 2009; SORDEN, 2005a, 2005b).

Concerning how the input provided can aid the process of L2 learning, Plass and Jones (2005) propose a model for the importance input can have and how it can aid learners in the process of improving their language production, that is, their output in any kind of multimedia material. In the case of this Doctoral Dissertation, the model proposed can be associated to information presented in videos as well. The model is presented in the following figure.

Figure 1 – Integrated Model For Input In The Process of L2 Learning



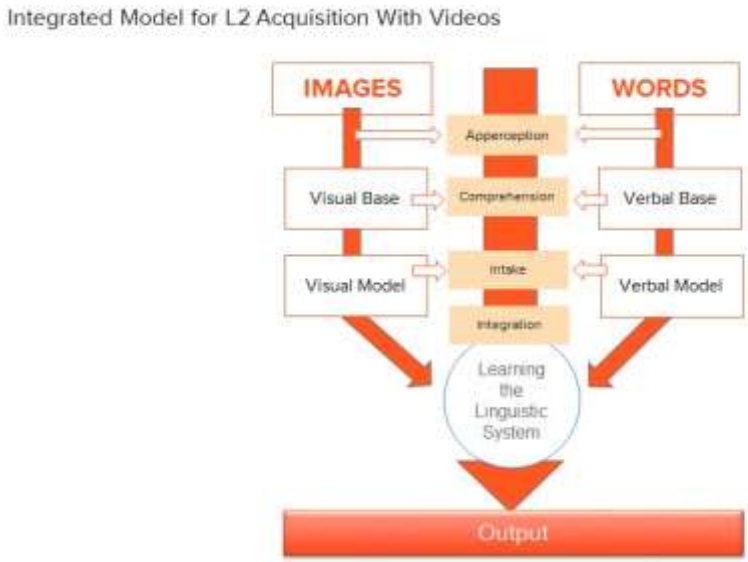
Source: The researcher, adapted from the diagram presented by Plass and Jones (2005)

As presented in Figure 1, videos can be a source of comprehensible input (PLASS; JONES, 2005) that is provided to learners in the form of imagery and verbal information. When exposed to the information on videos, learners engage in the apperception process in which the linguistic information is noticed (SCHMIDT, 1990; CHI, 2018), that is, when they become conscious (by means of noticing) and consciously select pieces of information they judge to be relevant based on their prior knowledge. Meaning is, then, attributed to the selected information concerning aspects associated to its semantic and syntactic meaning and structure. After the information is comprehended, that is, attributed with meaning, it becomes intake, which "refers to a process that mediates between the target input and the learner's internalized set of rules"

(GASS, 1988, p. 206), and is integrated to learner's linguistic system, becoming available in their repertoire for the production of output.

Acknowledging the elements that constitute the model of second language learning presented in Figure 1, and associating the concepts presented to videos, in which, as afore mentioned, may provide learners with input in form of images and oral words, Plass and Jones (2005) further propose an integrated model for L2 learning with multimedia materials. Although the researchers do not refer to videos specifically, they attribute the term Multimedia Materials to any types of educational materials that make "[…] use of words and pictures to provide meaningful input, facilitate meaningful interaction with the target language, and elicit meaningful output." (PLASS; JONES, 2005, p. 469), providing support for videos to be referred to as multimedia materials. Figure 2 presents a more detailed view on the processing of verbal and visual information and how these two types of information may be incorporated into the learner's linguistic repertoire.

Figure 2 – Integrated Model For L2 Learning With Videos



Source: The researcher, adapted from the diagram presented by Plass and Jones (2005)

As presented in Figure 2, once videos provide input in the forms of visual and verbal information, the combination of these can foster L2 learning. When words and images are consciously received by learners and the relevant visual and verbal pieces of information may be consciously selected (in the apperception process), a process of comprehension is activated, in which learners might build a verbal base for the verbal information and a visual base for the imagery information. The major aspect in this process is that both information bases (visual and verbal) are processed as complementing each other in order to aid learners in the process of

building coherent mental representations of the information being watched (MAYER, 2005a; 2005b; 2009) as a whole. Only information that is properly comprehended can be considered intake. In the intake process, learners may organize the visual and verbal information into verbal and visual models to be integrated to their linguistic system, which is, then, available for the production of linguistic output, either in the spoken or written forms.

Based on the assumption that videos may work as a powerful source of linguistic input, this input most times is presented in a contextualized way due to the fact that learners are presented with the information inserted in a certain context, or situation, so to make it easier to draw connections between what is being watched and the real world. Contextualization clues are hints that the language teacher, or video designer provides, in order to aid learners in the process of understanding a word or concept that may not familiar to them (INNACI; SAM, 2017; TUYEN; HUYEN, 2019). In addition to providing learners with aids to understand the information provided, the contextualization clues can also indicate the position of it within a sentence, in the case of a vocabulary item (AL-SEGHAYER, 2001). When understanding the meaning of words by the use of contextual clues, learning can take place not only concerning the lexical items themselves, but also the message as a whole as part of communication.

For the fact that learners are provided with input in the forms of imagery and verbal information, as mentioned before, such clues can be organized in the forms of oral and/or written language accompanied with images used to integrate the verbal information being presented. In the case, for example, of videos that aim at presenting grammatical explanations or examples of specific target language structures, contextualization clues can be associated to the written language being presented and explained. According to Tuyen and Huyen (2019) contextual clues can have an important effect on the process of learning new vocabulary items, once they provide learners with extra help in understanding and, at times, inferring the meaning of certain lexical items that may not be understood when considered out of context. Once videos can be taken as an environment in which communication is being made, that is, videos are a means through which a message is being transmitted, for the communication to take place effectively, viewers need to be able to attribute meaning to what is being read and heard. This concept leads to an important factor that is associated to learning an L2, and that with working with videos this concept should not be different. The information needs to be understood by the ones who are watching and listening to it. Putting it into the field of SLA (second language acquisition), videos should present learners with comprehensible input (HOQUE, 2017) so that learning can be fostered and the processing of the incoming information aided.

For the target L2 language to be learned, and for learners to move from their stage to the next level, they must be presented with the input that can already be familiar for them, added to a new information item, represented by Krashen (1981, 1982) as i+1 in his Input Hypothesis Theory. Such concept suggests that when learners are presented with a linguistic piece of information in the form of input that they are familiar with (the i), they can activate their prior knowledge (strings of knowledge already stored about that information in their long-term memory). It appears that the crucial point is concerning the extra information related to the next level, represented by the +1, intending to push learners into making sense of the new information based on their prior knowledge, increasing their strings of knowledge in long-term memory (LTM), providing them with a sense of unfolding in the learning process (KRASHEN, 1981; 1982). Based on the importance that contextual clues and input have for the learning an L2, it is possible, thus, to establish a connection between the input hypothesis theory and contextual clues. When considering that the +1 part of the input alone would be considered too complex for learners to comprehend, the i would have the role of the contextual clue, meaning that the i would be essential for the comprehension of the +1.

Concerning the development of learners' communicative competence, videos can also support the improvement of their listening skills and support the speaking production. Similarly to the process of learning an L1 – in which infants need to be exposed to oral input – the process of learning an L2 relies on the exposure to the oral language for the development of listening and speaking skills. As part of the input (KRASHEN, 1981, 1982) learners are exposed to in the process of learning, listening provides them with information about the target language, and messages that are being communicated, thus, making oral comprehension a key element to language development. According to Woottipong (2014), listening is a goal-oriented activity that individuals engage in, that is, individuals listen to an oral text in order to grasp meaning from what is being heard. In L2 educational settings, when learners are provided with oral input, they use a set of cognitive processes in order to organize the oral language being heard into meaningful messages to be understood, including lower-level processes, such as decoding, and higher-level processes, including inferential processes. Regarding speaking, videos can work as a source of linguistic input in which the information provided can serve as examples of the target language to be used in conversation tasks, for example.

A further important aspect of the development of listening skills is that, when learners are prepared, or at least aware of what they will listen to before listening to it, or even know what pieces of information they will be expected to obtain (understand) from the passage, the success of comprehending what is being heard (or at least a major part of it) can be achieved.

This is referred by Dunkel (1986) as script competence, in which the listener (and speakers, depending of the kind of task) knows in advance the topic about which they will be listening to. Knowing what the topic of the listening passage will be beforehand, and having a clear objective for listening or watching (in the case of videos), can be of importance once the listeners' prior knowledge can be activated as a preparation for the listening experience.

According to Schmidt (1990) and Ardeshir (2015), when learners receive certain stimuli, they activate knowledge they already have stored in their memory concerning that topic, and associate it to the new information in the text in order to make sense of it. In the case of learning from videos, learners' prior knowledge can help them in building a picture of the information being heard and seen and integrate it to their already existing schemas (DUNKEL, 1986), which can also be associated to the construction of coherent mental representation of the information (MAYER, 2005a; 2005b; 2009). The activation of learners' prior knowledge is grounded on the schema theory. According to Ajideh (2003), a schema is a set of mental strings of knowledge that are stored in long-term memory (LTM). Schemata are naturally built as a result of experiences, interactions or events that people experience in their everyday lives and have stored in their long-term memories structured as blocks hierarchically organized in superordinate and subordinate levels. For example, the word bathtub can evoke the superordinate schema bathroom, which, in turn, can also evoke the superordinate schema house. It can also evoke the subordinate schemata associated to the actions of bathing, brushing the teeth, washing the hair, as well as the words soap, towel, shampoo and toothpaste.

They can also be referred to as the organized background knowledge (AJIDEH, 2003) that is activated by means of a certain stimuli (AN, 2013). Ajideh (2003) further explains that the stimuli that can activate and help the process of schemata building can be of two kinds: the first kind is concerned with presenting learners with external information which, in turn, can be processed with already existing information (schemata), resulting in an expansion of the knowledge. The second kind is associated to building new schemata by means of the representations of new mental strings of information based on the schemata already existing. The basic difference between the two kinds of schema activation and building is that in the former learners receive external input, that is, from the world in the forms of experiences and interactions resulting in the expansion of their schemata, and in the latter, learners expand their schemata by the (re)processing the schemata they already possess. In the educational context, stimuli may be presented in the form of imagery and verbal content input whose objective is to provoke the retrieving of schemata already stored in LTM, so that they can serve as a basis for the building of new strings of knowledge, or the enlargement of the already existing knowledge.

In the case of learning an L2 with videos, the input provided in the videos may work as the stimuli.

The amount of information provided in videos as to their adequacy to learners' language competence is another issue to be addressed. Perhaps one of the most important feature online videos can offer is the control of the material and flow of information presented, as discussed by Benkada and Moccozet (2017), with the possibility of pausing and playing the video. Such feature can help learners once the information flow can be interrupted when learners feel the necessity, or even go back to a certain point in the video and re-watch for further understanding. An example of this is the study conducted by Hasler, Kersten and Sweller (2007), in which the researchers report on a piece of research conducted with young learners providing them with a video animation about the day and the night. The participants were organized in four different treatment groups: the first group watched the video in a system-paced fashion; the second group watched the video in a learners-paced fashion, organized in segments (blocks); the third group watched the video, having available play and stop buttons; and the fourth group was presented only with the listening passage. The results obtained by the researchers after applying tests to participants suggest that the two groups that were exposed either to a video organized in segments, or with the possibility of stopping and continuing the video, that is, the two groups that received the learner-paced treatment, performed better. The results suggest that when learners have the control, or even when the video materials follow a learner-paced fashion, learning can be enhanced.

In order to make the oral input more comprehensible and motivate learners into trying to grasp as much meaning as possible from the video, the use of subtitles[4] can result in a helpful source in the learning process. According to Pujadas and Muñoz (2019), "L2 audio-visual materials enhanced with subtitles or captions are robust tools for second language learning as learners are exposed to a large amount of input simultaneously through image, text and sound." (PUJADAS; MUÑOZ, 2019, p. 481). In addition, with the repeated exposure to materials containing subtitles, learners may develop a more effective processing of the written language (TALAVÁN, 2010) and, as a result, improve language learning by means of the integration of information that is presented in both visual and oral formats in a complementing way.

The use of subtitles also brings into discussion the fact that when learners are exposed to videos presenting information by means of images and in both oral and written forms, that

[4] Considering that there is the difference between subtitles and captions, in which the first (subtitles) are used in a language different from the one presented in the material, and the second (captions) are presented in the same language as the material. In this Doctoral Dissertation, the term *subtitles* will be used interchangeably.

is, in the spoken form accompanied by the subtitles transcribing what is being said, the two kinds of information (visual and oral) are processed and mentally integrated. The oral linguistic information (spoken words) and the visual linguistic information (written words and images) that enter the two channels – ears and eyes – as input, is processed and organized in a form complementing each other and aiding the building of mental representations of the information received (MAYER, 2009; MAYER; SIMS, 1994; PAIVIO, 1990) resulting as better learning.

The insertion of subtitles in videos can contribute for listeners to better understand the target language. The presentation of the written information helps listeners in making connection with the spoken language, thus complementing what they are reading to what they are listening and seeing (MITTERER; MCQUEEN, 2009). The use of printed words – subtitles – accompanying the spoken text can work an extra source to provide additional information about the language being spoken, which can also be referred to as lexically-guided learning, which may work as a resource to aid learners in the process of comprehension when the spoken language provided is complex for them to follow. The subtitle, then, has the function of integrating the written language to the oral language so that comprehension may be granted.

When subtitles are inserted in video materials, they can fall into two different categories: *interlingual* and *intralingual* subtitles (LAVAUR; BAIRSTOW, 2011; DANAN, 2004). The *interlingual* subtitles is the use of the written language inserted in videos that is different from the spoken language from the material. An example could be an instructional video in which the spoken language is English and the language of subtitles is Portuguese (viewers' L1). The *intralingual* subtitles, also referred to as *Same Language Subtitles* (KRUGER, 2013), present the subtitles in the same language that is spoken in the material. An example could be a video in which the spoken language and the subtitles are both presented in English, that is, both in the L2.

At times, comprehending information with listening-only tasks may result in an overwhelming task, especially for beginner learners because their language knowledge can be limited, and the provision of input only in the format of images might not be sufficient for learners' fully understanding and integration of the spoken words. Danan (2004) suggests that a strategy that may help in making incomprehensible input comprehensible is the use of interlingual subtitles because when exposed to oral input that can be too complex for learners, the use of subtitles in their L1 can work as a facilitator for comprehension of the message being transmitted. Opposing to this idea, Mitterer and McQueen (2009) argue that when materials are designed with the use of subtitles, the presentation of the written language should follow an intralingual fashion, that is, subtitles should be presented in the target language (learners' L2)

as well so to present viewers the words being spoken. The choice of whether to use inter- or intra- lingual subtitles can be strongly influenced by the level of language knowledge learners have. For example, for beginner learners, the use of interlingual subtitles might be helpful in order not to overload their attention and processing of the amount of information in the L2. For more advanced learners, on the contrary, as having a larger amount of language schemata, the provision of intralingual subtitles may be beneficial, once deeper connections between the language being heard with the language being read can be made, resulting in learning the target language (MITTERER; MCQUEEN, 2009).

When new vocabulary items to be learned are presented, different strategies can be adopted by learners, as for example using lists, or making associations between the target words and their translations (KIM; GILMAN, 2008). A number of studies have investigated the effect subtitles can have on the process of learning an L2. Farías, Obilinovic, Orrego and Gregersen (2013) conducted a study whose objective was to evaluate the effects of presentation modality in retention and transfer of ESL vocabulary. The researchers tested university participants who watched a multimodal PowerPoint presentation focusing on the acquisition of vocabulary. The participants were assigned to three different treatment groups: the first group watched a presentation containing on-screen text and narration; the second group watched a presentation containing on-screen text, narration and still images; and the third group watched a presentation containing on-screen text, narration and video. The results showed that the on-screen text, narration and videos showed to be more efficient either in retention and transfer of the new vocabulary presented when compared to the groups who were provided with still images, or no images in the multimodal presentations. This means that learning can be facilitated when learners are provided with subtitles to complement the oral language and the imagery input.

Perez, Noorgate and Desmet (2013) conducted a meta-analysis of studies that reported on the use of subtitles for the development of listening comprehension and vocabulary acquisition. The studies reported on the use of intralanguage subtitles, that is, presenting the video with the subtitles in the same language spoken on it, interlanguage subtitles, in which the subtitles presented where different for the language spoken on the video, that is, the subtitles were in the participants' L1, and videos with no subtitles. The results of the analysis found by researchers suggested that the use of intra- and interlingual subtitles can be effective for the development of listening comprehension and for vocabulary acquisition in the context of L2 learning depending on learners proficiency. Moreover, grounded on the Dual Coding theory brought by Paivio (PEREZ; NOORGATE; DESMET, 2013; CLARK; PAIVIO, 1991) the authors also provide a connection between the use of subtitles and the way information is

processed, in which “the presentation of an item in two modes (aural and written) stimulates the verbal and imagery system. This is expected to improve information processing, subsequently leading to greater depth of processing and better recall.” (PEREZ; NOORGATE; DESMET, 2013, 2013, p. 722). Such assumption offers the notion that when learners are provided with the target language as oral and written input, they may integrate each other and learning may be fostered.

Concerning the issue of intra- and inter- language subtitles, a study conducted by Hayati and Mohmedi (2010), investigated whether the use of films with and without subtitles could be effective to develop listening comprehension skills. The researchers collected data with participants who had Persian as L1 and were learning English as an L2. The participants were divided in three groups who watched six episodes of 34 minutes each of a DVD entitled “Wild Weather” each. All the three groups watched the same episodes, but each one received a different video treatment: the first group watched the videos with the spoken language in English and with English subtitles; the second group watched the video with the spoken language in English, but with Persian subtitles; and the third group watched the video with spoken language in English with no subtitles. The results obtained with this study were that participants could remember more items of the passage when they were provided with English subtitles, that is, intralingual mode. The study suggests that the use of subtitles, particularly intralingual, when presenting learners a video or a film may work as an aid to facilitate learners in the process of comprehension, and thus, learning the target language; once they are being exposed to the target language, connections between the spoken and written language can be easily made.

The research studies presented above present evidence that videos can result in a powerful source of input for the acquisition of an L2. Not only do videos present information relevant to the topics being studied and worked in the L2 classroom, but also, they provide valuable forms of linguistic input. When learners are provided with input in more than only one mode, for example, instead of being offered only with the information in the oral mode, they are provided with videos that present information making use of sounds, images and text in their design, more complex mind processes activated in the decoding of the information, resulting in a more efficient learning.

Once videos can work as a rich resource for learning an L2 (LONERGAN, 1984; PLASS *et al.*, 1998; PLASS; JONES, 2005; TUMOLO, 2017), there is the need for a closer look into how their design can impact on the way information is processed, and to observe the structural design as well as the principles that researchers have proposed. In the next section,

videos will be discussed as a multimedia construct and how learning can be affected by their design, including the features and principles proposed by Mayer (2009) that can be considered in the design of educational materials.

2.2 MULTIMEDIA AND PEDAGOGY: VIDEOS & LEARNING

This section presents reflections concerning how videos can be characterized as multimedia materials and aid learners in the process of learning English as an L2. For this objective to be achieved, firstly, subsection 2.2.1 presents definitions brought by Mayer (2009) and other scholars regarding instructional multimedia materials aimed at promoting learning, followed by arguments defending that videos can fall into the multimedia materials category.

The following subsection, 2.2.2, presents discussions concerning how educational multimedia materials, in this case, videos, can aid the process of knowledge building, grounded on the Cognitive Theory of Multimedia Learning (CTML), regarding how the human mind processes information and organizes it into knowledge by means of providing learners with input in the forms of visual and oral information.

In subsection 2.2.3, discussion is provided concerning the limitations of working memory and how educational materials, especially to aid learners in learning an L2, should be considered in their design so that learners' processing capacity is not exceeded, facilitating the process of learning.

Finally, subsection 2.2.4 presents the twelve principles proposed by Mayer (2009) for the design of multimedia materials and narrows the discussions down to the three principles of Signaling, Segmenting and Coherence.

2.2.1 Video as a Multimedia Resource

With the development of technology, the way information is delivered has changed, moving from words alone (provided by the teacher, or in books), to the use of countless resources involving computer graphics, images, photos and videos. This can be especially observed in the educational context, in which digital technology has been offering new ways to deliver information with different resources. With such, teachers and instructors have the opportunity to share information by using a range of visual aids that can complement the verbal input offered. Based on the notion that technology should be included in educational contexts, it is important to bring into discussion how videos fall into the category of multimedia materials

as well as the discussions presented in the previous section, concerned with a number of benefits that videos can offer for the learning of an L2, in this case English. By having presented a more general view of the features videos can afford to the L2 classroom, the discussions that follow will narrow down to what effects the design of videos can have on cognition, and ultimately on learning.

Firstly, it is important to define multimedia. Multimedia is the term used to define the use of words and images together to promote learning (MAYER, 2005a; 2009). This use can be combined in a computer-based learning material, an animated encyclopedia with instructions on certain topics, or a live presentation where there is a slideshow accompanied by a presenter speaking. According to Mayer (2005a), multimedia materials are a type of presentation that necessarily involves the use of both images and words to promote learning. In relation to images, the author refers to the use of photos, graphs or animations, and in relation to words, the author refers to presenting information in the form of written or spoken language.

The term multimedia also brings two other terms, which are concerned with the use of words and images in order to promote learning: multimedia learning and multimedia instruction. Multimedia learning is associated to the fact that learners build mental representations of the information presented in the multimedia material used, and multimedia instruction to the way words and images are presented in order to promote learning, thus, being the former focused on learners and the processes that involve learning, and the latter on materials design (ZHENG, 2009). Based on the definition presented, a video can also be defined as a multimedia material[5], for it may present information using the words spoken by a narrator or a character, along with images, graphs and other kinds of visual information. When videos are designed to meet educational objectives, based on Mayer (2009), three important characteristics should be considered.

a) The delivery media – videos need at least two delivery devices, as for example, a computer screen and the speakers, a television and the speakers, or a projector and speakers;
b) The presentation mode – videos are intended to present information using verbal and pictorial information, that is, in form of spoken or written words and images or

[5] Although the discussion in this section is about multimedia materials, once the objective of this Doctoral Dissertation concerns specifically with videos, the terms *videos* and *educational videos* will be used interchangeably when referring to multimedia materials.

motion pictures (also associated to Paivio's (1986) dual-coding theory – to be discussed in the next section); and

c) The sensory modalities – videos present information using both aural[6] and visual senses for the information to be grasped.

When learning is treated as a process of knowledge construction, the view upon videos changes. Once, as afore discussed, videos present formation by providing learners with a combination of images and words, learners are not considered merely passive watchers. Instead, they become active in the learning process, that is, learners actively engage in the process of making sense of the information being watched by consciously selecting the parts they consider to be relevant for their understanding. According to the Cognitive Theory of Multimedia Learning (CTML), proposed by Mayer (2005a; 2009), learning is the result of the mental representations that learners actively build based on the images and words they watch on videos. There are, however, issues concerned with how digital technology and, more specifically, videos, are designed and used in the educational context to better aid learners in the process of knowledge construction. Once videos can be defended as being complete environments that can aid learning for presenting imagery and verbal information, an important issue that arises with the deeper understanding of how the human mind processes information is to formulate principles for the design of multimedia materials so that learning can be better aided, thus, being the design of eduational videos learner-centered, rather than technology-centered.

The basic difference between the two approaches to educational videos is that technology-centered materials consider mostly how information could be incorporated into new communication apparatuses, as for example, the World Wide Web, the use of virtual realities, or even the use of materials in the digital formats. The dilemma with this approach may lie on who, or in this case what, is the focus being given to, arising the concern that how technology can be better accessed and incorporated seems to be more important than providing learners with videos that can actually aid the process of knowledge construction. A learner-centered approach (MAYER, 2009), on the other hand, takes the opposite pathway when considering and designing educational videos. The promise of this approach is that educational materials – in this case, videos – should be adapted following the human cognitive architecture. When videos are in consonance with the way the human mind processes information, learning can be fostered and the construction of knowledge is more likely to be successful.

[6] *Aural* refers to any kinds of sounds and is associated, more specifically, to the sense of hearing, whereas *oral* is associated to the mouth, that is, to speaking.

Considering learning to be the main objective when educational videos are designed, Mayer (2005a) discusses different concepts that can be attributed to the results obtained by presenting three metaphors for using such materials. The first, Response Strengthening, concerns with learners performing a response to a certain stimulus received. When the response is correct (or corresponding to the expected), learners receive a positive feedback from teachers or instructors – also referred to as a reward. When the response is not correct, then they receive a feedback aiming at orienting them towards the correct response – also referred to as punishment. In this metaphor, learners create connections between the information provided in the videos and their responses (or language to be used), and the feedback received. In this case, videos can work as a training material to practice certain skills. The second metaphor, Information Acquisition, concerns with incorporating information to learners' memory, being learning based on the information being presented and watched, and ultimately transferring it from the computer or television screen to the memory. This assumption is analogous to the empty vessel assumption, in which learners' minds are empty and need to be filled with the information to be learned. The third, and perhaps the most important, Knowledge Construction metaphor defines that learners engaging in a complex task of making sense of the information being presented and watched. For knowledge to be constructed in a meaningful and effective way, learners' personal knowledge of the world (prior knowledge, or schemata, already stored in long-term memory (LTM) (KALYUGA, 2010) plays a crucial role. This can provide insights concerning the pedagogical objectives that are attributed to videos in the educational context, and how they can work as activating learners' prior knowledge and aiding the process of knowledge construction based on the input received.

Acknowledging the fact that educational videos are aimed at promoting learning, three outcomes can result from providing learners with materials that present the target information by using images and words (MAYER, 2005a). The first, No Learning, as the name itself suggests, happens where there is no change in learners' memory concerning the information watched and it means that no new strings of knowledge were constructed and incorporated to their LTM. In the second, Rote Learning, learning is limited only to some fragments of information. In rote learning, learners can remember some incomplete pieces of information, but do not have enough schemata to use the information retained in different contexts, meaning that retention can happen, but transfer cannot take place, once the fragmentation of information, thus knowledge, is incomplete to be applied to different contexts. The third, Meaningful Learning, is when actual learning takes place, when learners construct knowledge based on the mental representations built, and can transfer the knowledge acquired and stored in their LTM

to other contexts. Meaningful learning can be associated to the third knowledge construction, presented above.

For effective learning to take place, Mayer (2009), and Mayer and Moreno (2010) present three principles that are based on cognitive theories for multimedia materials: The first principle, Dual Channels, explains that individuals have two separate channels for processing the oral information (auditory) and the imagery information (visual). The second principle, Limited Capacity, explains that individuals can only process a rather limited amount of information in each channel each time. The third principle, Active Processing, explains that for effective learning to take place, learners need to engage in a series of processes that include selecting information they consider relevant, organizing these pieces of information into coherent blocks of mental representations, and integrating the new information to already existing information in their LTM. When the choice of educational videos is grounded on their design considering how information enters their cognitive system, the limiations and processes involved in turning information into knowledge, learning can be aided and the construction of knowledge facilitated.

Hence, it can be defended that when some sort of information is not only heard, but actually heard and seen, learning may be fostered, once the information is processed by means of the building of coherent mental representations in learners' working memory, and their integration into LTM in the form of new, and/or expanded schemas (AYRES; PAAS, 2007).

Grounded on the notion that learning is aided when learners are provided with videos that present information using the combination of words and images, it becomes important to provide a panorama of how information is processed, considering their cognitive architecture. The next section will present discussions concerning how information is processed by the human cognitive system in order to be turned into learning.

2.2.2 The Relation Between Educational Videos and Cognition

As mentioned before, this section will present discussions on how information is processed by the human mind, and how videos as multimedia materials can impact the way the information presented is processed, and ultimately affect learning.

In the educational context, when words and images are used together in a digital format with a pedagogical purpose, this combination is named multimedia learning (MAYER, 2009), as presented in the previous subsection. Mayer (2009) claims that with the use of different modes of presenting information (for example, images, spoken and written language), learning

can be promoted in a more efficient way. Lonergan (1984) adds that when learners can experience new information by seeing and hearing what is happening, they can draw connections between the language and the actions and observe features such as the mood of participants and their body language which cannot be observed when simply listening to a conversation or information. It is possible to say, then, that when learners are exposed to educational videos, composed by images and words, the information received is not only presented twice, but in manner complementing each other (MAYER, 2009; SORDEN, 2005a, 2005b).

For educational videos to aid knowledge construction, there is the need for their design to consider the structure of human cognitive architecture (SWELLER, 2010; KALYUGA, 2010; 2015). Such notion means that materials should be designed taking into consideration the human information processing system and its organization. Considering the structure of human cognitive architecture associated to the biologically secondary knowledge[7], Sweller (2017) discusses five principles associated to these structures and the processes that constitute cognition. The five principles are explained as follows.

The Information-Store Principle establishes that in order to live and deal with any types of information and input, individuals receive throughout their lives large amounts of information needed to be stored in the forms of schemas in their long-term memory (LTM). In an L2 educational context, a practical example can be explained as one individual being more proficient in one topic than another individual as a result of the amount of schemas stored in LTM about that topic.

According to the Borrowing And Reorganizing Principle, information that is processed and organized in LTM in the forms of schemas is constructed by means of listening to and reading what others have said. When new information enters learners' cognitive system in the form of input, it is integrated to information already held in LTM and new schemas are constructed, or the already existing are expanded.

The Randomness-As-Genesis Principle establishes that when learning, individuals tend to test their hypothesis as to whether they are correct, or incorrect. Such principle is associated to the fact that, when no information can be borrowed from others, individuals need to try to

[7] According to Sweller (2017), the human cognitive architecture that deals with knowledge construction is organized in two categories: 1. The Biologically Primary Knowledge is the knowledge that individuals have acquired (with no conscious effort) through the generations, as is the case of an L1. 2. The Biologically Secondary Knowledge is the knowledge that individuals acquire (with conscious effort) by explicit instructions, as is the case of learning an L2 in an educational context.

generate it themselves. The issue about this principle is the fact that the results are often imprecise and imposes higher levels of cognitive load on learners.

Perhaps the most important, the Narrow Limits Of Change Principle, establishes that the limits of processing of new incoming information is immensely limited. The mind can only process three to four elements at a time for about fifteen to twenty seconds. In the case of learning an L2, it is possible to argue that learners are most of times dealing with new incoming information, imposing high cognitive load to them. The case for educational materials is that they be designed in a fashion that decreases the cognitive demands and frees up working memory (WM) resources.

Finally, the Environmental Organizing And Linking Principle establishes that, differently from WM whose processing limit is extremely reduced, there is no known limit for the amount of information that can be stored in LTM. Furthermore, the limitless capacity of processing is also associated to the processing of information that is already stored, that is, the larger the amount of information learners have stored, the less the cognitive load materials will impose on them for the processing of information. In the case of learning an L2, the more information about it learners have stored in LTM, the more proficient they will be in using the target language.

It becomes, then, possible to associate the five principles presented by Sweller (2017) to the context of learning with videos. In an applied view, Mayer (2005b) argues that when learners engage in learning from any multimedia material, in this case, videos, there are five steps, each one, leading to the next in which they engage in in achieving the educational objectives. They are 1. Selecting the relevant words (text or narration) from the material; 2. Selecting relevant images from the material; 3. Organizing the selected words in a coherent verbal representation; 4. Organizing the selected images in a coherent pictorial representation; and 5. Integrating the verbal and pictorial information to learners' schemata. The five steps take place in learners' cognitive system, organized in three major elements. These elements are sensory memory (SM), working memory (WM) and long-term memory (LTM) (SORDEN, 2005b; SWELLER, 2010; KALYUGA, 2010; 2015; BRAME, 2016).

The first element that works for external information to enter learners' cognitive system is the semsory memory. Although there are different sensory receptors through which external information can enter individuals' cognitive system (*e.g.* tactile, olfactory or gustatory), in the case of learning with videos, only two sensory receptors are considered: the eyes and the ears (MAYER, 2005b; SCHNOTZ, 2005; MAYER, 2009), that is, sight and hearing (considering videos providing information in the forms of images and sounds). Receiving information by

means of the visual and auditory channels is characterized as the dual-channel assumption. The dual-channel assumption defines that humans' cognitive system is organized by having an auditory/verbal channel, and a visual/pictorial channel through which the information enters the cognitive system as aural and visual input (SCHNOTZ, 2005). Information that is presented in the form of images, photos and graphs is received by the visual channel, and information that is presented in the form of speech or narration is received by the auditory (or aural) channel.

The key characteristic of the two channels is that each has limited capacity of receiving and holding information, being limited only to a small amount of information each time. Information that enters the visual channel is stored for less than one second. However, when attention is directed to that piece of information, it is, then, conducted to the visual WM where it is further processed in the form of mental representations. Information that enters the auditory channel is stored for less than three seconds. As with the visual channel, when attention is directed to the information, it is conducted to the auditory WM for further processing.

According to Sweller (2010), WM is responsible for the selection, processing and integration of information received. WM is also associated to the active processing assumption. Following this assumption, it is possible to argue that when learners watch a video aimed at learning a certain topic, they engage in cognitive processing in order to build coherent mental representations of the situations that they experience throughout their lives as well as the information that they are provided with in educational settings. These processes are organized in a series of stages working in a way that can complement each other. These are paying attention, organizing information that enters the cognitive system and integrating the new information to already existing information (SWELLER, 2010).

In short, active processing can be defined as individuals making sense of the information they receive as input in a process of building coherent mental representations of those instead of only receiving information passively. When WM is at the service of active processing and active learning, three major processes can take place: in the first, Selecting Relevant Material, learners consciously focus their attention (SCHMIDT, 1990; LORCH, 1989; KONING *et al.*, 2010) to specific parts or pieces of information (words and images) presented in the material. It is also concerned with bringing outside materials and information into the cognitive system WM. In the second, Organizing Selected Material, learners build relations between the information elements in WM, that is, they engage in making sense of the information by attributing coherent meaning to them. In the third, Integrating Selected Material with schemata, learners create connections between the novel information that is being processed in WM and integrate it with their schemata, already stored in their LTM. The issue, however, is that WM

has two limitations: the first is concerned with the amount of elements that can be processed each time – considerably limited concerning the novel information that enters in form of input, and the second is that the time span for the processing is extremely short. It means that, especially for learners who lack prior knowledge on the topics being watched, WM limitations may hurt learning.

After information is processed in WM, it may be stored in LTM in the form of knowledge. The main aspect that LTM presents is that, unlike WM, LTM seems to have no limited capacity. Information is stored in LTM in the form of strings of coherent information that are available to be accessed at any moment when needed, and can be referred to as prior knowledge or schemata. Learning can, thus, be considered an alteration in LTM (SWELLER, 2010). So it becomes possible to argue that when there is no change in LTM, then no learning can be considered to have happened.

When learners are provided with videos aiming at promoting learning, for comprehension to effectively take place, there is the need for learners to already have some degree of schemata stored in their LTM, as previously discussed. When learners engage in the processing of new incoming information in their WM, already stored schemata is also retrieved back into WM so that the new information is processed along with the prior knowledge in an interactive way. The ultimate importance of the schemata stored in LTM is that it can directly influence how the incoming input is processed and turned into new strings of knowledge (SCHNOTZ, 2005). In short, the larger the schemata learners have, the lower the degree of complexity attributed to the information watched, hence the importance of aiding learners in the process of knowledge construction.

With a deeper understanding how individuals' cognitive architecture is structured, new approaches to materials can (and should) be taken. This is the case of videos designed based on the Cognitive Theory of Multimedia Learning, that is, videos that are designed in order to explore the potential of using words and images together to promote learning (MAYER, 2009) taking into consideration how individuals process the information. The diagrams presented as follows represent the different ways images, written words, and spoken words are processed by the human cognitive system. The three ways information is processed are depicted in each of the diagrams explained as follows.

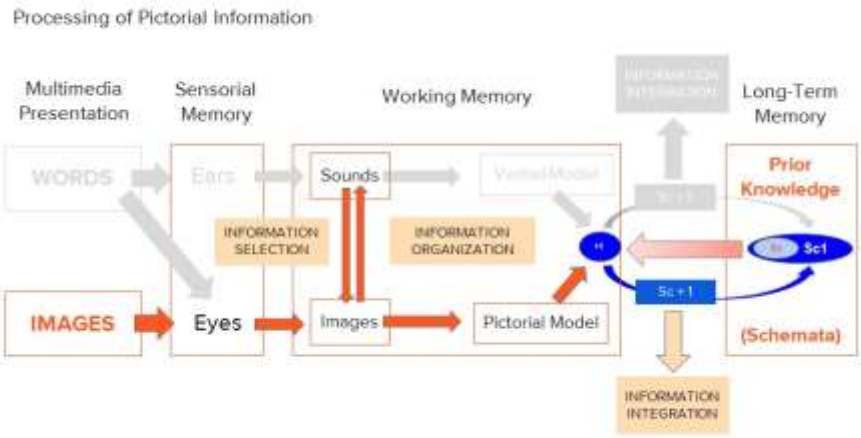
As can be observed in the Figures 3, 4 and 5, the dual coding assumption (PAIVIO, 1986; MAYER; SIMS, 1994; MAYER, 2009) assumes that information enters learners' cognitive system through two main channels: the eyes and the ears, respectively. Once information has entered, it is processed in the working memory (WM), where the sounds are

converted into a verbal model of information and images are converted into a pictorial model. It is also important to observe, as is represented in the three diagrams to be presented, that even though sounds and images enter two different channels, they are processed in a way integrating each other throughout a process in which coherent mental representations of the information presented are built. Then, prior knowledge stored in the LTM is activated and retrieved into WM so that it is processed with the new information in order to construct new strings of knowledge to be stored back into LTM. In this sense, the knowledge already stored in LTM, that is, the schemas learners have stored are represented in the diagrams as Sc in the light blue circle, and the new incoming information provided in the video is represented by the +1. In the learning process, the schemas (Sc) learners already have are retrieved back into WM to be (re)processed with the new information (+1), indicated by the arrow going from LTM to WM. The processing and integration of the new information to learners' schematic knowledge is represented as the Sc+1 and results in the construction of new strings of knowledge, which are re-stored back into LTM in the form of expanded schemas, represented by the bigger dark blue circle with the Sc1.The way the three different kinds of information (pictorial, verbal written and verbal oral) are processed in humans cognitive system are discussed as follows.

2.2.2.1 The Processing of Pictorial Information

Pictorial information concerns all the visual input that is presented to learners in the form of imagery information. The diagram below depicts how pictorial information is processed in learners' cognitive system.

Figure 3 – How the Processing of Pictorial Information Takes Place



Source: The researcher, adapted from the diagram presented by Mayer (2005a)

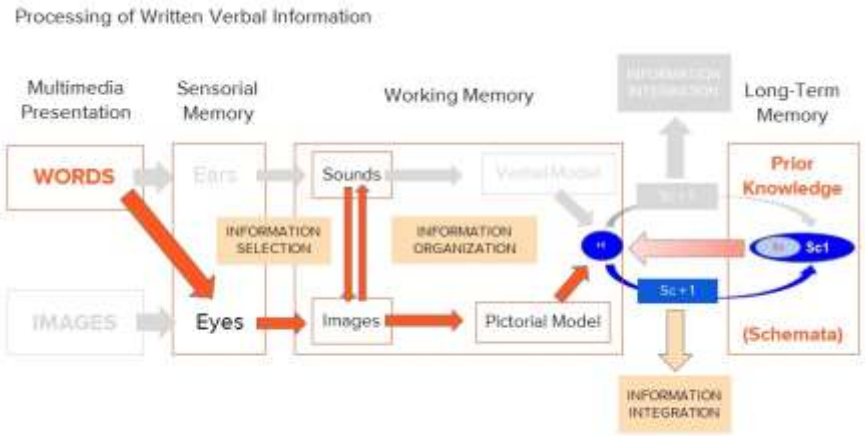
The initial element concerns with the input in the form of images that are presented on a video and enter the cognitive system through the visual channel, that is, the eyes. This first process occurs automatically on the part of learners, who do not need to invest a big amount of effort to see the information. After information has entered the visual channel, the processing stage starts. An important fact of this stage is that learners need to consciously[8] select the images they find relevant to be processed in WM in order to build coherent mental representations of the information being watched. It is important to observe that, when the visual information is processed in WM, there is also one arrow pointing to sounds and other pointing back to images, indicating that, when learners see the image of a tree, for example, they create a mental representation of the image, and the written form and sound of the word tree. After the images are converted into words, they are organized into a pictorial model of the imagery input received – a tree. In the processing that takes place within WM, already stored information from LTM, that is, learners' prior knowledge in the form of schemata, also plays a crucial role. The schemas (Sc) learners already have stored in their LTM are retrieved back into WM to be (re)processed with the pictorial mental representation built based on the new visual information being presented (+1). After these pieces of information are processed together and integrated (Sc+1), they are stored back to LTM in the form of expanded schemas (Sc1), resulting in learning. In addition to the imagery information, linguistic input can also enter lerners' cognitive system through the eyes, as explained bellow.

2.2.2.2 The Processing of Written Verbal Information

Written verbal information is the input presented to learners in the form of written text. The processing of written words in a video develops in a manner similar to the processing of the pictorial information due to the fact that written words are also conducted to learners' cognitive system through the visual channel, as is presented in the following diagram.

[8] *Consciously* refers to learners directing their attentional resources of pieces of information they judge to be important for the understanding of the input they received as they watch the videos and, consequently, aid the process of building coherent mental representation of them.

Figure 4 – How the Processing of Written Verbal Information Takes Place



Source: The researcher, adapted from the diagram presented by Mayer (2005a)

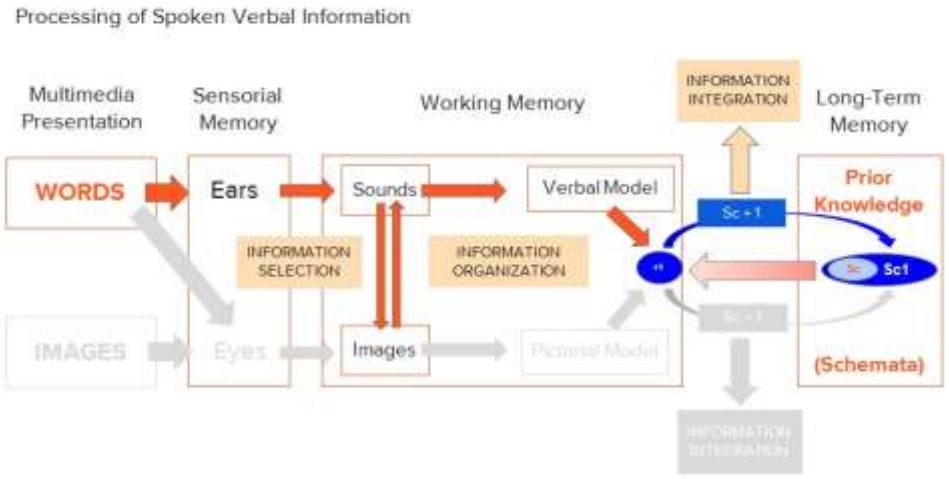
In this case, when learners are provided with written text, the words are conduced to their WM through the eyes. After seeing the written text, learners also need to consciously select the parts they consider relevant, that is, the relevant words to be processed. Once it is possible to argue that a written text is the visual representation of a word, the words selected are firstly processed as the images of the words, that is, the typographical representation of them. As can be observed, there is one arrow that points to sounds, and other arrow pointing back to images. When written words are processed in WM, connections between the typographical representations and the sounds of the words are also built, and the image of the written word is also converted into the image of the object represented by the word. For example, when learners see the word cat, they automatically convert the letters into the sound of the word /kat/. After converting the word into a sound, it is converted into the image of a cat in the pictorial model. Learners' schemata stored in LTM (Sc) is also important, once it is retrieved back to WM to be processed with the incoming information to be converted and integrated (Sc+1) as new strings of knowledge in LTM (Sc1). Input, however, enters learners' cognitive system not only through the eyes, but also through the ears, as discussed, which is explained as follows.

2.2.2.3 The Processing of Spoken Verbal Information

Spoken verbal information concerns the oral (spoken) input learners are exposed to while watching videos. When learners are provided with spoken text on a video, the input information enters their cognitive system through the auditory channel, that is, the ears. The

process that takes place is comparatively similar to the ones of the processing of images and written text, and can be seen in the following diagram.

Figure 5 – How the Processing of Spoken Verbal Information Takes Place



Source: The researcher, adapted from the diagram presented by Mayer (2005a)

Firstly, learners need to select the pieces of oral information that they may consider relevant. After the information is selected, it is conducted to WM, where the sounds heard are converted into images. These images are the written (typographical) representation of the sounds heard and the imagery representation of the objects, or ideas heard, as shown by the two arrows, one pointing from sounds to images, and the other back from images to sounds. For example, when learners hear the word school, they mentally convert the sound into the representation of the written word school and the mental image of a building with windows, doors and children representing the concept of a school. After this conversion, the sounds are organized in a verbal model of the word school. The schemata that learners have stored in their LTM (Sc) plays an important role, once it is retrieved back into WM to be processed with the new information and integrated (Sc+1) as new strings of knowledge to be stored in LTM (Sc1) as learning.

As afore presented, learners engage in different types of processes when they are exposed to videos presenting the target information making use of images and words in a combined fashion. The processing of the information being received takes place in the WM, which has a limited capacity of processing. This is further expanded in the next subsection.

2.2.3 Limitations of Working Memory

As discussed in the previous subsection, working memory (WM) is responsible for the processing of the information that enters learners' cognitive system through the eyes and the ears. The case is that not all types of video with the objective of aiding learners in the process of developing an L2 can have similar effects. Depending on the kinds of information presented in the video and on the way it is designed, it can impose a cognitive load on learners (PLASS; MORENO; BRÜNKEN, 2010).

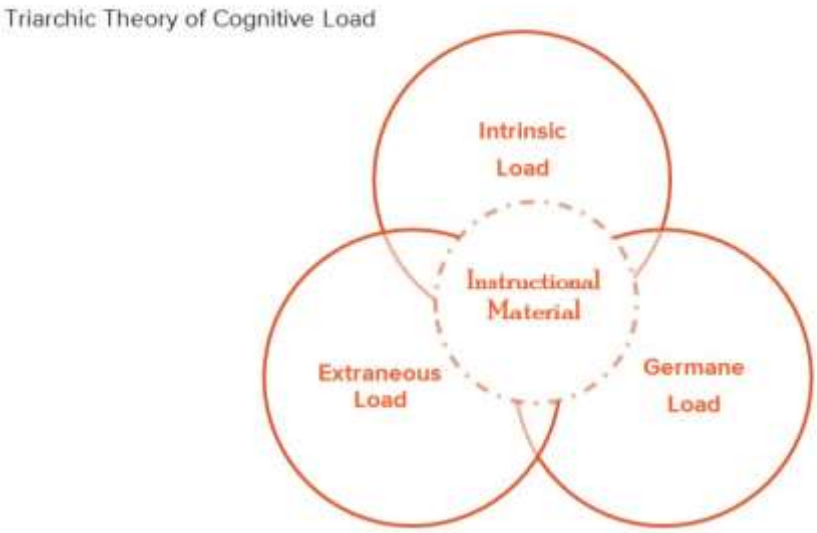
According to Lee and Mayer (2018), cognitive load is associated to the total amount of effort imposed on the WM during any given learning experience in making sense of the information received through the sensory memory (SM) for a moment in time. As explained by Sweller (2017), WM seems to have a considerably limited capacity (amount) and duration (time) when processing new incoming information. The degree to which the incoming information can be considered new as a whole or in parts depends on the amount of strings of knowledge learners have already stored in long-term memory (LTM) about the information being received as input. The case, then, becomes that the lesser prior knowledge on a certain topic learners have stored, the higher the cognitive load the information will impose on their WM.

Considering that the main objective of educational materials is to increase the strings of knowledge stored in LTM, Sweller (2010), further considering the limitations of WM, explains that

> All instructional material imposes a working memory or cognitive load, and that cognitive load can be divided into two independent categories – intrinsic and extraneous – with a third category, germane cognitive load, dependend on intrinsic cognitive load. (SWELLER, 2010, p. 40).

Acknowledging that in any learning experience, WM plays a crucial role, an examination of the three kind of loads aforementioned becomes needed. They are referred as the triarchic theory of cognitive load (MAYER; MORENO, 2010) once they are associated to the three different levels of processing that underlie learning, as presented in Figure 6. Each of the three elements is discussed as follows.

Figure 6 – Model of the Triarchic Theory of Cognitive Load



Source: The researcher, based on Mayer and Moreno (2010)

As presented in Figure 6, any type of instructional material can be loaded with the three aspects of cognitive load. Even though they have specific characteristics each, they all interact and are defined as being additive because, together, they regulate the total amount of cognitive load an instructional material can impose on a learners' cognitive system. Each one, Intrinsic Cognitive Load, Extraneous Cognitive Load, and Germane Cognitive Load are explained as follows.

2.2.3.1 Intrinsic Cognitive Load

Intrinsic cognitive load is the level of complexity between the elements presented in an instructional material (SWELLER, 2010), also referred to as essential processing (MAYER, 2009; MAYER; MORENO, 2010) due to the fact that all the elements presented in an instructional material – in this case, videos – are (at least should be) essential for the full understanding and learning. Intrinsic cognitive load can be defined as the connections learners must make between the information elements that are presented in the material (KALYUGA, 2010) in order to be understood. It referes to the inherent difficulty of the material. These connections are based on learners' LTM and WM sources that need to be activated during the information processing also concerning the new information density, that is, the amount of new information in relation to the old information (information already established in LTM).

The level of intrinsic cognitive load materials have can be defined based on the amount of schemas that learners have stored in their LTM. According to Sweller (2010), LTM can be

considered the base of human cognition, because it is not only used as a box for referring to past events, but it can work as a source for problem solving, thought and regulating present actions. Based on this concept, the schemas that are already stored in learners' LTM work at the service of how the new incoming information can be turned into learning, thus being also incorporated into it, enlarging the schemas already existing. For this process to take place, WM takes an equally important role. For new information to be integrated and stored into LTM, it first needs to be processed in WM with the already existing schemas that are retrieved from LTM back to the WM, as shown in Figures 1, 2 and 3. For example, if learners watch a video that presents vocabulary about house furniture, the elements in the video can be considered to have little interactivity among each other, being the video considered of low intrinsic load, once the vocabulary items can be learned independently from one another (AL-SEGHAYER, 2001). On the other hand, considering a different scenario, in which learners watch a video with a narrator describing his or her last vacation, bringing information of the activities done, the places visited and the dishes eaten, the level of element interactivity can be considered high, resulting in a high level of intrinsic load because learners need to process the language that is being used (the past tense) along with the vocabulary about activities, about names of places and food. In this case, all the elements (vocabulary and grammar) need to processed simultaneously, resulting in a high level of WM demands for the information to be assimilated and integrated.

An important issue that arises is: how is it possible to deal with the intrinsic cognitive load so that learning can be fostered? The element interactivity is inherent to the information being presented on the material. When learners are watching a video where the characters are talking about everyday routines, for example, it becomes inevitable that some amount of prior knowledge about everyday activities and the language structure of the present simple tense is needed. One possible alternative that can be adopted in order to decrease the amount on WM resources that (especially less proficient) learners need to use is to organize the information in a series of separated elements (segments) in which they do not interact between each other (SWELLER, 2010), and only in a later moment put them together as being part of a whole. The provision of information in segments can afford learners with a limited amount of information each time so that they can have time to digest and assimilate what they are watching and prevent them from engaging in extraneous processing by overloading their WM resources. Moreover, presenting only a reduced amount of information – in a concise way – in each segment can assure that they will not overload their WM capacity (SWELLER, 2010) by focusing their attention to one element at a time. When learners are presented with learning materials that do

not consider the elements that may cause extraneous load, learning may be hurt because learners' WM resources may not be sufficient to process and build the necessary coherent mental representations of the information intended to be learned (KALYUGA, 2010).

As discussed earlier, the case that element interactivity may result in WM overload does not apply to all learners equally. For learners with a higher language proficiency, the tendency is that element interactivity becomes lower. As proficient learners have a greater amount of schemas already stored in LTM, when these strings of knowledge are retrieved to WM in order to be processed with new coming information, they are treated as one single schema and also leave room for the processing of linguistic elements and not overloading WM resources. For novice learners, with a reduced amount of schematic knowledge stored in LTM, all the elements must be processed simultaneously and the material can be considered intrinsically high due to their WM demands)MAYER, 2005a; 2009).

It is important to reiterate, however, that the three cognitive loads (as presented in Figure 6) are additive and interact among each other, making the design of educational materials in segments not restricted to preventing learning from engaging in extraneous cognitive load, but also in managing the intrinsic load the information presented may impose to learners.

2.2.3.2 Extraneous Cognitive Load

Extraneous cognitive load concerns the processing of information that is not relevant for achieving the intended objectives imposed by the material or when the information itself presented is poorly designed demanding learners to interpret it themselves with no effective guidance. Mayer (2009) also refers to it as extraneous processing, once it results in the processing of irrelevant information, thus using valuable WM resources that, otherwise, would be directed to the processing of essential information.

A series of factors can cause extraneous cognitive load, attributed to learners themselves or to the instructional material being presented. In the former case, it is related to the limited capacity that WM has to process only a small amount of information each time (KALYUGA, 2010), or the amount of schemas related to the information presented. In the latter case, it concerns materials, videos in the case, that are not designed based on human cognitive architecture, presenting unclear instructions, or contradictions, demanding learners to try to build connections between the pieces of information, and find the coherence of the information.

An example would be a scenario where learners watch a video aiming at presenting grammatical explanations about the use of the simple past structure. While watching the video,

instructions are presented providing learners with a range of explanations of the use of irregular and regular verbs in affirmative sentences, the insertion of the auxiliary verb in negative sentences and questions, all together followed by examples and images to depict these exemples and accompanied by an exciting soundtrack to work as a motivational aspect. The probable result of such design would the extraneous processing for two reasons: the first is the provision of an enormous amount of information (probably) presented in an unstructured fashion, in which learners need to focus their attentional resources to many elements at the same time; and the second is the presentation of irrelevant information, which is the case of the soundtrack, or the insertion of images that have only an embellishment function, thus, competing for learners' attentional resources.

When videos are designed in a coherent fashion that provides learners with only the pieces of information that are crucial for the understanding of the content, the process of knowledge construction is faciliated, resulting in learning. With well-designed videos that take into consideration extraneous cognitive load, the attentional resources should be directed to the amount of knowledge learners already have and to the essential information so that learning can be fostered, not hurt, resulting in effective learning.

2.2.3.3 Germane Cognitive Load

Ideally, materials design should be intended to direct working memory (WM) processing to increase learners' schematic knowledge, that is, the generation of germane cognitive load, which is also referred to as generative processing (MAYER, 2009; MAYER; MORENO, 2010). The premise of germane cognitive load can be defined as the processing that leads to learning.

Germane cognitive load is associated to a series of cognitive processes learners engage with in the process of schemas construction such as selecting the information that is relevant for the building of coherent mental representations and the integration of the information (MAYER; MORENO, 2010). For example, assuming that learners watch a video that encourages them to picture the information presented in a way that does not overload their WM capacity by coping with intrinsic and extraneous loads, learning may be fostered. When learners are challenged and motivated in the process of building mental representations of the information watched in order to fully understand the ideas or concepts, and become successful in building clear representations that can be incorporated into their LTM as new, or enlarged

strings of schemas, germane learning can be achieved, once it is the result of the cognitive processing necessary to achieving the learning objectives (BRAME, 2016).

When the design of materials is intended to consider the extraneous cognitive load, and the intrinsic load and learners' WM resources can be at the service of processing the essential information, then the germane cognitive load can work in the generation of new schematic strings of knowledge (SWELLER, 2010).

The main objective of instruction is to aid learners in building new and enlarging already existing strings of knowledge they have stored in LTM. For this objective, that is, for learners to be able to convert information received in the form of input into expanded schemas, there is the need for effective processing of it, which takes place within WM. However, for the processing of the new information to be integrated to already stored information in LTM effectively, the limits of WM should not be exceeded. When providing learners, for example, with a video that presents irrelevant information or in larger amount at each time, exceeding the processing capacity of WM, learning can be impaired.

Sweller (2017) argues that one of the most important aspects in promoting the learning of an L2 is that instructions should be designed in a fashion that reduces WM load. The instructional implications that are associated to the limitations of WM consider the promotion of learning by means of freeing WM resources by reducing the cognitive load educational materials impose to learners. In considering the limitations of WM, five pedagogical procedures should be acknowledged (SUEK, 2018), as presented bellow.

The Worked Example Effect, in which learners are provided with exemplifying procedures or steps used in solving a problem, or, in the case of working with a specific language structure, the construction of that structure (*e.g.* examples of Simple Past sentences). It is important, however, to acknowledge that when worked example effect is applied to the instructional design, learners' prior knowledge be considered so that extraneous load be avoided by the possible high level on intrinsic load that the information may be charged with.

The Goal-Free Effect is associated with providing learners with the opportunity of solving a determined task or problem by themselves and finding as many possible outcomes as they can pose a positive effect on the cognitive load and possibly generates germane load, which leads to effective learning.

The Split-Attention Effect advocates that the video design should consider the sources from which learners should retrieve information. One example can be observed when learners are provided with visual and verbal aids, instead of presenting them as separate pieces of

information, they should be presented as integrated information, that is, they should be part of information as a whole.

The Modality Effect establishes that learning an L2 can be aided when instructional materials are designed in a fashion that provides the target information in more than one mode – visual and aural. By providing learners, for example, with the target language input with the use of visual sources and oral language, the cognitive load imposed is reduced so that working memory processing capacity is increased.

Finally, The Redundancy Effect, presenting learners with information in the form of a visual representation of an image, graph or video and then presenting the same information in the form of a written text, for example, may increase learners' cognitive load, thus, hurting learning.

The instructional recommendations provided by Sweller (2017) aim at the design of educational materials that can be considered effective in dealing with the three loads imposed by materials aiming at aiding learners in the process of learning an L2. Such recommendations can be argued to be in consonance with the principles for the design of multimedia materials proposed by Mayer (2009), with twelve principles that should be considered in the design of educational materials so that learning can be fostered. In the next subsection, the twelve principles that are intended at the reduction of extraneous cognitive load, the effective processing of the intrinsic cognitive load, and foster germane cognitive load will be presented. Moreover, a discussion of how these principles can be applied in the design of videos will be presented.

2.2.4 Principles for the Design of Multimedia Materials

As aforementioned, videos can be a source of input to aid learners in the process of L2 learning. In the discussions about how videos can best promote learners' development and the process of knowledge construction, very specific features must be taken into consideration, especially when it comes to their design. It becomes of high importance that, for educational videos to be effective, their design be based on how the human mind processes information. Based on this concept, Mayer (2005a; 2005b; 2009) has studied and conducted research with different multimedia learning environments, resulting in the book entitled Multimedia Learning (2009). Mayer has been interested in how different multimedia materials can aid the process of knowledge construction for longer than twenty years, and the result of his studies is the development of principles for the design of multimedia materials for pedagogical purposes. The

twelve principles proposed by Mayer (2009) are grounded on the Triarchic Model of Cognitive Theory, represented in Figure 6, which can be considered to be part of any educational context, more specifically, any educational video. The principles are intended at providing guidelines for the design of educational materials in regards to reduce extraneous cognitive processing, manage essential processing and foster generative processing. Each of the twelve principles is intended at aiding learners in the process of knowledge construction, and are presented as follows.

Chart 1 – Relationship Between Mayer's 12 Principles and the Triarchic Model of Cognitive Load

Principles for Reducing Extraneous Cognitive Load
1. The Coherence Principle – The materials should focus on presenting only crucial pieces of information, and avoiding any kind of extra (unnecessary) information. When extraneous material is avoided or left out in the multimedia material, learners can learn better. 2. The Signaling Principle – Learners' attention should be turned to important or relevant pieces of information presented in the material, by using different word sizes or colors, or changing the tone of the narrator's voice. 3. The Redundancy Principle – The presentation of the same information twice, for example, the very text presented as a narration, also being presented in the written form, as captions or subtitles, resulting in extraneous processing and poor learning[9]. 4. The Spatial Contiguity Principle – The location on screen where images and text are positioned have an effect on learning, that is, they should be positioned near each other or in a way that learners can easily make connections between them. 5. The Temporal Contiguity Principle – When there is the use of images and narration, or more than one image, these elements need to be presented simultaneously instead of in sequence so that learners can easily make connections between them.
Principles for Managing Essential Processing
6. The Segmenting Principle – Information should follow a learner-paced structure, in blocks or units so that all pieces of information can be more easily assimilated. 7. The Pre-Training Principle – A pre-viewing treatment can help prepare learners for the material they will be exposed to. 8. The Modality Principle – Multimedia materials should be designed in a way that provides learners with two sensory input modes: images (visual elements) and narration (spoken language) rather than images and written language.
Principles for Fostering Generative Processing
9. The Multimedia Principle – Learning is better aided when the multimedia materials present images and words, rather than only words. 10. The Personalization Principle – When the multimedia material is organized in a way that the narration takes place in a more conversational, informal way, learning is benefitted. 11. The Voice Principle – The presentation of the narration should be designed in a human voice rather than a mechanical computer voice to build social and affective engagement. 12. The Image Principle – The presence of the image of the narrator may present some sort of extraneous processing, turning learners' attention to unnecessary objects or information, hence not benefitting learning.

Source: The researcher, adapted from the tables presented by Mayer (2009)

[9] *The Redundancy Principle* as explained by Mayer (2009) may hurt learning in a context of learning where the information is presented to learners in their L1. In the case of the context of learning an L2, presenting redundant information may benefit learning, as discussed in the previous section.

Mayer's (2009) principles presented above can be identified in educational multimedia materials. Each principle plays a specific role concerning how information is presented in order to be processed for knowledge construction and, ultimately, stored in long-term memory (MAYER, 2009). They can be used for the design of videos for educational purposes, resulting in materials that can enhance the development of knowledge. Researchers (IBRAHIM , 2012; IBRAHIM *et al.*, 2012; BRAME, 2016) have focused on three major principles, namely: Signaling, Segmenting and Weeding, to be considered in the analysis and design of educational videos with the purpose of optimizing the process of knowledge construction, especially considering the transient nature of the flow of information presented in the video (LEAHY; SWELLER, 2016).

Based on the Cognitive Load Theory (SWELLER; AYRES; KALYUGA, 2011), Ibrahim *et al.* (2011) suggest that the three principles that guide the design of multimedia materials, in this case educational videos, can be organized in two groups.

The first group is aimed at reducing extraneous cognitive load and increasing learners' germane load. For extraneous processing to be avoided, two principles are considered in the design – Signaling and Weeding. In Signaling, essential pieces of information are made salient to direct learner' attention to them and help in the selection and building of mental representations of the information being exposed. The signaling principle has the further objective of saving working memory resources for the processing of the necessary pieces of information. Weeding is concerned with removing unnecessary information and elements so that only the essential information is processed, thus, increasing the germane load, which is crucial for learning.

The second group concerns with helping learners by organizing the material in blocks, or chapters, to help learners in the processing of the intrinsic load of the information presented in the video, that is, segmenting. The use of this principle in the design of educational videos provides learners with time to process each piece of information, presented in each segment, before moving to the next so to prevent the overload of precious WM resources and extraneous load, as argued by Ibrahim *et al.* (2012).

It is important to emphasize that the twelve principles proposed by Mayer (2009) and the SSW principles proposed by Ibrahim (2012) are aimed at working as a guide for the design of educational videos, that is, they are expected to be considered when educational videos are in the design process. The fact that these principles work as a guide for the design process can be especially observed in the term weeding (removal of irrelevant materials), which, when applied in the production process, aims at removing the parts that can result in extraneous

processing (MAYER, 2009; SWELLER, 2010). Since this Doctoral Dissertation will deal with the analysis of already-designed videos, the Coherence Principle rather than weeding is used, that is, dealing with the analysis of pieces of information that may be considered irrelevant to foster learning. All the three principles – signaling, segmenting and coherence – also referred to as SSC are explained in more detail in what follows.

2.2.4.1 Signaling

The first principle that has been extensively studied concerning its effects on learning with multimedia materials is signaling (MAYER, 2009; IBRAHIM, 2012). In the context of L2 learners, when learners watch videos that present information, that is, content and the target language in a visual and oral transient fashion, the processing of information can result in a relatively overwhelming task for learners due to the amount of information, the pace it is presented (KONING *et al.*, 2007) and the complexity of what to focus their attention to. Although the term signaling is largely used, the term cueing[10] (KONING *et al.*, 2010) is also used in research to refer to the same principle and refers to making specific pieces of information available in the video salient, turning learners' attention to them. In educational videos, signaling plays the role of directing viewers' attention to important or relevant pieces of information to facilitate learning.

The concept of signaling can be associated to Richard Schmidt's Noticing Hypothesis (1990). According to the scholar, learners' attention and awareness can have an important effect on learning. Schmidt's (1990) hypothesis has worked as a foundation for further theories concerning how attention and awareness can affect learning. Schmidt (1990, 2010) also defines three different levels of awareness (also referred as consciousness) that, according to him, are crucial when learners are in the process of acquiring an L2. The first level, referred to as perception, concerns with learners creating mental representations of external information that is presented in the form of input. The second level, referred to as noticing, or focal attention, is concerned with learners being aware of the event being performed or part of. The last level, understanding, takes place when learners engage in the process of comprehending.

Considering signaling as directing learners' attentional resources to specific pieces of information in order to facilitate the learning of an L2, Plass and Jones (2005) argue that educational multimedia materials (in the context of this Doctoral Dissertation, videos) can have

[10] In this Doctoral Dissertation, the term *signaling* will be used interchangeably, though.

a positive effect in supporting the learning of an L2. Educational videos that are designed following the signaling principle aim at helping learners' cognitive processing of the information presented to be processed and incorporated in a way that results in learning (KONING *et al.*, 2007). A point of departure of creating signals in a video is to help learners select relevant information and build coherent mental representations of the information presented. As learners watch the videos, the information watched needs to be processed simultaneously as new flow of information is presented because of the transient nature that can be attributed to videos. As transient information can become difficult for learners to process, especially because video contents can be complex (of high intrinsic load), the use of signals can offer opportunities to turn learner's attention to relevant pieces of information by making them salient.

According to Plass and Jones (2005), some of the signaling aids that can be inserted in a video design are "highlighting or otherwise marking words, sentences, and linguistic features, often in conjunction with providing comprehension aids in the form multimedia annotations and pairing words with images." (PLASS; JONES, 2005, p. 472), also discussed by Benkada and Moccozet (2017) as facilitating learners' comprehension. In addition, signals can offer a variety of visual aids that can help turn learners' attention to specific elements and pieces of information, such as adding colors, images, arrows, or positioning written text in a specific location of the screen. The visual element that is used in order to make specific information salient, especially in images, is referred to as visual contrast (KONING *et al.*, 2010). This visual effect makes the intended parts become visually more salient by contrasting their colors with the rest of the elements of the video. When it comes to signaling with verbal oral language, Chi (2018) explains that the use of different word stress or sentence intonation can make specific pieces of the target language salient, thus, directing learners' attention to them. Furthermore, presenting the verbal language in the end of the video, or of each segment, working as a recapitulating feature in order to summarize the main ideas or concepts previously presented, can also aid learners in the process of knowledge construction.

Regarding the verbal written language as a signaling design feature, the Wochit Company web page (www.wochit.com) brings an extra explanation using written text in videos. Although the company is not aimed at developing educational videos, its web page provides a term for the insertion of text in video design, referred as overlay text (WOCHIT, 2022), whose objective is to call viewers' attention to video content. In this dissertation, however, the insertion of written verbal information in videos design will be henceforth addressed as on-screen target text, as it concerns with presenting learners with target written verbal information

on the screen in a direct fashion for learners to consciously see, select the relevant pieces of information for the comprehension of the information being watched, and integrate with the oral information. The insertion of such signaling features in the design of educational videos is grounded on the notion that when learners notice, or become aware of the information being watched, learning of the target language is more likely to happen.

One of the crucial aims of signaling, then, becomes the turning of learners' attention to specific parts of information, or to the relation among the pieces of information presented. However, considering the human cognitive architecture (SWELLER, 2010; KALYUGA, 2009; 2015), depending on the amount and the way information is presented, learning can be affected because learners' working memory can become overloaded in trying to process all the information, having difficulties in selecting which pieces of information to focus on (HASLER; KERSTEN; SWELLER, 2007). Such effect can be the result of materials that fail in the design that follows the human cognitive architecture. When learners are aided in the process of selecting the information to turn their attention to, processing the new information with already stored schemata, and integrating it into new strings of schemas to the long-term memory, the result may be better learning. For effective learning to take place, it is important that the signals match or accompany the thematic relevant aspects, that is, the signals should be in accordance with the mental representations that learners (should) build with the help of the WM processing.

Studies have investigated how different types of signaling can affect learners. Koning *et al.* (2007), for example, conducted a study in order to investigate whether signaling presented in the form of visual cues about the cardiovascular system and signaling its functioning by making the subsystems salient could have a positive effect on the comprehension as well as retention of the information presented in an animation. The results suggested that participants who were presented with animations with signaled information had better performance in comprehension and retention compared to participants that were presented with a non-signaled animation, suggesting that signaling can work as a helpful design feature to enhance learning.

Another study by Koning *et al.* (2010) investigated signals in the form of oral instructions. The aim of their study was whether providing animations with oral instructions may be more effective for learners to build mental representations of the information provided in animations with oral instructions compared to animations with no signaled instructions, where learners were demanded to infer the information through self-explanations. Moreover, the researchers investigated the effects of guiding attention in relation to the signaling effect in the oral instruction. After testing for retention, inference and transfer for the information

presented, the results indicate that the signaled animations provided more effective support for learners to build mental representations of the information provided.

The influence signaling has on learning can be further exemplified with a study conducted by Lorch (1989). In the study, the author investigated how signaling in written input can influence the comprehension and information retention effect in text reading. According to the author, signals are "devices that emphasize aspects of a text's content or structure without adding to the content of the text" (LORCH, 1989, p. 209). The author presents a list of different kinds of signal devices that can be incorporated to facilitate comprehension. These devices include titles, headings and subheadings, repetition of a statement for emphasis, function indicators, relevance indicators, enumeration devices, which can be presented as typographical cues, or as oral cues – as in the case of teachers' or instructors' narration. Although the signaling devices presented by Lorch (1989) are concerned with written texts, they can also be in educational videos in both oral and written formats.

The studies presented above provide evidence that the use of the signaling principle in the design of educational videos can result in aiding learners in turning their attention to relevant parts of the information being provided on videos. In addition, signals can help learners in building effective mental representations of the information reducing the overload of the WM processing resulting in learning.

2.2.4.2 Segmenting

The second design principle in the SSC model is segmenting. The most important characteristic that the segmenting principle holds is to help learners in the processing of the information presented by managing the intrinsic cognitive load the information depicted on videos may impose (as mentioned, depending on learners level of language proficiency). Thus, the segmentation design refers to breaking the video, or the information provided, into chunks or blocks in order to avoid information overload, and ultimately, cognitive overload, granting learners time do digest the information watched and build coherent mental representations of it. According to Moreno (2007, p. 768), "in this way, students have the opportunity to reflect on smaller portions of the information before they move on to view the next portion."

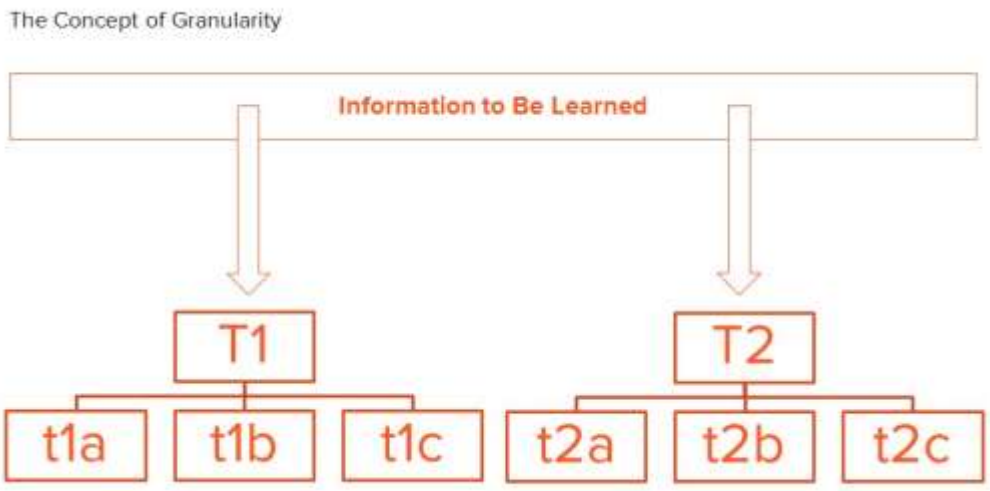
Videos present information in a transient fashion and, because of that, they are different from books. The greatest difference between these two media is that in the latter, information is presented in a still or static way. When studying a book that brings information in the form of written text accompanied by images, learners have the opportunity to go back to any parts

that are unclear and revisit the parts they judge to be necessary for understanding, or even move forward when they are ready for the next pieces of information. When working with static information, learners can use the amount of time they judge necessary for the understanding of the information provided, thus having more time for active processing (AYRES; PAAS, 2007). Videos, on the other hand, as previously said, present information in a transient way, in which the information previously presented needs to be processed at the same time the newly information is presented. Since videos present transient information that, at times, result in a difficult task for learners to process due to the pace (or flow) and/or amount of information, the result may be extraneous processing (AYERS; PAAS, 2007) resulting from exceeding their WM resources, thus hurting learning.

As discussed in previous sections, it is important to consider that the human cognitive system can engage in the active processing of only small amounts of new information each time (IBRAHIM *et al.*, 2012). Based on this premise, when videos are broken into smaller blocks, learners can be provided with time to process the information being presented in order to draw connections among the pieces of information and reduce the amount of intrinsic cognitive load (MORENO, 2007). In addition, with videos that are organized following clear-cut start and end points, the sense of organization and structure can help learners build not only the necessary coherent mental representations that can result in learning, but also create connections among those pieces of information.

The segmentation design can be compared to the concept of learning objects (LO), once LOs are organized in small pieces of information, also defined as granules (GRIFFITHS; STUBBS; WATKINS, 2007). Based on the assumption that the smaller the LOs are the more effective they can be, aiding the learning process and also be re-used in different contexts (NOOR; YUSOF; HASHIM, 2011), Thompson and Yonekura (2005, p. 169) explain that "[e]ach learning object can address only one learning objective", that is, they are designed in a way as small as possible, in which each one is limited to no more than a single piece of information, in a considerably granular size. Segmentation, can, hence, be compared to the principle of granularity, in which the main aim is to present learners with small pieces of information. This idea is depicted in the following figure.

Figure 7 – The Concept of Granularity Applied to Educational Topics



Source: The researcher

As presented in Figure 7, granularity can be associated to segmenting in the design of educational videos since both refer to breaking a topic down into smaller pieces, organized in blocks. For example, when a certain topic (T1) is to be learned, alternatively to being presented to learners as one whole piece of information, it can (and should) be organized in smaller segments (t1a, t1b and 1tc), each one to be focused one at a time. Only after having gone through all the (granular) segments that compose T1, can learners move on to T2, which is also organized in smaller segments in order to not to overload their WM processing capacity (JUFFS; HARRINGTON, 2011) and provide them with enough time to mentally establish the necessary relationships among the (pieces of) information. The granular organization of the information that learners are expected to learn can provide them with time to assimilate all granular pieces of information, and incorporate them altogether as one, as T1 and T2.

In addition to being a built-in feature of the design, segmenting can also be characterized by learner-control (HASLER, KERSTEN; SWELLER, 2007; AYRES; PAAS, 2007). When learners are provided with an educational video, they can exercise control to pause the video and go back to re-watch any part that resulted unclear for them, or with the next button, where they may move forward only when they feel ready for the next block. That is, they can control of the pace of the information being provided in organized segments as well as adapting the pace of the information to their cognitive needs (HASLER, KERSTEN; SWELLER, 2007).

An example of the effects of segmenting in educational videos is reported by Moreno (2007), who investigated the effects of segmenting accompanied by signaling on information retention by presenting learners with video and animation. Participants were divided in four

groups: signaling/no segmentation (group 1), no signaling/segmentation group (group 2), no signaling/no segmenting (group 3) and signaling/segmenting (group 4). The results suggested that the groups that received the segmented-design videos (groups 2 and 4) had a better performance in learning and showed to have a decreased processing load than groups 1 and 3.

Another study conducted by Hasler, Kersten and Sweller (2007) investigated the effects of learner-control could have on primary school learners by means of a stop and play button (group 1), next button (group 2), system pacing animation (group 3), and segmenting and narration-only (group 4). The results suggested that when learners had the control over the video, their learning was higher compared to when they did not have the control.

With the examples mentioned above, it is possible to argue that when educational videos are designed following the segmenting principle, learning can be enhanced by the reduction of working memory processing load, especially for less proficient learners, who may have difficulties in establishing connections between the information presented, especially when the level of intrinsic load may be high.

2.2.4.3 Coherence

Finally, the third principle in the SSC design is coherence. The coherence principle is part of the design of educational videos aimed at reducing or preventing extraneous processing and, thus, allowing for germane processing, that is, the construction of schemas in long-term memory (KALYUGA, 2010) by eliminating unnecessary pieces of information such as soundtrack or background noises or pieces of information that do not contribute to achieving the educational objective. In an applied view, as explained by Mayer (2009), coherence in the design of educational multimedia materials, in this case, videos, involves the elimination of elements such as background sounds or music and visual elements that may have only aesthetic function. In the case of this study, coherence principle will deal with the identification of any pieces of information in the videos that can be considered irrelevant, thus, causing extraneous cognitive processing (MAYER; MORENO, 2003).

Considering that coherence aims at preventing learners from extraneous processing, Ibrahim (2012) presents three types of processing of extraneous materials: the first type of processing takes place when videos present information that may be interesting, but irrelevant for the educational purposes; the second type takes place when redundant information is presented; and the third type takes place when redundant information already familiar is

presented, competing for attention and WM processing. As is reported by Ibrahim (2012), redundancy may result in extraneous processing, thus hurting learning.

The coherence principle justifies the presence of relevant information and, thus, the avoidance of unnecessary, or redundant information. This may be applied to the presence of written texts repeating the information given in the oral mode. There are studies showing that presenting information in the forms of oral text accompanied by the written language may hurt learning, causing cognitive overload (MAYER; HEISER; LONN, 2001; MAYER, 2005a; 2005b; 2009). However, there are situations that may challenge this claim, including the case of L2 learning contexts.

The presentation of information by means of visual aids accompanied by spoken text along with the written form may help learners in the learning process and can be exemplified by a study conducted by Moreno and Meyer (2002). In their study, the researchers conducted three experiments in order to investigate how the presentation of information by means of images, spoken text and written text would affect learning of scientific information. For such, participants were provided with a material that presented the process of lightning formation and were organized as two major groups. The first received only spoken information, referred to as non-redundant group. The second received spoken and written information, referred to as redundant group. After testing for learning by applying retention, transfer and matching tests, the researchers found that participants that were part of the redundant group performed better than the ones in the non-redundant group, suggesting that presenting information in both oral and written forms can, indeed, foster learning.

It is important to consider, however, that the arguments presented by Ibrahim (2012) and Mayer(2009) are concerned with materials aimed at providing domain knowledge such as the formation of lighting or the working of a pump, as is clearly reported in their research studies. The discussion of the principles in this Doctoral Dissertation are concerned with the design of educational videos aimed at aiding the process of learning English as an L2. In this case, redundancy, rather than possibly provoking extraneous processing, can be argued to foster learning, as is the case of the use of captions or subtitles in videos (HAYATI; MOHMEDI, 2010), as previously discussed.

In sum, based on the evidence reported by researchers that have their foundation on the cognitive load theory, the information that is presented in educational videos is processed in the WM. However, when it comes to limitation, it is argued that WM capacity is limited (KALUYGA, 2009; MAYER, 2009; SWELLER, 2010;) and extraneous information competes to be processed with essential information, thus, making the coherence principle as important

as the signaling and segmenting principles for the design – and analysis – of educational videos and for these materials to effectively aid learning.

This chapter provided a review on the literature of the relevance videos can have as a pedagogical tools in the context of L2 learning. Acknowledging videos as multimedia materials, in which the use of images and written, and/or oral verbal language in a combined form can enhance learning, they can be defended to provide opportunities for learners to build coherent mental representations of the information watched, in the case of an L2, the target-language. Design principles are argued by researchers to play an important role in the design of educational videos, once they are based on the human cognitive architecture, and grounded on how people learn. Although the principles presented in this chapter work as a guide for the process of designing educational videos, in the case of this Doctoral Dissertation, already-produced videos aimed at the teaching/learning English as an L2 will be analyzed.

In the next chapter, the methods and procedures taken for the selection (and exclusion) of the videos to be analyzed as well as the video analysis framework, are described in detail.

3 METHOD

In this chapter, the procedures applied for the development of this Doctoral research study are explained as to the characteristics of the study, the choice of the videos analyzed, and the framework developed to work as a guide for the analysis.

The first section, 3.1, presents the characterization of this study including explanations about the procedures for piloting the analysis and the decisions taken concerning the exclusion/inclusion of the sources from which the analyzed videos were taken. Moreover, the criteria for the selection and exclusion of videos are explained.

The second section, 3.2, presents information about the selection and exclusion processes that led to the choice of two sources of the videos analyzed. The first source concerns videos that were approved by the PNLD (*Programa Nacional do livro e Material Didático*) program and the criteria used for the selection and exclusion of the videos. The second source concerns videos that are part of the Interchange Series by Cambridge Publisher.

Finally, section 3.3 presents the video analysis framework developed based on the discussions presented in the Review of Literature chapter that guided the analysis and discussions presented.

3.1 CHARACTERIZATION OF THE STUDY

Grounded on Dörnyei (2007), this study can be characterized as being of an exploratory and interpretative nature aimed at investigating digital videos developed for the teaching and learning English as L2. For achieving this objective, studies that investigated videos as multimedia materials and how they can enhance learning (KIM; GILMAN, 2008; ZAREI; KHAZAIE, 2011; IBRAHIM, 2012; IBRAHIM *et al.*, 2012) worked as a basis for the development of an analysis framework used for analyzing the videos. Furthermore, this research study aimed at investigating which features the digital materials, in this case, videos, present to foster the learning of English as an L2, following the analysis criteria proposed in the framework.

Initially, the need to determine the sources of the videos to be analyzed was necessary. There were four main sources: 1) The videos that accompany the textbooks approved by the PNLD program; 2) Videos available on YouTube; 3) Videos from Khan Academy; and 4) Videos that accompany the textbooks produced by National Geographic. As four sources would result in an excessive quantity of videos to be analyzed, and the need for narrowing the sources

became essential, National Geographic as a source was excluded for the fact that the videos that accompany the textbooks are excerpts from documentaries produced by the publisher, thus, not being pedagogically designed for language learning.

In a second stage, a pilot study was conducted for the Qualifying Exam. As part of the Doctoral Program from the Post-Graduation Program in English (*Programa de Pós-Graduação em Inglês – PPGI*) from the Federal University of Santa Catarina (*Universidade Federal de Santa Catarina – UFSC*), the Qualifying Exam is a pre-requisite for the conclusion of the program. In the process, the Doctoral candidate is expected to have a part of the dissertation written, including the review of literature, and when possible, preliminary data analysis to be read by two professors who provide feedback and suggestions about the ongoing research study. In the case of this Doctoral research study, for the pilot study, three videos were analyzed, each from one of the three remaining sources: one video from a PNLD textbook, one from Khan Academy, and one available on YouTube.

The three videos were analyzed following the SSC framework (already presented) that was developed by the student-researcher based on previous studies that aimed at providing principles for video design (IBRAHIM, 2012). The analyses involved the signaling devices they presented, the segmentation structure they had (if any), and coherence concerning the pieces of information or features that could be considered irrelevant, thus, allowing for extraneous processing. Not only did the results of the pilot study as well as the feedback provided by the two professors who were part of the committee lead to a change in the sources of videos for the analysis, but also to the refinement of the video analysis framework.

The Khan Academy videos and the YouTube videos were excluded from the analysis of this Doctoral research study. The exclusion criterion concerned that the Khan Academy videos all follow the same blackboard explanation design pattern, not offering a diversity in features to viewers. The criterion for excluding the YouTube videos concerned that videos from any sources become available on the site, as well as a countless number of videos presented only by typing in the search bar a language topic, (for example, simple present), making it problematic to establish a criterion for selecting the videos to be analyzed. Based on the pilot study of the three videos, only the analysis of the audiovisual materials that accompany the PNLD textbook was maintained.

Moreover, based on the preliminary results from the pilot study conducted, the video sources were reconsidered and reorganized in two main sources of videos to be analyzed. The first source concerned the audiovisual materials that accompany the textbooks approved by PNLD program. Videos from this source were kept for the analysis once a number of English

programs were approved, presenting a diversity in audiovisual materials and ultimately in design features for the analysis. The second source considered the videos part of the Interchange English Series for its well-known relevance in the context of English teaching and use around the world. The criteria adopted for the selection and exclusion of videos in each of the two sources are explained in detail as follows.

3.2 THE SOURCES OF THE VIDEOS FOR THE ANALYSES

3.2.1 The First Source – Videos from Materials Approved by the PNLD Program

According to information on the Ministry of Education web site, this program is one of the oldest programs still running in Brazil, whose aim is to offer educational materials – textbooks – for learners from public schools from the country, being first launched in 1937 (BRASIL, 2022). Throughout the years, the program has had different classifications and different ways of organization and distribution of the textbooks. In 2017, the name attributed to the Program was extended from *Programa Nacional do Livro Didático* to *Programa Nacional do Livro e Material Didático*, which expresses the expansion to supporting materials for the teaching and learning, such as computer software, DVDs and educational games. In order to encompass all students from the public education sector in Brazil, this program organizes the distribution of textbooks and materials in four alternate cycles, ranging from children education, elementary school, which is divided in the initial years, (from first to fifth year), final years (from sixth to ninth year), and high school. One of the main features that the textbooks need to present to be selected to participate in the Program for 2020 was the offer of audiovisual materials available in digital formats. These videos aim at assisting the development of the topics worked in the classroom besides the materials in the analogical format, that is, the textbook.

One criterion for the pedagogical materials to be approved by the PNLD program is that, in addition to having the hardcopy of textbook to be used in class accompanied by the audio CD, or digital audio materials, the pedagogical materials should also present a range of extra materials in digital format to complement the topics to be worked in class. Based on this criterion, the extra materials that accompany the textbooks include a range of audiovisual materials, that is, videos as being part of the pedagogical package to be used in class.

For 2020, nine textbooks were approved. Teachers from public schools, then, received them to analyze and select to be implemented in each school. The list of the textbooks approved

was made available at the PNLD digital guide website[11]. The videos, which are part of the materials will, then, be the focus of the analysis and discussion in this research study.

In this study, the analysis of the materials from the PNLD program dealt with three textbooks. As there are nine textbooks (Alive!, Become, Beyond Words, Bridges, English and More!, It Fits, Peacemakers, Time to Share and Way to English for Brazilian Learners) on the site of the PNLD program that were approved to be part of the program, only three textbooks were selected. The criteria used for the inclusion and exclusion of the textbooks from which the video was analyzed is explained in what follows.

a) For the textbooks to be included in the list of materials to be analyzed, the audiovisual materials that accompany them needed to be available and accessible online[12], on the textbook website, and the textbooks English and More!, Beyond Words, Peacemakers and Alive! were excluded from the list due to the fact that the audiovisual materials could not be accessed online.

b) The textbook from which the video was analyzed in the pilot study of this Doctoral Dissertation was not used for the final version of this dissertation, that is, the textbook Way to English for Brazilian Learners was excluded.

c) After the initial exclusion, four textbooks remained: It Fits, Become, Bridges and Time to Share. From these four materials, two of them are published by the same company, FTD: Become and Bridges, making the exclusion of one of the two textbooks advisable. The exclusion criteria for the selection of the textbook to be removed from the analysis list was based on the titles of the two textbooks. Following the alphabetical order as the inclusion criteria, the textbook Bridges was excluded.

Grounded on the criteria for the inclusion and exclusion of the textbooks, the result was that three textbooks remained: It Fits[13], Become[14], and Time to Share[15]. As the materials are aimed at teaching/learning English in regular public schools, each one is organized in four levels, that is, one textbook for the sixth grade, one textbook for the seventh grade, one textbook for the eighth grade, and finally one for the ninth grade.

---

[11] https://pnld.nees.ufal.br/pnld_2020/componente-curricular/pnld2020-lingua-inglesa

[12] It is important to clarify that, although the videos were accessed online in 2020, they may possibly no longer be available at the publishers' websites by the time of this Doctoral Dissertation defense.

[13] https://pnld2020.smeducacao.com.br/colecao-It-Fits_Ingles

[14] https://s3-sa-east-1.amazonaws.com/pnld2020.ftd.com.br/ingles/P20-2-ING73-6-01-CDO-001_generico_1562614195/index.html

[15] https://educacaobasica.editorasaraiva.com.br/pnld/edital/pnld-2020/oeds/?disciplina=154

The analyses of the materials from the three textbooks were conducted with five videos. At a first moment, the analyses were structured with the selection of two videos from each one of the three textbooks. However, each textbook (course) is organized in four levels (for example, It Fits *6º Ano*, It Fits *7º Ano*, and so forth), which also needed to be taken into consideration for the selection criterion. Thus, the videos analyzes in relation to the textbooks and levels took the following organization.

a) From the textbook It Fits, the two videos analyzed were one from the sixth grade level and one from the seventh grade level.
b) From the textbook Become, the two videos analyzed were one from the seventh grade level and one from the eighth grade level.
c) From the textbook Time to Share, the two videos analyzed were one from the eighth grade and one from the ninth grade.

Each textbook includes a number of videos that are available online. The criterion for the selection of the videos from each grade level was the first available online, for example, from the textbook It Fits, the videos analyzed were the first video available for the 6th grade and the first available for the 7th grade, on the textbook program site.

For a better visualization of the videos selected from the textbooks approved by the PNLD program, the information is presented in the following chart.

Chart 2 – Videos Analyzed from the PNLD Textbooks

Textbook	Videos Corresponding to Grades
It Fits	1 (One) Video from the 6th Grade[16]. 1 (One) Video from the 7th Grade[17].
Become	1 (One) Video from the 7th Grade[18]. 1 (One) Video from the 8th Grade[19].
Time to Share	1 (One) Video from the 8th Grade[20]. 1 (One) Video from the 9th Grade[21].

Source: The researcher (2022)

[16] Video no longer available.
[17] Video no longer available.
[18] https://s3-sa-east-1.amazonaws.com/pnld2020.ftd.com.br/ingles/P20-2-ING73-7-01-CDO-001_generico_1562614308/index.html
[19] https://s3-sa-east-1.amazonaws.com/pnld2020.ftd.com.br/ingles/P20-2-ING73-8-01-CDO-001_generico_1562614457/index.html
[20] https://plurall-content.s3.amazonaws.com/oeds/NV_ORG/PNLD/PNLD20/_videos/Time_To_Share/9ano/Caracterizados/Falando_sobre_obrigacoes_e_necessidades.mp4
[21] https://plurall-content.s3.amazonaws.com/oeds/NV_ORG/PNLD/PNLD20/_videos/Time_To_Share/8ano/Caracterizados/Usando_going_to_e_will_para_falar_de_futuro.mp4

The criteria presented above worked as an organization structure for the selection of the three textbooks and also of the six videos selected from them.

3.2.2 The Second Source – Videos from the Interchange Textbooks

Interchange is an English Series developed by the Cambridge University, authored by Jack Richards with the collaboration of Jonathan Hull and Susan Proctor, published by the Cambridge University Press. The course is organized in four levels (composed by the Student's Book and the Workbook), that is, in four textbooks – Interchange Intro, Interchange 1, Interchange 2 and Interchange 3, each one organized in sixteen units following a communicative approach by presenting learners with, and encouraging, the use of the target language in contextualized situations.

The textbooks were first published in 1971 and have gone through improvements and refinement throughout the years, including the addition of extra materials and resources to be used in class. The Interchange English program is currently in its fifth edition and in the textbook website[22] teachers and students have access to a range of materials in the digital format, including audio files that accompany the activities in the textbook. Extra worksheets are available for further practice concerning grammatical aspects, listening activities worksheets, vocabulary lists, and project worksheets. In addition, there are videos available on the website (one video for each unit from the textbook) to be also incorporated into the classes for further exposure to the target language, in DVD format and also on a student self-study website that provides activities for learners to further practice and review the language worked in class.

Although there is a range of audiovisual materials that accompany the Interchange textbooks, the analysis of the Interchange audiovisual materials dealt with videos from the four textbooks that are available on the Interchange 5th Edition website. Once each textbook is organized in sixteen units, and each unit has a corresponding video to be used in class, a criterion for the selection of the videos for the analysis was needed. In order to better accommodate the number of videos *vs.* the number of textbooks available, the criteria followed was based on the idea of taking the two opposite extremes, that is, the first and last textbooks of the program (Interchange Intro and Interchange 3), two videos were selected from each, also from the first and the last units. The aim of this criterion is also to present a contrast in the

[22] https://www.cambridge.org/br/cambridgeenglish/catalog/adult-courses/interchange-5th-edition

audiovisual materials that are part of the textbook for beginner English learners and the textbook for more experienced English learners. The two other videos selected from the middle of the Interchange English Series, which were Interchange 1 and 2, followed a criterion of selecting the videos from the middle of the two textbooks, that is, from Unit 8 in each one. The result is described as follows.

a) From the Interchange Intro textbook, the videos analyzed were the one corresponding to Unit 1 and Unit 16;
b) From the Interchange 1 textbook, the video analyzed was the one corresponding to Unit 8;
c) From the Interchange 2 textbook, the video analyzed was the one corresponding to Unit 8;
d) From the Interchange 3 textbook, the videos analyzed were the ones corresponding to Unit 1 and Unit 16.

For a better visualization of the videos selected from the Interchange textbooks, the information is presented in the following chart.

Chart 3 – VideosAnalyzed from the Interchange Texbooks

Textbook	Videos Corresponding to Units
Interchange Intro	1 (One) Video from Unit 1[23]. 1 (One) Video from Unit 16[24]
Interchange 1	1 (One) Video from Unit 8[25].
Interchange 2	1 (One) Video from Unit 8[26].
Interchange 3	1 (One) Video form Unit 1[27]. 1 (One) Video from Unit 16[28].

Source: The researcher (2022)

After the analyses, for the final version of this Doctoral Dissertation, two analyses were removed, resulting in a number of ten (n=10) videos, as shown in the following chart.

Chart 4 – Analyses of Videos Included in this Doctoral Dissertaion

PNLD Textbooks	Videos Corresponding to Grades
It Fits	1 (One) Video from the 6th Grade. 1 (One) Video from the 7th Grade.
Become	1 (One) Video from the 7th Grade.

[23] https://www.cambridge.org/files/3115/9299/9817/Interchange_video_Intro_Unit_1.mp4
[24] https://www.cambridge.org/files/6215/9300/5267/Interchange_4_Intro_Level_Unit_16_Video.mp4
[25] https://www.cambridge.org/files/2315/9300/8222/Interchange_4_Level_1_Unit_8_Video.mp4
[26] https://www.cambridge.org/files/5415/9301/2486/Interchange_4_Level_2_Unit_8_Video.mp4
[27] https://www.cambridge.org/files/7415/9301/4127/Interchange_4_Level_3_Unit_1_Video.mp4
[28] https://www.cambridge.org/files/6015/9301/6777/Interchange_4_Level_3_Unit_16_Video.mp4

	1 (One) Video from the 8th Grade.
Time to Share	1 (One) Video from the 8th Grade.
Interchange Textbooks	Videos Corresponding to Units
Interchange Intro	1 (One) Video from Unit 1. 1 (One) Video from Unit 16
Interchange 1	1 (One) Video from Unit 8.
Interchange 2	1 (One) Video from Unit 8.
Interchange 3	1 (One) Video form Unit 1.

Source: The researcher (2022)

The first video, Video from the 9th Grade from the textbook Time to Share, was excluded for its design is particularly similar to the previous one analyzed, possible for the fact of being from the same textbook, thus, produced by the same designers. The second video, Video from Unit 16 from the textbook Interchange 3 was excluded for being an excerpt from a television show, thus not being a pedagogically designed video. Hence, for the analysis to be carried out, a total of ten videos (n=10) were selected: 5 videos from the PNLD materials, and 5 videos from the Interchange English Series. Next, the process of the framework development for the analysis of the videos is explained.

3.3 THE DEVELOPMENT OF THE VIDEO ANALYSIS FRAMEWORK

Grounded on the principles for multimedia design discussed in previous studies (MAYER, 2009; FARÍAS *et al.*, 2013; HAYATI; MOHMEDI, 2010; IBRAHIM, 2012), a framework for the analysis of videos as multimedia materials for the teaching and learning of English as an L2 was developed. The video analysis framework is structured into three major principles which help in the analysis of the videos: a) Signaling; b) Segmenting; and c) Coherence, based on the SSW principles design proposed by Ibrahim *et al.*, (2012).These three major blocks consider different approaches to the design of videos, thus, working as a guide for the analyses. In addition, the nine remaining principles proposed by Mayer (2009) are briefly described in the Discussion section as to whether they can be indetified in the videos designed. The framework that guided the analysis is presented in the following chart.

Chart 5 – Video Analysis Framework

Signaling		
Signaling with Written Language (Typographical Cues)	Does the video make use of subtitles?	Are the subtitles interlingual?
		Are the subtitles intralingual?
	Does the video present written pieces of information?	Does the video make use of different word fonts?

		Does the video make use of different word sizes?
		Does the video make use of different word colors?
		Is the written information positioned in a specific location on the screen to make it salient?
		Does the video present the written language to direct learners' attention to the relevant information?
		Does the video present the written language in order to direct learners' attention and recapitulate the information presented?
Signaling with Oral Language	Does the video present information by narration?	Is there the use of different intonations?
		Does the narrator call learners' attention to the information to come?
		Does the narrator explicitly call learners' attention to specific pieces of information?
		Does the narrator call learners' attention in order to recapitulate the information presented?
	Does the video present information by dialogs?	Is there the use of different intonations?
		Is there the use of conversational strategies, such as repetition?
		Is there the use of conversational marks?
Signaling with Visual Aids	Does the video make use of imagery aids?	Does the video present images/motion pictures/graphs to make specific pieces of information salient?
		Are the images/motion pictures/graphs positioned in a specific location on screen to make it salient?
		Does the video present visual effects to make specific pieces of information salient?
Segmenting		
Segmenting with Written Language	Is written language used in the segmentation design?	In beginning the next segment?
		In ending a segment?
	Do the segments aim at presenting and aiding the comprehension of the written target language structure by managing the amount of information being presented?	
Segmenting with Oral Language	Is oral language used in the segmentation design?	For introducing the next segment?
		For ending a segment?
	Do the segments aim at presenting and aiding the comprehension of the oral target language structure by managing the amount of information being presented?	
Segmenting with Visual Aids	Are visual aids used for introducing the segments?	Are images used?
		Are editing effects used?
	Are visual aids used for ending the segments?	Are images used?
		Are editing effects used?
Coherence		
Coherence Related to Written Language	Does the video present irrelevant written language?	Can the information be considered to hurt learning?

	Is the written language presented essential for the comprehension of the information?	
Coherence Related Oral Language	Does the video present irrelevant oral language?	
	Is the oral language presented essential for the comprehension of the information?	
Coherence Related to Visual and Aural Aids	Does the video present irrelevant visual information (images/motion pictures/graphs)?	
	Does the video present irrelevant visual effects?	
	Does the video present irrelevant music as soundtrack?	
	Does the video present irrelevant noises as soundtrack?	

Source: The researcher

Each of the 10 videos were analyzed following the questions from the SSC principles analysis framework presented in the chart above. Accompanying the analyses, a discussion will be afforded as to the features each one of the ten videos present in their designs as well as whether the features can be argued to aid the process of learning English as an L2, based on the three research questions presented in the Introduction chapter.

4 ANALYSES OF THE VIDEOS AND DISCUSSION

This chapter presents the analyses and discussions of the 10 videos selected according to the criteria described in the Method chapter, organized in two main sections, 4.1 Analyses of the Videos and 4.2 Discussion. Section 4.1 is organized in two subsection: 4.1.1 Videos from the PNLD Textbooks, and 4.1.2 Videos from the Interchange Textbooks. Section 4.2 Discussion, is organized in four subsections: 4.2.1 Discussing Signaling, 4.2.2 Discussing Segmenting, 4.2.3 Discussing Coherence, and finally, 4.2.4 A Few Words on Other Principles.

In the first subsection, 4.1.1 Videos from the PNLD Textbooks, the analyses of 05 videos from the PNLD textbooks are presented. Subsubsections 4.1.1.1 and 4.1.1.2 deal with the analyses of two videos from the textbook It Fits, from the 6th and 7th grades respectively. Subsubsections 4.1.1.3 and 4.1.1.4 present the analyses of two videos that are part of the textbook Become aimed at the 7th and 8th grades, followed by subsubsection 4.1.1.5 presenting the analysis of one video that is part of the material accompanying the Time to Share textbook, from the 8th grade.

The second subsection, 4.1.2 Videos from the Interchange Textbooks, deals with the analyses of 05 videos from the Interchange textbooks. In the subsubsections 4.1.2.1 and 4.1.2.2, the analyses of two videos from Interchange Intro textbook are presented, one from unit 1 and one from unit 16, respectively. In subsubsection 4.1.2.3, the analysis of one video from the Interchange 1 textbook, unit 8 is provided. Subsubsection 4.1.2.4 presents the analysis of one video from the Interchange 2 textbook, unit 8. Finally, the subsubsection 4.1.2.5 provides the analysis of one video from the Interchange 3 textbook from unit 1.

Each analysis will begin by presenting general description of the video in order to contextualize the language features it aims at developing, followed by the analysis of its design features based on the SSC video analysis framework proposed. The analyses of the ten videos are followed by a discussion regarding the three research questions that guided this Doctoral Dissertation, to be presented in section 4.2.

4.1 ANALYSES OF THE VIDEOS

4.1.1 Videos from the PNLD Textbooks

In this subsection, the five subsubsections providing the analysis of one video each that are part of the three selected textbooks (described in the Method chapter) approved by the

PNLD Program are presented. Each analysis aims at providing a detailed examination of the features that are part of the videos design and how these features can aid the process of learning an L2.

4.1.1.1 Video from the Textbook It Fits – 6th Grade

The video from It Fits 6th grade textbook, which is entitled Introducing Yourself, has the length of two minutes and forty-nine seconds. The context the video presents depicts two children having a video call in which they introduce themselves and provide information about their families, friends and where they live. The design of video follows a cartoon-like fashion and presents the characters as sitting in front of the screen, suggesting they are seeing each other, resembling video calls in real life. As follows, the video will be analyzed as to its design features according to the SSC framework.

Signaling

Signaling could be observed in this video in a number of ways: a) as typographical cues; b) as oral language to signal (or wrap-up) what was presented in the video; and c) as visual elements. All devices are discussed as follows.

Concerning the typographical cues, the first written language input (PLASS; JONES, 2005) element that the video makes use of is the presentation of the title Introducing yourself. The presentation of the title may work as a signaling device once it offers learners the sense of structure, in which the title of a text, or video is normally expected to come in the center of the page (in case of a written text) or in the center of the screen (in the case of a video). Moreover, the title is positioned in the center of the screen, possibly to direct learners' attention directly to it and to aid the process of apperceiving[29] (PLASS; JONES, 2005) topic and language that will be presented in the video and of activating prior knowledge which, according to Ambrose *et al.* (2010), can aid the process of learning. It is important to clarify that the fact of activating learners' prior knowledge cannot go beyond an assumption in this analysis, for the fact that learners' prior knowledge activation would depend on whether they had worked with this topic previously to watching the video.

Also as regards the typographical cues that can be defined as a signaling device presented in the design, the video makes use of subtitles presented in the color yellow. The

[29] Based on Plass and Jones (2005), *apperception* can be referred to the process in which learners direct their attention to certain pieces of information being presented, and select the ones they judge to be crucial for the understanding of the target information.

choice of presenting the subtitles in this color may be of two reasons. The first one may be because, in general, most subtitled videos and films present their subtitles in the color yellow, suggesting that the video design intends to follow such an audiovisual pattern. The second one (and most important) is that, in most part of the video, the important pieces of information are located in the center of the screen, leaving the boarders with a darkened color, especially the bottom part of the screen where the subtitles are located, making the subtitles more salient. Such saliency effect, according to Lorch (1989), and Koning *et al.* (2010) can work as directing learners' attention to the intended information. Furthermore, although the font used for the subtitles can be considered small, they can be visible for learners to read the information presented. In addition, the kind of font used is one that presents the letters in a clear way, with no use of stylish or embellishment design, which guarantees learners with a minimal effort for the reading and decoding of the written language presented. Thus, the contrast of the color yellow, and the use of an easy-to-read font make the subtitles more visible and readable, as can be observed in Figure 8.

Figure 8 – The Subtitles in it Fits 6th Grade Video Introducing Yourself

Source: It Fits Website (2019)

The type of subtitles can be characterized as intralingual, that is, according to Danan (2004) and Lavaur and Bairstow (2011), they are presented in the same language as the oral language spoken by the characters. One reason for presenting subtitles is the fact that, according to one of the characters (Peter), he is American, thus, speaking English with a rather fast pace and with an accent that resembles a non-Brazilian speaker. The insertion of the subtitles can aid the process of comprehension, once, at times his speaking may be considered too fast for

learners to follow (considering that this video is aimed at 6th grade learners), resulting in poor comprehension. The function of the subtitles, thus, can be argued to help learners in the process of comprehending the spoken language, that is, the information being presented orally, as well as of apperceiving the target structure of the language that they (probably) will be expected to use in a later moment. According to Plass and Jones (2005), learners may apperceive the target structure due to the fact that they direct their attention to the information and need to select the parts of it they judge to be important. Although the video presents the target language in the written form in two ways – the title and the subtitles – there are no other typographical devices used in order to signal specific linguistic aspects.

Concerning signaling information presented orally, because of the context of the video, that is, a conversational situation where two children are talking, there is no use of narration, but an interaction the characters perform. Furthermore, there is no observable change in intonation in order to signal, or in order to direct learners' attention to a specific piece of information from the video. Once they are having an informal conversation, the way in which the spoken language is presented follows a conversational style. One reason for the lack of oral signaling can be due to the presentation of the target language in the form of subtitles, presenting learners with the target language in the two modes (oral and visual).

Another aspect concerning the use of oral language, which works as a signaling device, can be observed at the ending of the video. When the dialog finishes, there is a change on the target language information presented. First, the information is presented in the form of narration, once the image of the person talking is not visible on screen. Second, the information presented makes a wrap-up of the video, saying that now they already know how to introduce themselves. Such oral design, as explained by Hasler, Kersten and Sweller (2007) and Koning *et al.* (2010), is an aid to call learners' attention to the topic of the video. Third, an instruction is presented orally asking learners how they would introduce themselves and talk to people they know to show that they have learned how to use this target language. Fourth, however the video is aimed at developing the language knowledge of how to introduce themselves in the target language, that is, in English, the language is presented in Portuguese (learners' first language). Presenting the narration in learners' first language can be to certify that learners understand the instruction presented by the end of the video. Throughout this last part of the video, subtitles are still presented, yet, in Portuguese to follow the spoken language.

Regarding the visual elements signaling specific pieces of information, it is possible to argue that the video Introducing yourself presents a number of elements that help direct learners' attention to what/who the characters are talking about. The aim can be to aid the

process of apperception of the linguistic input, as explained by Plass and Jones (2005). Once the conversation represents a video call, the characters make use of pictures and an arrow resembling the computer cursor to signal, that is, to direct learners' attention to what and who they are talking about. In the case of imagery elements, information is presented as they click and show pictures of the place they live on a map, the members of their families, the school, teachers and pets. When one character finishes talking about her image that was at first only appearing on screen, her figure becomes smaller appearing side by side with the image of the other character, suggesting that her turn to speak and present her personal information is open.

Visual aids, as explained by Schmidt (1990) and Koning *et al.* (2010) can have direct impact in directing learners' attention. In the case of the video, another visual aid that is used to direct learners' attention to the elements they are talking about in the photographs is an arrow that resembles the mouse cursor that appears in every computer screen. As the characters talk about and describe the elements, the arrow moves about the screen and points at the elements they are describing, as presented in Figure 9.

Figure 9 – Visual Signaling Devices in it Fits 6th Grade Video Introducing Yourself

Source: It Fits Website (2019)

As can be observed, the visual elements help direct learner' attention so that they do not focus on unnecessary visual elements, language features or search of the correct elements to associate to the information being presented orally. According to Sweller (2017), Mayer (2009) and Kalyuga (2010), such design can function as leaving working memory resources for the processing of the essential language elements to be comprehended.

The signaling devices that could, thus, be examined in the Introducing Yourself video are a) the use of typographical cues, as is the case of the use of a title and subtitles in the video design; b) oral language in wrapping-up the video and proving learners with instruction for practice the language learned; and c) visual elements, which aid the process of directing learners' attention to the people the two characters are talking about.

Segmenting

Segmenting could be seen in the videos in various ways: a) orally with the utterances produced by the characters, and visually, when they present pictures of the people/elements they are describing; b) with the aim of providing learners with small (and repeated) portions of the target language structure and vocabulary used for presenting people as to their names, ages, occupations and how such target language is related to the people introducing them; and c) as a closing with instructions for the use of the language presented. The segmentation structure analysis is presented as follows.

Although the video Introducing Yourself is organized in what Leahy and Sweller (2017) define as a transient fashion once it presents a conversational situation, in which two children are interacting, based on arguments of Sweller (2010), Ibrahim (2012), and Ibrahim *et al. (*2012), it follows a segmentation design that can be observed to be organized in 20 segments, or blocks. The first segment presents the title of the video, and the second starts by establishing the context of the dialog being present, that is a video call between two children. In the following segments (3 through 19) the target language about introductions is presented, as is shown in Chart 6.

Chart 6 – Segmentation Structure in it Fits 6th Grade Video Introducing Yourself

Segments	Information Presented
Segment 1	Introducing yourself
Segment 2	Computer screen is presented showing a video call being started.
Segment 3	**Peter:** Hello, I'm Peter. I'm twelve years old and I want to be your friend! What's your name? **Isabela:** Hi Peter. I'm Isabela. I'm twelve too and I'm from Brazil.
Segment 4	**Peter:** I'm from the United States and I live here, in the state of California. I'm happy you are my new Brazilian friend!
Segment 5	**Peter:** I'm going to introduce you to my family, friends and teachers. This is my brother, Phillip. He's fifteen years old.

Segments	Information Presented
Segment 6	**Peter:** This is my best friend, Sarah, she's eleven years old.
Segment 7	**Peter:** This is my school. It's in San Francisco, California. And this is the Pacific Ocean.
Segment 8	**Peter:** These are my school teachers And they're very cool! I love all of them. My favorite one is Mr. Johnson. He's an English teacher And he's very funny.
Segment 9	**Peter:** Let me show you my parents. This is my father Joseph. He's thirty six years old and he's a salesperson.
Segment 10	**Peter:** This is my mother, Samantha. She's thirty five and she's a teacher.
Segment 11	**Peter:** This is Barney, my dog. He's four years old and I play with him every day.
Segment 12	**Peter:** You're my only friend from Brazil. Tell me about you, Isabela.
Segment 13	**Isabela:** I'm from Natal, Rio Grande do Norte.
Segment 14	**Isabela:** I have a big family. This is my sister Andrea. She's eighteen years old.
Segment 15	**Isabela:** These are my brothers. This is Gustavo. He's fourteen years old. The other one is Jonas, he's sixteen.
Segment 16	**Isabela:** These are my parents. This is my father, Luciano. He's thirty nine and he's a dentist.
Segment 17	**Isabela:** This is Sheila, my mother. She's thirty seven and she's a chef.
Segment 18	**Isabela:** This is my family. We're a very happy family!
Segment 19	**Isabela:** I'm happy we are friends now, Peter. **Peter:**

Segments	Information Presented
	Me too. Let's talk again soon. **Isabela:** Ok, see you tomorrow! **Peter:** Bye-bye!
Segment 20	**Narrator:** Agora que você já viu como se faz, quero saber: Como você se apresentaria? Fale com seu professor, seus amigos, pais, irmãos, tios, avós e mostre que você já sabe se apresentar em inglês.

Source: The Researcher (2021)

As can be observed in the segmentation design, each of the segments presents certain utterances that are associated to introducing not only the characters themselves, but other people, such as their family members, teachers and pets as well as providing information about them.

The introduction of the segments and transition between the segments is made in two ways: orally with the utterances produced by the characters, and visually, when they present pictures of the people/elements they are describing. One example of the introduction of a segment by means of an utterance produced can be observed in segment 9. It is introduced when the character Peter says "Let me show you my parents."(1'10). This utterance can be observed to establish a link to the next topic to be explored, that is, the parents, ultimately to the next segment. In addition to the oral introduction of the segment, the character clicks on a link on his screen and a picture of two people (his parents) is presented.

The visual elements (pictures) also represent the transition between each segment, once when the characters finish talking about the people in the pictures being shown, they click again on the link, making the picture being shown move away from the screen to make room for a different picture to be presented.

It seems that the video is organized in segments in order to present learners with the target language in smaller amounts, which, according to Koning *et al.* (2007), and Ibrahim, *et al.* (2012), would benefit learning. One example for this is the segment 3, as is presented.

(Segment 3)

Peter:
Hello, I'm Peter.
I'm twelve years old
and I want to be your friend!
What's your name?
Isabela:
Hi Peter. I'm Isabela.
I'm twelve too and I'm from Brazil.

Segment 3 can be defined as aiming at providing learners with the target language used in the exchange of personal information about name, age and place of origin. The first evidence can be observed when Peter produces the utterance "I'm Peter. I'm twelve years old" (0'21) providing learners with input exemplifying the target language used to refer to one's name and age. In addition, the exchange of personal information can be further observed when he asks "what's your name?" suggesting that in a personal information exchange between two strangers introducing each other, the names and ages are likely to be the first pieces of information to be talked about, especially when the participants are children. According to Plass and Jones (2005), the oral design can help learners in the apperception of the target language, in this case associated to personal information. In addition to presenting their names, the participants also talk about their places of origin, as is presented in the next segment.

(Segment 4)

Peter:
I'm from the United States and I live here, in the state of California.
I'm happy you are my new Brazilian friend!

In segment 4, learners are provided with a rather limited amount of language, which, as explained by Ibrahim, *et al.* (2012) can benefit learning. However, it is possible to see that the language is concise enough to present learners the language input used to talk about one's place of origin, such as the state and the country, working as the necessary input of the target language to be learned, as explained by Plass and Jones (2005). The second utterance Peter produces can also play an important role. The sentence "I'm happy you are my new Brazilian friend!" (0'37) can emotionally engage learners, once the child Peter is talking to (Isabela) is from Brazil, and learners may become more involved because of her background.

The next segments are aimed at presenting learners with examples of the target language used when introducing a third person. In the context of the video, Peter and Isabela are providing information about their family members (Peter and Isabela) and about friends and teachers (Peter), as can be observed.

(Segment 5)

Peter:
I'm going to introduce you to my family, friends and teachers.
This is my brother, Phillip.
He's fifteen years old.

(Segment 8)

Peter:
These are my school teachers
And they're very cool!
I love all of them.
My favorite one is Mr. Johnson.
He's an English teacher

And he's very funny.

As discussed, the introduction of the segments in the video Introducing yourself are made visually and orally. In the case of the oral language produced by the characters, segment 5 exemplifies the introduction of a new segment when Peter produced the utterance "I'm going to introduce you to my family, friends and teachers." (0'41) and "Let me show you my parents" (1'10). The oral language presented on the two utterances can be claimed to have two important roles: the first one is about introducing and making the transition to the next segment. The second role can be associated to the first one, that is, their function can be defined not only as making the transition to the next segment, but most importantly, making the transition to the next language segment. When Peter says that he will introduce the people who part of his social circles, this leaves room for learners to prepare for the information that will come, as well as provide them with time to process the new language structures or vocabulary input, as discussed by Zarei and Khazaie (2011). Such processing time is argued by Sweller (2010), Mayer (2009) and Kalyuga (2010) to be crucial for the effective processing of the information.

The language structure used to talk about and provide information about family members is also presented in the segments that follow segment 8. The target language is exemplified in the segments that follow.

(Segment 9)

Peter:
[…]
This is my father Joseph.
He's thirty-six years old
and he's a salesperson.

(Segment 10)

Peter:
This is my mother, Samantha.
She's thirty five and she's a teacher.

(Segment 11)

Peter:
This is Barney, my dog.
He's four years old
and I play with him every day.

(Segment 14)

Isabela:
I have a big family.
This is my sister Andrea.
She's eighteen years old.

(Segment 15)

Isabela:
These are my brothers.
This is Gustavo.
He's fourteen years old.

The other one is Jonas, he's sixteen.

(Segment 16)

Isabela:
These are my parents.
This is my father, Luciano.
He's thirty nine and he's a dentist.

(Segment 17)

Isabela:
This is Sheila, my mother.
She's thirty seven and she's a chef.

As can be observed in segments 9 through 17, based on Zarei and Khazaie (2011), the aim is to present learners with vocabulary about family members, (for example, brother, sister, father, mother) jobs, (as is the case of dentist and chef) and present the grammatical structure to be used to provide information about them, such as the names and the ages and what they do preceded by the utterance "this is…". This can be observed, for example, in segment 15, when Isabela produces the utterance "This is Gustavo. He's fourteen years old." (1'48). Such segment presented the language structure necessary when one wants to provide information about a family member, in this case, about a brother. Moreover, in the second part of the utterance, learners are also provided with another example of how to talk about the age of a third person. Segments 16 and 17 present the target language grammatical structure in a similar way. However, they present learners with an extra piece of information, that is, about the parents' jobs. When the utterances "This is my father, Luciano. He's thirty nine and he's a dentist." (2'48) and "This is Sheila, my mother. She's thirty seven and she's a chef." (2'05) are produced, the first target language feature is that they are very similar in their structure, following the same pattern. Once, according to Kalyuga (2010) working memory resources should not be exceeding while processing information, it can be defended that the utterances have been designed this way in order not to overload learners' working memory (WM) resources by presenting them with the relevant pieces of information and grammatical structure in order to aid the process of attention and retention. Another important aspect is that, once the two structures (talking about ages and jobs) can be considered concise in the information presented, they leave room for learners to process the news about Isabela's parents' jobs and, as explained by Plass and Jones (2005), to apperceive the target language by assimilating it to the imagery information provided.

Finally, the last segment that is part of their interaction provides learners with the target language that can be used to say good bye to people. The utterances produced "Let's talk again soon.", "Ok, see you tomorrow!" and "Bye-bye!" (2'18) have the function of ending the

interaction between Peter and Isabela and also presenting learners with the target language that can be used to end a conversation, as is the case of the utterance “I’m happy we are friends now, Peter.” And present three different ways of saying goodbye, as can be observed in the following segment.

(Segment 19)

Isabela:
I’m happy we are friends now, Peter.
Peter:
Me too. Let’s talk again soon.
Isabela:
Ok, see you tomorrow!
Peter:
Bye-bye!

The last segment presented in the video does not follow the design of the conversational situation presented. After the video call is ended, the two conversation windows fade out and a narrator presents instructions, as below.

(Segment 20)

Narrator:
Agora que você já viu como se faz,
quero saber: Como você se apresentaria?
Fale com seu professor, seus amigos,
pais, irmãos, tios, avós
e mostre que você já sabe
se apresentar em inglês.

The last segment of the video Introducing Yourself is aimed at making a wrap of the language worked. Furthermore, it is arguable that the utterances “*Como você se apresentaria? Fale com seu professor, seus amigos, pais, irmãos, tios, avós e mostre que você já sabe se apresentar em inglês.*” (2’35) are produced aiming at encouraging learners to use the target language presented in the video productively to talk not only to the teacher and colleagues, but also, to bring the language to contexts outside of the classroom, in this case, to their families. Another difference of segment 20 compared to the others is that it presents the narration in Portuguese, that is, in learners’ L1. One possible reason for this is that only the segments that presented the target language structure and vocabulary were in the target language, that is, in English. As segment 20 presented a recall of the language learners were exposed to and were invited to use it with the people from their social circles, Portuguese may have been the choice of narration to make sure learners would comprehend the instruction of talking to their acquaintances.

Another important element is that, because of the transient nature of the information being presented in the video (LEAHY; SWELLER, 2017), learners are provided with the play/pause button, so that they can have control over the information flow, pausing or going

back to any part of the video they judge to be necessary to review. It is important to consider, however, that the play/pause button can be used by learners when they are watching the video outside the classroom. When the video is used as an in-class resource, it is under the teacher's control.

One reason that can be argued for the design of this video to be in smaller chunks is that, in each segment, learners will be presented with a reduced amount of language, based on Mayer (2009) and Sweller (2010), leaving working memory sources for them to apperceive (PLASS; JONES, 2005) and process the information being presented. Although the segments are not presented in clear-cut fashions, the segments can be argued to be organized in a clear format that in each one focus is given to the structure to be used when introducing a third person, using the correct form of the verb BE (grammar) and words to refer to occupations and family members (vocabulary).

The segmentation design of the video, thus, could be observed to aim at providing learners with the target language structure and vocabulary in small chunks repeatedly as a number of people is introduced. For the fact that the two characters interact presenting the people who are part of their social circles, the grammatical structure is presented for talking about one's age with the verb BE, and the vocabulary necessary for providing information about their occupation and/or presenting people as family members.

Coherence

As discussed in the previous chapter, the coherence principle can be defined as aiming at preventing learners from engaging in extraneous processing, that is, coherence concerns with eliminating unnecessary pieces of information that may disturb the learning of the target language presented in the video.

As regards the language presented in the written form, the only written language presented is in form of subtitles. Once they follow the oral utterances that are presented in the video, the subtitles can be considered essential, once they may assist the comprehension of the target language that may not be clear only by listening to the characters producing it, as discussed by Danan (2004), Hayati and Mohmedi (2010), and Lavaur and Bairstow (2011). As has been discussed based on Plass and Jones (2005), in the case of learning English as an L2, using videos as a form of language input which provide subtitles can be argued to have positive effect in the process of comprehending the language presented, especially because the video Introducing Yourself is part of the textbook aimed for beginners who are in the 6th grade.

Concerning the oral language presented in the video, it can be considered essential for the comprehension of the target language. Considering that the video presents a conversational

situation, the target language is not only restricted to utterances presented as isolated chunks to be used when talking/introducing one's family/teachers, but also presenting conversational strategies for introducing a topic, or to present a certain language structure to be used. This can be observed when, in segment 5, the character Peter produces the utterance "I'm going to introduce you to my family, friends and teachers." (0'42), suggesting that Peter is preparing his hearer (in the case of the interaction, Isabela, and in the case of the video, the learners) for the language structures he will use from that moment forward, that is, to use language to provide information about his family, teachers and friends.

Regarding the visual effects, the ones that become more evident in the video are the transitions between the pictures the characters Peter and Isabela present. As they introduce the members or school teachers, the pictures jump in the screen, and as they starting talking about different people, the pictures jump off the screen. This transition can be considered important, due to its attention directing (SCHMIDT, 1990; LORCH, 1989; KONING *et al.*, 2010; HASLER; KERSTEN; SWELLER, 2007) nature, as discussed in the signaling section based on Lorch (1989), Schmidt (1990), Hasler, Kersten and Sweller (2007) and Koning *et al.* (2010). The sound effects in the design of the video can be considered reduced because they are limited to the sounds of a keyboard typing when the login and password are filled in on the screen and the sound of a video call being started, which, in the context of a video call, seem to be complementary. In addition, when the characters click on the new image button, no sound is produced. One possible explanation for this is that the sound effect of the clicking would not aid the process of comprehending the target language of the video, leaving room for the processing of the essential oral information, preventing that the two types of aural information (the oral language + the sound effects) compete for the working memory resources (KALYUGA, 2010; MAYER, 2009; SWELLER, 2010), thus assisting learners in the generative processing.

4.1.1.2 Video from the Textbook It Fits – 7th Grade

The video analyzed from the It Fits 7th grade textbook, entitled Things They Can or Can't Do!, has the length of three minutes and eleven seconds. The context of the video presented is different from the video from the 6th grade analyzed once it does not present a conversation situation in which people are interacting. The video follows a cartoon-like design in which learners are presented with the information by means of a narrator introducing two children and providing information about the skills they have using the target language to be

learned as shown in the analysis of the video concerning its design features based on the SSC framework.

Signaling

Regarding the typographical cues presented in the video, the first written language input is presented in the title of the video Things They Can or Can't Do!. The title can be claimed to work as a signaling device once it aims at, besides presenting the title of the video, directing learners' attention to the topic (language) the video will present, once the part of the title "can or can't do!" is presented in boldface, according to Lorch (1989) and Koning *et al.* (2010) aiming at making it more salient. Furthermore, the title is positioned at the center of the screen, following the structure that normally a written text would be expected to have, in this case, the video.

Still about the use of written language as a signaling device, the video makes use of subtitles presented in the color yellow. The choice for using the subtitles in this color can be for two reasons: the first reason can be that the video is intended to present the subtitles following the design of regular films so as to provide them with a rather cinematic fashion. The second reason can be associated to the color of the background. As the whole video is designed presenting a green background, the use of the subtitles in the color yellow can be defended to make them more salient, thus easier to read by learners, as presented in Figure 10.

Figure 10 – Subtitles in it Fits 7th Grade Video Things they Can or Can't Do!

Source: It Fits Website (2019)

The use of subtitles in the design of the video can also be claimed to have an important role concerning the language presented. One important aspect of the subtitles to take into consideration is that they are presented in the same language presented orally in the video, falling in the category of intralingual subtitles, as brought by Danan (2004), Hayati and Mohmedi (2010), and Lavaur and Bairstow (2011). One important affordance that presenting intralingual subtitles in the video design can have is to aid learners in the comprehension process. As the oral language is presented in a pace that may be difficult for learners to follow, and the vocabulary and language structures presented can present difficulty for learners to understand what is being said, the subtitles can compensate for the lack of understanding the oral information, thus integrating the written language to the oral language.

Regarding the size and the fonts used, although the size of the subtitles can be defined as being of a small size, they can be considered easy for learners to read. Also regarding the font used, the subtitles are presented on a visually clear font, with no use of stylish or embellishment design, so to demand learners with a low reading effort.

The use of subtitles, however, is not the only typographical device to signal the language being presented in the video. When the narrator starts the video, she introduces the two characters whose information will be provided about by saying their names “Daniel” and “Charlotte”. As the names of the characters are presented orally, their names are presented in the written format above their images. In addition, there is the presentation of on-screen written language throughout the video with the target language structures, as is presented in Figure 11.

Figure 11 – Typographical Signaling Devices in it Fits 6th Grade Video Things they Can or Can’t Do!

Source: It Fits Website (2019).

As the title presented in the video suggests, the context of the video deals with language used to provide information about skills people may have regarding what they can do and cannot do; questions and answers are presented about the two characters concerning different skills that are discussed and contextualized before presenting the questions and answers. The presentation of on-screen target text aimed at directing learners' attention to the target language structure in the written form can work as a device aiding the process of apperception (PLASS; JONES, 20050) of the language structure used to refer to skills.

Detectable due to the technical design, the font size, colors and type of the on-screen target text questions and answers are different from the subtitles presented. As to the size, it is possible to observe that the questions and answers are presented in a size bigger than the subtitles. According to Lorch (1989) and Schmidt (1990), the choice for such design features can be associated to the saliency aspect, in which learners' attention is directed to the target language structure in order to aid the apperception process after being exposed to the input (PLASS; JONES, 2005). Regarding the color of the questions and answers, they are presented in two colors: the questions are presented in black and the answers in white. One possible reason for that may be to direct learners' attention to how a question is structured, in which the modal "can" comes before the subject, and how the short answer is structured, in this case "yes, she/he can." or "no, she/he can't.". Regarding the types of fonts, it is observable that the types of the fonts used in the subtitles and in the questions and answers are different. A possible argumentation for such can be based on the fact that they can have different purposes. The subtitles are intended to provide an affordance to compensate the possible lack of comprehension of the oral language presented, whereas the questions and answers are intended to present learners with the target language structure to be used to talk about people's skills.

Further about directing learners' attention to specific pieces of information by the use of written language as linguistic input, as argued by Plass and Jones (2005), in this case, the presentation of questions and answers, the positioning of the questions and answers on the screen can be said to have an effect on directing learners' attention. As can be observed in Figure 11, both the questions and answers are positioned on the higher part of the screen. In addition, when the questions and the answers are compared to each other in regards to their location on screen, the questions are positioned in a higher level, possibly, based on Plass and Jones (2005), aiding learners in the apperception process of understanding the hierarchical organization of the language structure in which the question is presented first, followed by the answer.

Regarding the use of oral language as signaling devices, it is possible to observe that the video is designed in a fashion that does not provide learners with a conversational situation where the characters are interacting. Instead, the video presents the two characters (as described earlier) and a narrator provides learners with information about them concerning the activities they can and cannot perform.

Moreover, the use of intonation in order to signal important pieces of information is also evident in the video design. In the beginning of the video, the narrator introduces the two characters and contextualizes the information that will be presented by defining them as two "active teenagers" and that "every day they try to do something different". Following this introduction, the narration calls learners' attention to the activities that the two characters can and cannot do, aiding the apperception process of the language structure (PLASS; JONES, 2005). At this moment, when the utterance "Let's see what they can (…) and what they can't (…) do." (0'18) is produced, a micro pause can be detected after the words "can" and "can't" in the narrator's voice. One argumentation for the use of the micro pauses after these words can be associated to the topic that the video presents. Once the target language presented refers to what people can and cannot do, which is also clearly presented in the title of the video, the main language items are the words "can" in order to describe an activity that one is able to perform and "can't" to describe an activity one is not able to perform. Moreover, as "can" and "can't" inserted in sentences to describe people's abilities are presented throughout the video, directing learners' attention to these language items at the beginning of it can be considered to aid learners in the comprehension of this structure.

The use of visual aids can also be observed in the Things They Can or Can't Do! video. Regarding visual information, the video presents the use of imagery features in order to depict the target vocabulary, that is, the images work as an aid for learners to make connections between the activities being described and the target words to describe the activities they are being able to do or not, as also observable in Figure 11. For example, since the beginning of the video, the character Charlotte is presented wearing black sunglasses and holding a walking stick, presenting evidence that she is visually impaired. When the narrator provides information about Charlotte having been born blind, the target language is presented with the question "Can Charlotte read?" and the answer "Yes, she can." (1'25), learners are (or are expected to) drawn to make connections between the main verb read, which expresses an action, in this case, an activity, and the image being presented of her reading a book using her tactile sense. The imagery features in each of the segments presented in the video (to be discussed in the segmenting section) are described as follows.

Daniel dressed in a swimming suit (segment 3).

Charlotte dressed in a typical Brazilian carnival dress (*porta-bandeira*) (segment 4).

A stone which Daniel tries to move (segment 5).

Charlotte sitting at a table with her hands on a book suggesting she is reading it (segment 6).

Daniel wearing a helmet and gloves and a wall (segment 7).

Charlotte singing and a table with glasses and a juice jar breaking as she raises her voice singing (segment 8).

Daniel holding a plate with a slice of chocolate cake and chewing followed by him wearing an apron and holding a spoon suggesting he is cooking. His head is surrounded by question marks, following the question about him being able to cook (segment 9).

Charlotte wearing a party hat and the images of cakes, sweets and drinks, which can be associated to a party (segment 10).

Daniel standing between a toolbox with a hammer laying next to it, and a door that seems to be broken followed by Daniel with a screwdriver fixing the door (segment 11).

Such design structure presenting images that correspond or depict the activities being presented is part of the whole video. It is possible to argue that the use of the imagery pieces of information are necessary in presenting the target language to learners so to aid the process in which they, according to Plass and Jones (2005), apperceive the target language structure presented and engage in the process of, as put by Mayer (2009), integrating the language they are listening to and reading, and the images they are seeing. This integration, as discussed, can be considered important in the process of building coherent mental representations not only of the activities described in the video (for example, reading, cooking, climbing), but also the target language structure used to describe activities one is able to do or not. Based on Mayer (2009) and Mayer and Moreno (2010), when learners engage in such cognitive processes, the generative load the video can afford them with may result in learning.

Regarding the use of visual effects in order to direct learners' attention to special pieces of information in the video, that is, to help them in the apperception process by means of visual and verbal input as discussed by Plass and Jones (2005), it is possible to argue that the only visual effect that is part of the video design is the fading in and fading out effects when changing the information being provided. These visual effects, however, cannot be defended as being crucial for aiding learning, nor as overloading learners' working memory (WM) resources thus, according to Mayer (2009) being considered irrelevant.

Segmenting

The video Things They Can or Can't Do! can be defined as organized in segments based on Sweller (2010), Ibrahim (2012), and Ibrahim *et al.* (2012), or blocks, each one presenting specific target language. The video is organized in a total of 11 segments, being the first one presenting the title, suggesting learners what language will be presented, as shown in Chart 7.

Chart 7 – Segmentation Structure in it Fits 7th Grade Video Things they Can or Can't Do!

Segments	Information Presented
Segment 1	Things they can or can't do!
Segment 2	**Narrator:** This is Daniel and this is Charlotte. They are very active teenagers. Every day, they try to do something different. Let's see what they can and what they can't do.
Segment 3	**Narrator:** Since Daniel was a little boy, he has always loved going to the swimming pool. Can Daniel swim? Yes, he can.
Segment 4	**Narrator:** Charlotte's parents are mestre-sala and porta-bandeira of a samba school. She is always at the rehearsals. Can Charlotte dance samba? Yes, she can.
Segment 5	**Narrator:** This is a very heavy stone, it weighs 300 kilos. Can Daniel move the stone? No, he can't.
Segment 6	**Narrator:** Charlotte was born blind. she learnt how to do a lot of things with the tactile sense. Can Charlotte read? Yes, she can.
Segment 7	**Narrator:** Daniel enjoys practicing adventure sports. Can he climb this wall? Yes, he can.
Segment 8	**Narrator:** Charlotte loves dancing, but she never learnt how to sing. Can Charlotte sing an opera? No, she can't.
Segment 9	**Narrator:**

Segments	Information Presented
	There is one thing Daniel loves: Eating his grandmother's chocolate cake. Every week, he asks her to bake a cake. He never sees how she prepares it. Can he bake a chocolate cake? No, he can't.
Segment 10	**Narrator:** When Charlotte's family throws a party, it's not her father, or her mother who are in charge of it. She is the one who organizes everything. Charlotte is a very good hostess. Can she organize a big party? Yes, she can.
Segment 11	**Narrator:** Daniel's father is a handyman. He can fix absolutely everything and Daniel is learning with his father. Can Daniel fix this door? Yes, he can.

Source: The Author (2021)

The video Things They Can Or Can't Do! presents the target language in segments, as can be observed in the chart presented above. The title is the first segment presented in the video, and can be defended as being the only segment that makes use of the written language to introduce the next segments. No other segment can be observed as being introduced or ended with the use of written language.

When it comes to the use of oral language input to introduce a segment, it becomes possible to observe that there is the use of the segment introduction with oral language. The only part of the video introducing the segments with the oral language can be identified in segment 2, as presented.

(Segment 2)

Narrator:
This is Daniel and this is Charlotte.
They are very active teenagers.
Every day, they try to do something different.
Let's see what they can
and what they can't do.

The segment presented above can be defined as having two main functions: the first function is to introduce the two characters, and as brought by Innaci and Sam (2017), and Tuyen and Huyen (2019) to start contextualizing the target language to be presented. The second function can be examined based on Plass and Jones (2005) as directing learners' attention to

the target language input to be learned. The contextualization of the target language that will be presented in the video can be observed in the sequence of utterances “This is Daniel and this is Charlotte. They are very active teenagers. Every day, they try to do something different.” (0’14). Such contextualization, when introducing the characters, aims at establishing a social engagement with learners and for them to become acquainted with the two characters – Daniel and Charlotte – the video will present information about. The last utterance produced in the segment “Let’s see what they can and what they can’t do.” (0’18) invites learners to watch situations describing what Daniel and Charlotte are and are not able to do and to continue watching the video, moving to the next segments. In addition, this segment can work as activating learners’ prior knowledge concerning the language used with the words “can” and “can’t”.

The next segments presented in the video design can be argued to follow a similar structure. Considering segments 3 through 11, each one presents a specific activity that Daniel or Charlotte are able to perform or not by the use of the modals “can” or “can’t”. The abilities vocabulary presented in the video are observed as swim (segment 3), dance samba (segment 4), move a stone (segment 5), read (segment 6), climb a wall (segment 8), bake a chocolate cake (segment 9), organize a big party (segment 10) and fix a door (segment 11).

The segments, however, not only present the target language grammatical structures to make questions about what activities people are able to do or not. Instead, each segment presents a contextualization of the activity that will be associated with the character being talked about. The aim is, according to Hasler, Kersten and Sweller (2007) and Koning *et al.* (2010), to direct learners attentional resources to the abilities that are subsequently presented.

(Segment 4)

Narrator:
Charlotte’s parents are mestre-sala
and porta-bandeira of a samba school.
She is always at the rehearsals.
Can Charlotte dance samba?
Yes, she can.

In segment 4, two important aspects of the utterances produced can be observed. The first utterance “Charlotte’s parents are mestre-sala and porta-bandeira of a samba school.” (0’40) presents information about Charlotte’s parents being mestre-sala and the mother being porta-bandeira. This information is clearly associated to a Brazilian context, once it is referred to the carnival where these roles can be seen. Moreover, once the information presented involves a Brazilian context, the role of the father and the mother are implicit, suggesting the

learners can identify who is the mestre-sala, in this case, her father, and who is the porta-bandeira, in this case, her mother.

Furthermore, the utterance that is produced next, “She is always at the rehearsals.” (0’41) can be defined as establishing a clear association between the samba school and Charlotte’s dance skills and the contextualization of the whole picture making it consistent to the fact of asking about her knowing how to dance the samba. After the information is contextualized, the target language structure is, then, presented, in which learners are first presented with the question “Can Charlotte dance samba?” (0’44) and the answer “Yes, she can.” (0’49) showing learners the target language structure in which the modal is used to describe abilities in a clear-cut way.

Another example of how the target language question and answer structure is preceded by a contextualization can be observed in segment 9.

(Segment 9)

Narrator:
There is one thing Daniel loves:
Eating his grandmother’s chocolate cake.
Every week, he asks her to bake a cake.
He never sees how she prepares it.
Can he bake a chocolate cake?
No, he can’t.

The segment presented above, which is concerned with presenting a small portion of information about Daniel being able to bake a cake, contextualized by the use of personal information about the family, more specifically, to the grandmother. Such choice can have an emotional effect engagement, once learners may identify themselves with the food their grandmothers prepare. The utterances “There is one thing Daniel loves: Eating his grandmother’s chocolate cake. “Every week, he asks her to bake a cake.” (2’10) aid the establishment of the context, which is about Daniel’s preference for chocolate cake and asking his grandmother bake one for him. The utterance that is produced next “He never sees how she prepares it.” (2’13) can be argued to create an association between the context of Daniel being fond of chocolate cakes, and being able to cake one. This utterance is, then, followed by the question “Can he bake a chocolate cake?” (2’15) and the answer “No, he can’t.” (2’18), providing learners with the information that Daniel does not know how to bake a cake, and the target language structure using the modals “can” and “can’t” to describe the activities one can do. Once again, it is possible to analyze that the segments aim at providing learners with target language input used to describe the abilities one may have or not by presenting short and straight to the point example questions and answers, as pointed out by Plass and Jones (2005), in order

to aid the process of apperception and, as discussed by Ayers and Paas (2007), Mayer and Moreno (2010) and Sweller (2010), to prevent learners from engaging in extraneous processing.

In regards to the segmentation structure with the use of visual effects, the video Things They Can or Can't Do! can be argued to present a rather simplistic design in what concerns visual effects to introduce or end each of the segments. Throughout the video, each segment is introduced by a fade-in effect, and when the target language structure is presented, that is, the question and answer in each segment, there is a fading out effect, following an equal design sequence, for example, (segment 2) fades in → information presentation → fades out → (segment 3) fades in → information presentation → fades out. Such structure can be considered simple once it does not seem to overload learners' WM capacity, thus, as is explained by Sweller (2010) and Kalyuga (2010), leaving room for the active processing of the necessary pieces of information in order to make sense of the language being presented and (probably) to use it in a later moment.

As presented, the video Things They Can or Can't Do! is organized in a rather straight to the point structure, in which the use of the written language is presented by means of the subtitles and, as part of the video, as the presentation of the target structures. Furthermore, no substantial visual effects are inserted in the video design, preventing learners from extraneous processing, exceeding their WM resources to process the incoming information (SWELLER, 2010; MAYER, 2009; KALYUGA, 2010).

Coherence

As discussed, the written language presented in the video is associated to the target language questions and answers and the subtitles. It is possible to argue that the written language presented in the video can be considered essential for aiding the comprehension process. Firstly, the target language utterances in form of questions and answers presented as on-screen information can be considered essential once they present learners with the language structure concerning the uses of the modals "can" and "can't" in order to describe people's skills in performing certain tasks. Secondly, the subtitles in the design of the video can aid in the comprehension of the information being provided by the narrator, making them also essential to learners, thus not presenting any pieces of irrelevant written language.

The oral language presented in the video can also be defined as having two important aims: the first aim is to present the target language, that is, making questions and answering them by using the modals "can" and "can't". The second aim is to contextualize the skills being described in each one of the segments (as discussed in the previous section). The oral information presented in the video can, thus, also be characterized as essential in aiding learners

in apperceiving the linguistic input as discussed by Plass and Jones (2005), and possibly having a full understanding to the skills being described and somehow drawing a relationship between them and the characters. As in the example of Daniel being fond of cakes, but never seeing how his grandmother prepares them, thus, not being able to bake a cake. Perhaps the only piece of oral information that is possible to be considered not essential is when Charlotte starts producing unintelligible rhythmic sounds, resembling she is singing an opera.

Lastly, regarding the visual and aural aids in the video design, it is possible to observe that the video makes use of these two sources in its design. Concerning the images used, it is arguable that all the imagery information presented can be considered essential for the comprehension of the information presented, once, as argued by Mayer (2009), they aid the process of integration with the verbal information.

Aside from the fading in and fading out visual effects, no other visual effects are presented in the video, making it possible to defend that those effects do not seem to cause overload in learners' WM processing capacity, which, according to Ayers and Paas (2007) and Mayer (2009), would result in extraneous processing.

Regarding the soundtrack, two kinds of aural information are presented: the music as soundtrack and noises as soundtrack. The only music soundtrack that is part of the video design is presented in segment 4, depicting Charlotte being able to dance a samba. When the question is answered, she starts dancing and a soundtrack resembling the carnival melody is played, suggesting she is dancing to that sound. Noises as soundtrack can be detected in segment 8, when the narrator provides information about Charlotte singing skill, which is depicted as negative (she can't sing) with the images and sound of two glasses and a juice jar breaking as she sings. Furthermore, in segment 10, when the narrator asks about Charlotte's skill of organizing a party, and providing a positive answer, there is the sound of applause suggesting that she is very good at performing such task. The two background noises, however, do not seem cause extraneous processing by overloading learners' WM resources. Once the target language questions and answers can be considered rather concise for being presented in segments, and are constructed in a similar way, in which they are presented in each one of the segments, being only the activities different, there can be argued to be room for the processing of all the incoming visual and aural information.

4.1.1.3 Video from the Textbook Become – 7th Grade

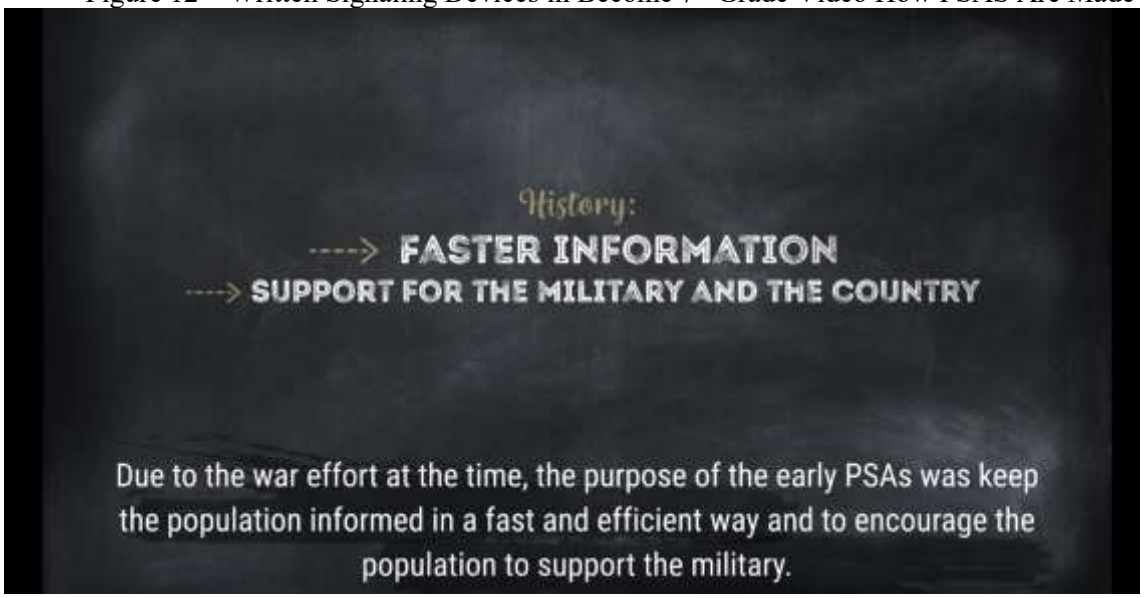
The video from the textbook Become 7th grade is entitled How Psas Are Made and has the length of six minutes and eight seconds. The video is designed in a way that involves a narrator presenting explanations of what PASs (Public Service Announcements) are, the objectives of this kind of text and what they are used for, accompanied by examples. The video also presents a number of guidelines that learners can follow in order to create their own PSAs and how to make them available at their schools as can be seen in the analysis presented.

Signaling

According to Plass and Jones (2005), the use of typographical cues can work as written verbal input. The first written language presenting information in the video is the title How PSAs are made. The title is positioned in the center of the screen under the acronym PSAs. The title can be claimed to be structured in a hierarchical structure, in which the first element presented is the acronym, about which the video presents information about, and under the acronym, the title of the video is presented. Moreover, the title of the video can also be argued to situate learners about the information that will be provided with throughout the video, that is, that they will work with information about how the PSAs are created.

In order to signal the information presented with the use of written verbal language, the video How PSAs are made makes use of subtitles in its design, which are positioned at the bottom center of the screen. The subtitles on this video, however, do not totally follow the regular structure subtitles present, usually restricted to one or two lines. In the case of the video analyzed, as can be observed in Figure 12, at some moments, they range from three to four lines. The fact that the subtitles, at some parts may become long can be considered to have a negative effect in learning. Once learners have to listen to the information being presented orally, and also, in the case of having subtitles in three or four lines, there is the possibility of learners having to search for the information by reading the lines and the on-screen information, possibly resulting in extraneous processing, thus, as discussed by Sweller (2010) and Kalyuga (2010), exceeding their working memory (WM) processing capacity.

Figure 12 – Written Signaling Devices in Become 7th Grade Video How PSAS Are Made



Source: Become Website (2019)

Based on Danan (2004), Hayati and Mohmedi (2010), and Lavaur and Bairstow (2011), the subtitles presented on the How PSAs are made video follow an intralingual design. Such design choice can be observed to aid the comprehension process of the spoken language, which can be considered of a high level of density in the video. The insertion of the subtitles can provide learners with affordances to make sense of the information being heard, which, at some times, may result, thus, in poor comprehension, by aiding them in the process of building coherent mental representations of the information, as extensively discussed by Mayer (2005a), (2005b) and (2009) once the oral and written input complement each other.

The color, the font and the size of the subtitles can have important roles in the video design. The first aspect, the color white, in which the subtitles are presented, can be examined due to the background of the screen. Throughout the video, the background image resembles a blackboard, displaying some kinds of white faded stains, suggesting it has already been used and erased several times. Since in many classrooms there still are blackboards, the main chalk color that is used in the teaching context can be white, making it possible to draw a connection between the background image of the video and choice of color for the written verbal information, that is, white written information on a blackboard. Another fact is that once the blackboard has a rather darkened color, the choice of the color white to provide the written verbal information can make the written words more salient, that is, according to Lorch (1989), Schmidt (1990), and Koning *et al.* (2010), more visible to learners due to the contrast of the two colors. In respect to the font used and size of the subtitles, it is possible to argue that the font chosen is one that does not present stylish, or embellishment features in order to aid learners

in clearly detecting the letters and words to be read and understood. Although the size can be considered small, it is possible to say that learners' effort to see the written pieces of information presented is reduced. Still considering the font of the subtitles, there are two moments in which the subtitles present three words in capital letters, which are CHANGE, ALERT and the expression PROMOTE CHANGE. The choice for presenting these words in capital letters in the subtitles can be associated to the fact that these words summarize the objectives of creating PSAs, directing learners' attention to these concepts.

Apart from the subtitles, the video also presents written verbal information as on-screen target text as part of its design. Throughout the video, as the narrator provides information and explanations about what PSAs are with examples and how to guidelines, sentences with key-concepts are presented as on-screen target text. As can be observed in Figure 12, based on Hasler, Kersten and Sweller (2007), the written verbal information presented on screen as the video advances with the information can direct learners' attention to the main facts, when presenting the explanations about the creation and use of PSAs, and the steps necessary for developing a PSA.

Moreover, it is possible to observe that the written verbal information presented on screen is designed with the use of different fonts, sizes and colors. Concerning the fonts, all key-concept sentences are presented in capital letters. Furthermore, it is also observable that within the key-concepts there is a noticeable difference in the sizes of the sentences, as, for example, can be observed in the sentences in Figure 12, in which the sentence *FASTER INFORMATION* has a bigger size than the sentence *SUPPORT FOR THE MILITARY AND THE COUNTRY*. The use of different font sizes suggests a hierarchical organization of the information, in which the first key-sentence establishes a general concept or context, and the following sentence(s) provide(s) extra information about that context, which, as defended by Mayer (2009), can aid learners in organizing the information in coherent strings of cause and effect relations. One reason for presenting the key-concept sentences in capital letters can be to make them more salient for learners. According to Lorch (1989) and Schmidt (1990) typographical resources can aid in directing learners' attention to the information being presented. In addition, once the amount of information presented in the video can be considered massive, the key-concepts can be argued to provide the main ideas in a more concise way, argued by Mayer (2005a), (2005b) and (2009) to aid learners in building coherent mental representations of the information.

The use of stylish fonts for the embellishment of the written information can be detected in certain parts of the video. The first is in the title of the video "What are PSAs?" with two

different fonts. The first part of the sentence “What are” is presented in a font that resembles the cursive writing style, and the second PSAs is presented in the same font as the other key-concepts presented throughout the video. Another moment where embellishment fonts are used can be observed when the narrator explains the history of PSAs, the words history and today are presented in different font, size and color compared to the subtitles and the key-concepts positioned on the top of the screen, above the concepts, resembling a text structure, in which they work as a title. The sentence Do It Yourself is also presented in a bigger font, in which the colors of the first letters, in the case, the D, the I and the Y are faded. There is, however, no arguable reason for presenting the sentence with such design choices. The only reason that can be observed in the choices of stylish fonts is to the embellishment of the titles presented. When it comes to aiding the learning process, such design choices cannot be defended as (theoretically) having a concrete effect on promoting, or even facilitating learning in any way.

Regarding signaling with the oral language, the information throughout the video is presented in the form of narration, in which there is a voice providing all the information concerning the topic of the video, that is, about PSAs. It is possible to observe that at some moments of the video, there is the use of intonation change in the utterances produced by the narrator. The narrator produces a slightly increased intonation in the content words (signaled here as underlines), as for example, in the utterance “They are usually brief and can be in the form of a video, a radio announcement, or even a poster or image.” (0’32). Comparable change in intonation can also be observed in the utterance “Determine a goal for your PSA. What do you want to change or call attention to? What do you want people to do?” (03’26) and in the utterance “Another thing to keep in mind is the language you will use. Use simple and clear language. Use the language of the target audience so they can understand you” (4’20).

It is possible to argue that the use of intonation change can work as an aid to direct learners’ attention to the information being presented, working as a device to aid the apperception the comprehension processes, as explained by Plass and Jones (2005). In the case of the two examples provided, the intonation stress is applied to the words that can be observed to be charged with meaning to aid learner’s general comprehension. For example, it can be more effective to call learners’ attention to the word “brief”, rather than to the word “form”, once this word has meaning on its own, whereas the word form probably needs other words so that its meaning can be effectively attributed to it.

In addition to directing learners’ attention with the use of oral language as linguistic input there are parts of the video where the utterances produced by the narrator are built in a way that suggests she is talking to learners when presenting the utterances “Do you know what

a PSA is? Did you know that they are really helpful and can even save people's lives?" (0'10), "Now that you've learned about what PSAs are and how they are used, let's look at how to create one!" (2'56), "Now that you have learned what PSAs are and how to make them, it's your turn!" (5'43) and "After you create your own PSA with different media, such as audio, video, or even drawing, you can upload it to your school blog or webpage and share it with your friends." (5'55). With the examples provided, it is possible to observe that there is an attempt to engage learners to the topic, once they are being questioned about their knowledge of what PSAs are. Furthermore, the guidelines to be considered when producing a PSA bring learners to be position of authors, that is, they are placed in the role of creators, being responsible for their own productions, encouraging their autonomy in a hands-on task.

Regarding the use of visual aids to signal the information being presented, it is possible to observe that the video How PSAs are made uses imagery aids in its design in order to depict some of the information being presented, as can be observed in Figure 13.

Figure 13 – Exemplifying Visual Signaling Devices in Become 7th Grade Video How PSAS Are Made

Source: Become Website (2019)

As can be observed in Figure 13, the video presents imagery information to complement the information being presented by the narrator. The information provided by the narration is aided in two forms: the first, as afore discussed, is presenting written information as subtitles and on-screen text. The second is presenting visual information in the form of images, as can be observed in the figure presented above. Although most of the video presents information in the written form, there is the use of images in nine moments of the video whose aim is to depict or provide examples of the information being presented.

It is important to observe, however, that there are images that are also charged with written information, as is the case of the hurricane sign in the figure presented above. Providing learners with images that not only present imagery pieces of information, but also written information may result in overload in WM resources due to what Kalyuga, Chandler and Sweller (1999) refer to as split attention effect once learners may want to read the information composing the hurricane image and need to read the subtitles in order to understand the information being presented.

An example of the use of images to represent and help contextualize the ideas and concepts being presented in the video can also be observed in Figure 14.

Figure 14 – Contextualizing Visual Signaling Devices in Become 7th Grade Video How PSAS Are Made

Source: Become Website (2019).

The image presented can work as a visual device to aid the contextualization of the topic about social issues, such as bullying, as shown in Figure 14. In this case, the image presented in the video depicts a girl sitting on the floor, holding a sign with the written utterance STOP BULLYING. Such images can work as a contextualization device, once bullying can be closely linked to the educational context, that is, the schools.

The video How Psas Are Made also presents visual effects as composing its design to work as signaling devices. Throughout the video, the visual effects design applied to both the written and imagery information follow a pattern in which they fade in and fade out rapidly to accompany the narrator when she starts and stops presenting information.

The use of arrows can be considered a visual aid used in the video design when the written information is presented as on-screen target text. When the information presented by the narrator is complemented with written verbal information on screen, arrows roll onto the screen from the left to the right in the direction of the written information being presented, in a dynamic non-static movement, moving forward and backward.

Visual effects can also be observed in the way the subtitles are presented, not following the regular fashion in which they are presented. Rather than presenting the subtitles in the fashion making them appear when the spoken language is being presented and disappear when it stops with no visual effects, in the video design, they are shown gradually as the narrator provides the information. Moreover, it is possible to notice that in the lower part of the screen, where the subtitles are presented, there is a visual effect on their background, resembling the movement of a brush painting the board in black in the felt-to-right movement turning their background black. The choice for such visual design for the background can be of two reasons: the first, in which when the subtitles start appearing on the screen, the brushing movement can work to call learners' attention to the information being presented. The second (turning the background black) can be associated to making the subtitles, which are white, more salient for learners to read because of the color contrasting effects.

Segmenting

Although the video How Psas Are Made presents the information in, based on Leahy and Sweller (2016), a transient fashion, according to sweller (2010), Ibrahim (2012), and Ibrahim *et al.* (2012), it can be argued to be designed in segments, in which the segments present contextualization information about what PSAs are used for with a brief history of them, and guidelines for the development of PSAs, as can be observed in Chart 8.

Chart 8 – Segmentation Structure in Become 7th Grade Video How PSAS Are Made

Segments	Information Presented
Segment 1	**Narrator:** PSAs and how they are made.
Segment 2	**Narrator:** Do you know what a PSA is? Did you know that they are really helpful and can even save people's lives?
Segment 3	**Narrator:** PSA is an acronym that stands for Public Service Announcement
Segment 4	**Narrator:** They are usually an institutional effort to raise awareness about a cause or event and promote positive changes in society.
Segment 5	**Narrator:** They are usually brief and can be in the form of a video, a radio announcement or even a poster or image.

Segments	Information Presented
Segment 6	**Narrator:** They are usually transmitted voluntarily in the media by radio stations, TV networks, newspapers, magazines, social networks, etc…
Segment 7	**Narrator:** Public service announcements were first made around the beginning of Second World War in both the UK and the United States.
Segment 8	**Narrator:** Due to the war effort at the time, the purpose of the early PSAs was to keep the population informed in a fast and efficient way and to encourage the population to support the military.
Segment 9	**Narrator:** Today, PSAs are used all around the world by government agencies, non-profit organizations and non-governmental organizations
Segment 10	**Narrator:** To promote CHANGE or to ALERT the population by spreading reliable and trustworthy information.
Segment 11	**Narrator:** There are many examples of how the PSAs are used around the world…
Segment 12	**Narrator:** One of the most common types of PSA is a government campaign to prepare the population for a potential natural disaster, such as a hurricane, tornado, flood, etc.
Segment 13	**Narrator:** In the case of imminent danger, the government can quickly and easily make announcements on TV and radio stations to notify the population of the situation and guide them to safe areas.
Segment 14	**Narrator:** So, for example, if flooding is about to occur, the government body responsible for monitoring the risks will issue an alert or signal to warn the population of the affected areas and advise them on whether to leave their homes and where they can find safety and shelter.
Segment 15	**Narrator:** Another common type of PSA are the campaigns that aim to raise the awareness of social issues like violence, bullying, smoking, drug abuse, drinking and driving, etc.
Segment 16	**Narrator:** In this case, the objective is to inform the population and PROMOTE CHANGE in people's behavior or have them take action.
Segment 17	**Narrator:** Governments also make use of PSAs to raise a population's awareness of their rights.
Segment 18	**Narrator:** In this case, the goal is to inform and educate the population about their rights, so they assert and exercise their citizenship.
Segment 19	**Narrator:** Now that you've learned about what PSAs are and how they are used, let's look at how to create one!
Segment 20	**Narrator:** First of all, you have to pick a single issue you want to address. There's no time to talk about several topics in a single PSA, so you will need to focus on one issue.
Segment 21	**Narrator:** You can talk about a natural disaster, social cause, civil rights, etc…

Segments	Information Presented
	Choose something you really believe in or care about.
Segment 22	**Narrator:** Determine a goal for your PSA. What do you want to change or call attention to? What do you want people to do?
Segment 23	**Narrator:** Then you have to determine your target audience. -Who are the people you want to affect with this? -Why is it important for them to listen to this message?
Segment 24	**Narrator:** Once you have your goal, define your message. What is the message you want to send? What do you want people to understand? The message in a PSA should be very clear and have a call to action.
Segment 25	**Narrator:** For example, if you want to encourage people to stop smoking, you can explain the harmful side effects.
Segment 26	**Narrator:** Don't forget: You have to have a single strong message. Don't try to deliver more than the one message to the audience, otherwise they amy* get confused.
Segment 27	**Narrator:** Another think to keep in mind is the language you will use. Use simple and clear language. Use the language of the target audience so they understand you.
Segment 28	**Narrator:** Verbs in the imperative are useful to add emphasis to your message, and for giving warnings or advice to your audience.
Segment 29	**Narrator:** Next, you have to find a way to get people's attention. Use creativity, humor, surprise, and emotion. Create a slogan or phrase that people will remember.
Segment 30	**Narrator:** Remember, if you are using a picture or design: Choose a powerful image and pair it with the text to deliver with the message. There needs to be a 'dialogue' between the image and the text so that it makes sense. You can also use bright colors and large fonts to emphasize you message.
Segment 31	**Narrator:** If you are recording an audio or video: Use strong diction and enunciation: speak clearly and make sure to use the correct words and pronunciation. Giving emphasis and stress to certain words in sentences, helps you express what is most important. Use emotions when you are reading, as they help to convince an audience. If they are appropriate to the message you're delivering, you can also use sound effects, music and intonation to draw attention to your PSA.
Segment 32	**Narrator:**

Segments	Information Presented
	Now that you have learned what PSAs are and how to make them, it's your turn! After you create your own PSA with different media, such as audio, video, or even drawing, you can upload it to your school blog or webpage and share it with your friends. Remember: with the right information you can save people's lives or even change the world!

*Misspelled word in the video subtitles.
Source: The researcher.

Although the video How PSAs are made follows a continuous, transient information flow design, it is possible to observe that the information presented is organized in segments, as presented in the chart above. It is also important to examine that the aim of the video is not to provide learners with information about grammatical structure as language input. The video How PSAs are made is concerned with providing learners with content information, in this case, more than teaching the use of language structures, its objective it to teach learners the use of a certain kind of media and how to create one.

Through its length, as the video provides contextual information about PSAs, including historical facts and guidelines for how to develop them, the segments, besides providing the oral information provided by the narrator, also provide written information in order to aid learners' comprehension. Concerning introducing the segments with written verbal language, it is possible to distinguish only one segment that presents such design, that is, the first segment.

(Segment 1)

Narrator:
PSAs and how they are made.

As can be observed, the first segment has the function of presenting learners with the title of the video, which is first presented visually, that is, in the written form, then the name of the video is produced orally by the narrator. This design choice for the title can be examined due to being first presented as written verbal language so that learners can read it, then presented orally, in order to integrate the written and the oral forms of the utterance presented.

Except for segment 1, it is possible to argue that no other segments are introduced with the written language. This argument can be supported once the written information presented throughout the video in no segment is presented before the narration starts. Adversely, the narrator first starts producing her speaking, and then after the oral language has started being produced, the written information starts to be presented as on-screen target text, making it an aid for the spoken language to be understood in case it is not enough to promote learners' comprehension, as discussed earlier. Grounded on the arguments by Plass and Jones (2005), it

is possible to observe that the written language presented in the video works as input presented in each segment, which is organized in the subtitles and on-screen target text, having the aim of aiding learners in the process of information comprehension, as is observed segment 3.

(Segment 3)

> **Narrator:**
> PSA is an acronym that stands for Public Service Announcement

When the information in segment 3 is presented by the narrator, the utterance "PSA = Public Service Announcement" is shown on the screen, providing learners with the written information to support what is being said. Moreover, even though the name of what PSAs are is presented on the subtitles, in segment 3 this is the most important piece of information for learners to understand, as it is also presented as on-screen target text. Such design, in which the written information in a segment is presented in a fashion to aid learners comprehension of the information being presented can be observed in segment 5 and segment 16.

(Segment 5)

> **Narrator:**
> They are usually brief and can be in the form of a video, a radio announcement or even a poster or image.

(Segment 16)

> **Narrator:**
> In this case, the objective is to inform the population and PROMOTE CHANGE in people's behavior or have them take action.

In segment 5 the narrator provides learners with explanations about the designs that PSAs can have, in which different media and sources can be used for their production. When the narrator starts producing the oral information, the utterances "Brief Format" and "Video, Radio, Posters, Images, Etc." are presented on screen. The choice of presenting the written target text on screen can be aimed at supporting learners' comprehension of what media they can use when designing a PSA. Once the information presented in segment 5 can be considered specific for the design of PSAs, the written information is presented in order to aid learners in making sense of the basic concepts explained in this segment by engaging in generative processing, according to Mayer (2009), and Mayer and Moreno (2010), the processing essential to learning, and not to turn their attentional resources to unnecessary pieces of information. Segment 16 follows a similar design, in which the written words that are provided in capital font in the subtitles are also provided in the form of written utterances on screen as "Promote CHANGE", "ALERT the Population" and "Reliable Information", once again being

responsible for directing learners' attention to the main objectives PSAs may have by means of aiding learners in the apperception process as discussed by Plass and Jones (2005).

In regards to introducing segments with the oral language, it is possible to detect that only one segment introduces a group of segments orally.

(Segment 19)

> **Narrator:**
> Now that you've learned about what PSAs are and how they are used, let's look at how to create one!

As can be observed in segment 19, not only are the segments that follow it introduced, but also, the segments that were presented before it ended. When the narrator produces the utterance "Now That You've Learned About What PSAs Are And How They Are Used", it becomes recognizable that a closure of the former segments is being made. In this case, the utterance suggests that the information provided up to this moment will be replaced by a new or different kind of information. Furthermore, when the narrator produces the second part of the sentence "Let's Look At How To Create One!", she is inviting learners to a next step in the video, which is to observe the information that will teach or provide guidelines with the steps to be taken to create a PSA. When the oral utterances are produced by the narrator, the written sentence "Do It Yourself" is shown on the screen. It can be observed that the choice for presenting this utterance suggests that not only will learners be provided with the guidelines for developing a PSA only for the sake of knowing them, but also, they will be expected to develop one.

Regarding the use of visual information, that is, the use of pictures to introduce or end the segments, some of the segments make use of visual information to support the oral information being presented, as is the case of segment 12 (Also observable in Figure 14).

(Segment 12)

> **Narrator:**
> One of the most common types of PSA is a government campaign to prepare the population for a potential natural disaster, such as a hurricane, tornado, flood, etc.

Although the segment makes use of visual information in order to support the information being provided by the narrator, and also to show learners examples of PSA, it is not possible to argue that the images introduce the segment, or any other segments that make use of imagery information, once the images do not start the segment. Rather, they are presented after the narration of each segment has started.

In regards to visual effects for the introduction or ending of the segments, it is possible to say that the video How PSAs are made does not present significant effects. The segments are started and ended with visual effects, which may have the aim at providing learners with time to process the information presented, and also at preventing information processing overload in their WM, according to Sweller (2017) and Kalyuga (2010), causing extraneous processing. Each segment starts by showing the empty background, and as the narrator starts providing the information, and the subtitles start appearing on screen, the written utterances or the imagery information appear on screen in a dynamic popping in movement. When the narrator finishes producing the information of the segment, the written information or the images make the opposite movement, disappearing from the screen in a popping out movement. Moreover, it can be argued that the video is designed in a way that makes use of simple effects, thus, not causing extraneous processing (AYERS; PAAS, 2007; MAYER, 2009; MAYER; MORENO, 2010; SWELLER, 2010) due to distracting effects, ultimately, not hurting learning.

Coherence

When the coherence principle comes into analysis, there are features in the How PSAs are made video design that can be observed. The use of subtitles and the written information presented on screen can be defended to be essential as aiding learners' comprehension of the information presented. In regards to presenting irrelevant written language, when contextualizing what PSAs are and providing learners with examples (see Figure 14 previously presented), the video makes use of subtitles and the use of imagery aids to depict examples of PSAs. It is important to observe, however, that the images exemplifying PSAs also present written information. Such imagery design can also be observed in a subsequent moment in the video (2'01) (segment 14), when the narrator explains the government spreading alerts, and the image of a cell phone with a speech balloon with written text is presented. According to Sweller (2010) and Kalyuga (2010), providing learners with the images that also contain text may result in overload of WM resources due to the split attention effect (KALYUGA; CHANDLER; SWELLER, 1999) that, based on Kalyuga, Chandler and Sweller (1999), can result from learners not knowing what pieces of information to turn their attention to, thus, affecting their comprehension, and ultimately hurting learning. One alternative would be to present images with a reduced amount of written language, and with a more concise amount of information, as for example, the image providing a girl holding a "Stop Bullying" sign (2'24). An alternative, perhaps, would be to present the examples of PSAs that contain a larger amount of written information after the narrator has done the explanation and stop providing oral information, leaving room for learners to focus on the imagery and written information.

As regards the oral information provided, the video cannot be defined as presenting learners with irrelevant oral information due to its aim at providing general information about PSAs, which can be defended to demand a different design concerning the amount of information provided, thus, as explained by Plass and Jones (2005), generating linguistic input that can be considered of a large amount. Regarding the visual effects that compose the design of the video, two features can be seen. The first is the effect applied to the subtitles and the second is the use of soundtrack.

In the case of the subtitles, the design chosen, in which the letters appear on screen one by one as the narration proceeds can have a negative impact on learning. For the fact that they are designed with a visual effect, learners may become distracted due to paying attention to the effects and to how the information is gradually presented, resulting in extraneous processing, as discussed by Ayers and Paas, (2007), and Mayer (2009), that is, paying attention to how the subtitles are presented rather than *to* the information itself. Furthermore, the complexity and amount of information presented in the video is also discussed by Koning *et al.* (2007) as well as Ibrahim, *et al.* (2012). The choice of subtitles design that follow a regular appear-disappear fashion can have more beneficial effects on the comprehension of the relevant information with the removal of the gradual appearing effects design.

The second feature, the soundtrack, in which the video presents a light background music as soundtrack throughout its full length, can be considered irrelevant for the comprehension of the information. One reason for the insertion of this musical feature can be associated to traditional audiovisual materials which, at most times, present the oral and written verbal information with a soundtrack in order (perhaps) to make it more enjoyable, or even as an aid to suggest a smoother flow between the segments. However, as the information presented in the video can be considered of a high intrinsic load, as explained by Mayer and Moreno (2010), and Sweller (2010), in which learners may need to have their WM resources fully available for the process of sense-making and comprehension, the use of soundtrack may compete for their WM resources, resulting in extraneous processing, thus, hurting learning.

4.1.1.4 Video from the Textbook Become – 8th Grade

The video from the textbook Become 8th grade is entitled The Life of Nelson Mandela and has the length of four minutes and fifty seconds. The video design presents a narrator providing information about the history and achievements of South African activist Nelson Mandela presented in the following analysis.

Signaling

The first typographical signaling device in the video is the title The Life of Nelson Mandela. It is positioned at the center of the screen, in order to suggest learners a sense of organization structure and also, as explained by Hasler, Kersten and Sweller (2007), to direct learners' attentional resources to it. In the case of an audiovisual material, as in this case, the title can be defended to follow a similar hierarchical structure, being presented in the center of the screen, and the imagery and verbal texts (words + images) to be presented in a next segment. Moreover, the title as a signaling device can also be defended as preparing learners for the information they will be exposed to, that is, based on arguments by Schnotz (2005) and Ayres and PAAS (2007), the content words Life and Nelson Mandela can work to start activating learners' prior knowledge about events from the past (History) associated to an important figure (Nelson Mandela).

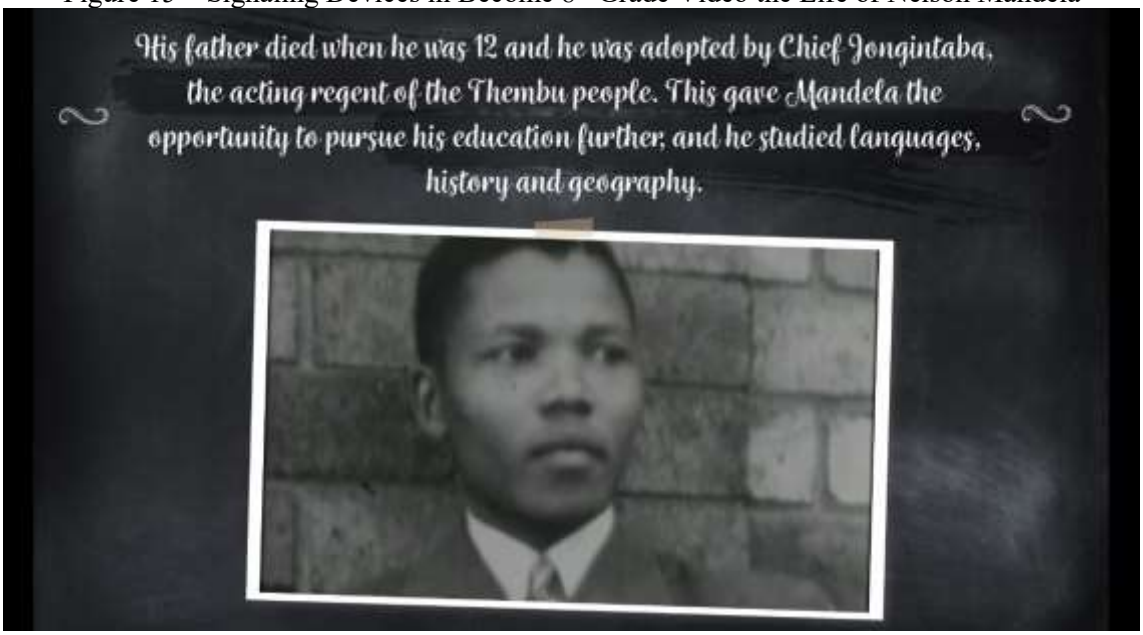
Other typographical signaling devices presented in the video are subtitles, which, according to Plass and Jones (2005), can work as written language input. Throughout the video, the information provided orally is accompanied by intralingual subtitles, that is, as explained by Danan (2004), Hayati and Mohmedi (2010) and Lavaur and Bairstow (2011), the oral language presented in the video is English, and the subtitles are designed in the same language. The design of the subtitles in the intralingual fashion can be defended to be due to the fact that, once the video is concerned with presenting learners with general information, that is, alternatively to being limited to presenting examples of grammatical structures to talk about past events, for example, the video is aimed at providing historical events about Nelson Mandela's life. In addition, the video can be argued to be charged with a high level of intrinsic load, which, according to Mayer and Moreno (2010), demands a high level of essential processing in which learners need to engage in a series of cognitive processing for successful comprehension. Considering this, the subtitles can aid this process. Due to the high element interactivity of the information presented, that is, for the comprehension of the video, learners not only need to be familiarized with the use of the past structure, but also, elements as the possessives, and events reporting rhetoric, a large amount of vocabulary items, and some degree of knowledge of historical nomenclatures, such as racial segregation, the information provided orally may not be sufficient for them to build coherent mental representations of the information being watched, defended by Mayer (2005a) and (2005b) as crucial for effective learning.

The subtitles can, then, work as aiding learners in the comprehension process when the oral information may be difficult to understand or to follow, granting them with an alternative to further understand what is being said in the video by, according to Mayer (2009), integrating

oral and written language. Although the subtitles can aid the comprehension process, they are not presented in the regular fashion subtitles are (expected to be) presented in most audiovisual materials. Firstly, they are presented on the top of the screen, apparently, to provide a sense of hierarchy of the information presented, or even, to follow the regular reading style, which goes in a movement starting from the top going downwards in order to make it more salient to learners, assumedly, being those the pieces of information they may direct their attention first for the information to be apperceived (PLASS; JONES, 2005). Secondly, although most (if not to say all) subtitles are presented in an maximum of two lines, in the video The Life of Nelson Mandela at moments the subtitles can range from three to four lines, as observable in Figure 4.7. One issue about exposing learners to four lines of subtitles at once is concerning whether they may aid or hurt comprehension. Based on Ibrahim *et al.* (2012), it can be assumed that the longer the subtitles, the greater the amount of information presented. On the counterpart, it is also important to consider that, as the video is concerned with providing general knowledge about events that took place in the past, although the information presented may be longer, based on Mayer (2009), the mental representations learners build in the process of making sense may be concise.

Furthermore, the subtitles are in a white, small size font that resembles the handwriting style, as can be observed in Figure 15. One reason for choosing the color white may be due to the background of the video. Throughout the length of the video, the same background is used in its design, which resembles a blackboard with white strains that suggest there has been a lot of writing and erasing with white chalk. The color white of the subtitles can make the written information more salient to learners due to the contrast of the white letters towards the darker background. In the case of the font type, the handwriting style can also have two reasons: one reason can be limited to having an aesthetic function for the embellishment. Another reason can be associated to the background, once suggesting a teacher is writing the information on a blackboard for learners.

Figure 15 – Signaling Devices in Become 8th Grade Video the Life of Nelson Mandela



Source: Become Website (2019)

Apart from the subtitles, the use of written language on screen to signal a piece of information to learners is observed once. In the segment in which the narrator provides learners with information about the autobiography Nelson Mandela has written, the name of the book is presented both in the subtitles and written on the screen so to direct learners' attention to its title, showing the utterance "Long Walk to Freedom" (4'08). The font is the same as used for the subtitles; however, it is presented in a bigger size compared to the subtitles. Moreover, the color used is a faded light yellow color. The choices of font size and color may be due to directing learners' attention to the name of the book, which, as can be inferred from the video, is Nelson Mandela's most remarkable production.

As the video is designed in a way that does not present an interaction situation where people are talking, but presenting a narration providing learners with information about the life and achievements of Nelson Mandela; no dialogs are part of its design, being it limited to narration only. Concerning the use of oral language that, according to Plass and Jones (2005), can work as input as to signal, that is, to direct learners' attention to important pieces of information presented, some features can be observed in the The Life of Nelson Mandela video. In the oral language design, the change of intonation is used in order to put emphasis on content words during the narration, in a way to aid learners in the apperception process (PLASS; JONES, 2005), and directly making a question in the beginning and in the ending of the video.

Throughout the video, since the oral information is provided by means of narration only, it is possible to observe that, in each segment, the narrator puts emphasis (specified as

underlined words) on content words that can be observed to be charged with meaning as the case of nouns of adjectives, for example. When the narrator produces the utterances "He spent 27 years in prison. While imprisoned, Mandela had the support of a big portion of the South African population, as well as the support of the international community, which condemned the Apartheid." (2'32) and "After leaving office in 1999, he retired from politics but continued his work as a philanthropist and advocate of peace and social justice. The Nelson Mandela foundation was established that year." (3'50), it is possible to observe that emphasis is applied to content words, or expressions that can have strong meaning to the context and the information presented at each segment of the video. Such emphasizing feature can be observed throughout the full video, whose aim is to direct learners' attention to the pieces of information presented, important for the understanding of the context of the situation reported, as well as aiding them in building coherent mental representations of the events, as argued by Mayer (2009)

The second feature used in order to direct learners' attention concerns the narrator presenting questions in two moments. The first moment is at the beginning of the video, after the title is presented, when the narrator questions learners with the utterance "What do you know about Nelson Mandela, the civil rights activist who was crucial to the dismantling of Apartheid in South Africa? Let's learn a little about his life and his struggle?" (0'13). As can be observed, based on Schnotz (2005) and Ayres and Paas (2007), the intention can be to activate or even, as argued by Ambrose *et al.* (2010), expand learners' prior knowledge about Nelson Mandela. Moreover, as the question itself provides a few important pieces of information in its structure with the words "civil rights activist" and "Apartheid in South Africa", it can be defended that not only its objective is to invite learners to reflect and to activate their prior knowledge, but also provide them with cues about the person they are being questioned. In the second question, it is observable that learners are invited to proceed to watch the video in order to learn more information about his life history and his achievements. Furthermore, it is possible to argue that the insertion of these questions in the video design has the function of directing learner's attention (SCHMIDT, 1990; LORCH, 1989; KONING *et al.*, 2010; HASLER; KERSTEN; SWELLER, 2007) to information that is to come throughout the video and, perhaps, to aid them in the apperception process (PLASS; JONES, 2005) of the information that they will watch.

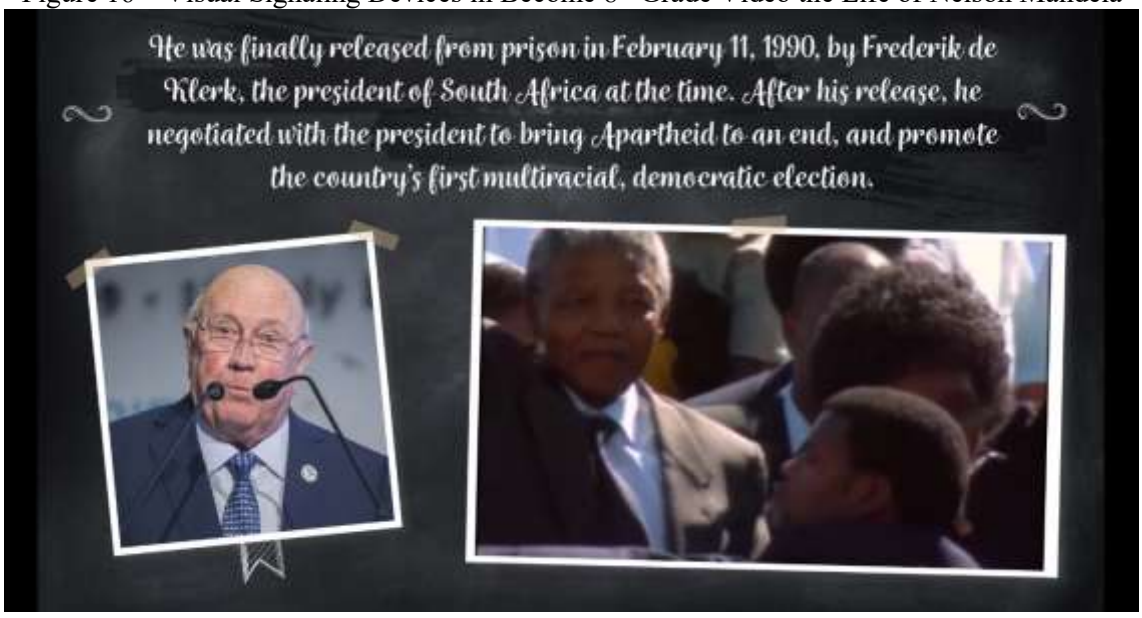
Another moment in which the narrator addresses questions directly to learners can be observed in the end of the video. After the historical events and achievements about Nelson Mandela are presented, the narrator produces the utterances "What about you? Are you inspired by Mandela's life? What can you learn about the importance of fighting for freedom, democracy

and equality from his story?" (4'45). They can have two functions. The first is to signal to learners that the video may be finishing and that there will be no further information about Nelson Mandela. The second function is to invite learners to reflect about the journey Nelson Mandela has gone through during his life concerning the difficult moments he went through and correlate these events to their own lives, associating them to fighting for people's rights.

The video The Life of Nelson Mandela further makes use of visual aids in its design in order to signal the information presented. Throughout the video, the visual aids incorporated can be observed as the presentation of images, which can be of two types: one type is the presentation of still images, that work as photos, and the other type is the presentation of fragments of motion pictures.

All the imagery pieces of information follow the same structure, in which the written information is on the top part of the screen and the images and motion pictures on the lower part of the screen, as can be observed in Figure 15. The imagery information is presented in a fashion which resembles a photograph, with the white boarders glued on the blackboard with a scotch tape and the motion pictures with frames resembling a photograph frame, as also observable in Figure 16.

Figure 16 – Visual Signaling Devices in Become 8th Grade Video the Life of Nelson Mandela



Source: Become Website (2019)

The design positioning images and pictures in a certain place on screen, that is, on the down part, suggests a hierarchical organization. Learners first need to look at the written information presented on the upper part of the screen before moving to the visual information on the lower part, suggesting that the images and motion pictures in the video work as a

complementary aid to depict the information by the narrator, having a secondary role in aiding learning. Moreover, the choice of the use of still and motion pictures can be due to granting the video a more dynamic design and, perhaps, present learners with a more realistic view of the information. Once the images and motion pictures work as an aesthetic aid, they can be defended not to have a crucial role for the comprehension of the information presented. It is, however, important to observe that presenting the images and the motion pictures at the same time as the narrator is providing the oral explanations and the subtitles are shown on screen can have negative effects on learning. According to Kalyuga, Chandler and Sweller (1999), presenting all the pieces of information with no effective time for learners to listen to and read the information and look at the imagery information, they may exceed their working memory (WM) resources as a result of the split-attention and general cognitive overload effect.

Segmenting

The video The Life of Nelson Mandela is designed in a fashion similar to the previous video analyzed, once its aim is not to provide learners with isolated grammatical structures, but to provide them with content information about the history and achievements of Nelson Mandela. Although the information is presented in a transient fashion as explained by Leahy and Sweller (2016), according to Ibrahim (2012) and Ibrahim *et al.* (2012), it is possible to argue that the video follows a segmented design, as is presented in Chart 9.

Chart 9 – Segmentation Structure in Become 8th Grade Video the Life of Nelson Mandela

Segments	Information Presented
Segment 1	**Narrator:** The Life of Nelson Mandela
Segment 2	**Narrator:** What do you know about Nelson Mandela, the civil rights activist who was crucial to the dismantling of Apartheid in South Africa? Let's learn a little about his life and his struggle?
Segment 3	**Narrator:** Mandela was born Rolihlahla Mandela in July 1918, in the village of Mvezo, South Africa.
Segment 4	**Narrator:** He was the first of his family to attend school. It was one of his teachers who named him "Nelson".
Segment 5	**Narrator:** His father died when he was 12 and he was adopted by Chief Jongintaba, the acting regent of the Thembu people. This gave Mandela the opportunity to pursue his education further, and he studied languages, history and geography.
Segment 6	**Narrator:** It was thanks to his interest in African history and to his elders, who told him that the Apartheid system made them "enslaved in their own

Segments	Information Presented
	country", that Mandela slowly became interested in contributing to the freedom of his people.
Segment 7	**Narrator:** In 1939, he enrolled at the university of Fort Hare, the only center for higher learning for black South Africans in those times of racial segregation. Before completing his degree, he was expelled for engaging in student protest.
Segment 8	**Narrator:** Following that, Mandela ran away from home to escape a marriage arranged by the Chief, and went on to live in Johannesburg, the largest city in South Africa, where he worked many jobs and went to law school.
Segment 9	**Narrator:** It was during this time that he became involved in the anti-Apartheid movement, joining the African National Congress (ANC), a black liberation group.
Segment 10	**Narrator:** The ANC fought against the Apartheid, a racial segregation system consisting of many discriminatory policies and laws against black South Africans.
Segment 11	**Narrator:** They demanded full citizenship and the same rights as white people.
Segment 12	**Narrator:** Mandela was arrested many times for his anti-Apartheid activism, and was eventually sentenced to life in prison for political offences in 1964.
Segment 13	**Narrator:** He spent 27 years in prison. While imprisoned, Mandela had the support of a big portion of the South African population, as well as the support of the international community, which condemned the Apartheid.
Segment 14	**Narrator:** With the launching of an international campaign for his release, the South African government was increasingly pressured to release Mandela.
Segment 15	**Narrator:** He was finally released from prison in February 11, 1990, by Frederik de Klerk, the president of South Africa at the time. After his release, he negotiated with the president to bring Apartheid to an end, and promote the country's first multiracial, democratic election.
Segment 16	**Narrator:** They were both awarded the Nobel Peace Prize in 1993, for working for the end of the Apartheid system.
Segment 17	**Narrator:** In 1994, Mandela was elected and became South Africa's first black president, at the age of 77. He worked to lead the country in a peaceful transition to nonracial democracy. He signed the country's new constitution, which guaranteed the rights of the minorities and freedom of expression.
Segment 18	**Narrator:** After leaving office in 1999, he retired from politics but continued his work as a philanthropist and advocate of peace and social justice. The Nelson Mandela foundation was established that year.
Segment 19	**Narrator:**

Segments	Information Presented
	He died on December 5, 2013, in his home in Johannesburg, at the age of 95. He published many books during his life, including his famous 1994 autobiography, Long Walk to Freedom, which later inspired the movie Mandela: Long Walk to Freedom, released in 2013.
Segment 20	**Narrator:** Mandela's legacy continues to inspire activists in the fight for freedom, democracy and equality. In his own words: "To be free is not merely to cast off one's chains, but to live in a way that respects and enhances the freedom of others."
Segment 21	**Narrator:** What about you? Are you inspired by Mandela's life? What can you learn about the importance of fighting for freedom, democracy and equality from his story?

Source: The researcher (2021)

Based on the chart, it is possible to observe that the video is organized in segments, each one providing content information by means of imagery and, according to Plass and Jones (2005), reporting the life history of Nelson Mandela. Although, as discussed before, the video presents the written language mostly in the form of subtitles, only one segment is introduced by presenting the written language, which is the title presented in segment 1.

(Segment 1)

> **Narrator:**
> The Life of Nelson Mandela

The written information in segment 1 is presented in both written and oral forms. It is possible that the information is presented as written language input due to its being the title of the video. The title can be argued to be presented not only orally, but also in the written form, to follow a regular audiovisual fashion in which the title of films and video are normally presented in the written form.

The title of the video can be considered the only segment in which the written language plays a primary role of presenting learners with the information to be watched through the video as on-screen target text. The next segments present the written information in the form of subtitles, having the function of integrating the written to the oral information. Segment 2 presents an example of such feature.

(Segment 2)

> **Narrator:**
> What do you know about Nelson Mandela, the civil rights activist who was crucial to the dismantling of Apartheid in South Africa? Let's learn a little about his life and his struggle?

As can be observed in segment 2, the narrator uses a language addressing learners, resembling an interaction, in which someone is asking them a question. Furthermore, in the

second questions presented, “Let’s learn a little about his life and his struggle?” the oral language is used to invite learners to continue watching the video in order to learn pieces of information about the character, and probably, based on Koning *et al.* (2010) and Hasler, Kersten and Sweller (2007), to demand from them to have their attentional resources directed to the information to be presented.

Another example in which the narrator turns her speech to address to learners is in the last segment.

(Segment 21)

> **Narrator:**
> What about you? Are you inspired by Mandela’s life? What can you learn about the importance of fighting for freedom, democracy and equality from his story?

Segment 21 presents evidence of the oral language input used in order to end, not only the segments, but also the video. The utterances suggest a movement in which, after provided with information about Nelson Mandel’s life journey, it is the learners’ turn to reflect on the information and associate it to their own realities. Another possible aim for the final utterances can be to aid learners in developing a critical view on the social issues that are part of the world, as well as take an active stand in reflecting about those.

Throughout the video, the segments present information of events that are part of Nelson Mandela’s life, raging from his birth to his death. The first piece of information that can be observed is his birthplace.

(Segment 3)

> **Narrator:**
> Mandela was born Rolihlahla Mandela in July 1918, in the village of Mvezo, South Africa.

By presenting learners with information about Nelson Mandela’s birthplace, not only is the segment providing contextualized content about the character, but also, learners are being exposed to the grammatical structure used to talk about one’s hometown. Presenting the information about the place of birth in a segment can be to provide learners with a small amount of language input, characterized in two ways. The first is limited to the information about where he was born. The second is about providing learners with the language structure used to provide information about one’s place of birth. Presenting the birthplace in a segment can grant learners with a concise example of the structure, thus, not overloading their WM processing capacity, which, according to Kalyuga (2010), Mayer (2009), and Sweller (2010) is crucial in the process of knowledge construction.

Throughout the video, the pieces of information are concerned about reporting past events, thus, presenting the necessary language structures to report past events (Simple Past) as can be exemplified.

(Segment 5)

> **Narrator:**
> His father died when he was 12 and he was adopted by Chief Jongintaba, the acting regent of the Thembu people. This gave Mandela the opportunity to pursue his education further, and he studied languages, history and geography.

(Segment 7)

> **Narrator:**
> In 1939, he enrolled at the university of Fort Hare, the only center for higher learning for black South Africans in those times of racial segregation. Before completing his degree, he was expelled for engaging in student protest.

(Segment 12)

> **Narrator:**
> Mandela was arrested many times for his anti-Apartheid activism, and was eventually sentenced to life in prison for political offences in 1964.

(Segment 15)

> **Narrator:**
> He was finally released from prison in February 11, 1990, by Frederik de Klerk, the president of South Africa at the time. After his release, he negotiated with the president to bring Apartheid to an end, and promote the country's first multiracial, democratic election.

(Segment 17)

> **Narrator:**
> In 1994, Mandela was elected and became South Africa's first black president, at the age of 77. He worked to lead the country in a peaceful transition to nonracial democracy. He signed the country's new constitution, which guaranteed the rights of the minorities and freedom of expression.

The segments presented provide examples of the language used throughout the video, which, once reporting events about the life of an important figure, present the necessary grammatical structure for such. In addition, by making use of the necessary grammatical structures to report past events, the video can be observed to have its main focus on information about the events and achievements involving Nelson Mandela. Moreover, the video presents learners with examples of the rhetoric used in the narration of facts, as presented in segment 8.

(Segment 8)

> **Narrator:**
> Following that, Mandela ran away from home to escape a marriage arranged by the Chief, and went on to live in Johannesburg, the largest city in South Africa, where he worked many jobs and went to law school.

The expression “Following that” that starts segment 8 can have two functions. One is to create a connection between segments 7 and 8, presenting a sense of progression of events and providing learners with an expression that can be used when reporting a series of events of the past. Another one is to introduce the segment, in which learners’ attention is directed to the next event that follows the event described in segment 7.

Once the video reports Nelson Mandela’s journey, throughout its length it presents information in a chronological way, being his birth reported in the beginning of the video, and his death at the end, as presented in segment 19.

(Segment 19)

> **Narrator:**
> He died on December 5, 2013, in his home in Johannesburg, at the age of 95. He published many books during his life, including his famous 1994 autobiography, Long Walk to Freedom, which later inspired the movie Mandela: Long Walk to Freedom, released in 2013.

Reporting a person’s life can be argued to always present some important pieces of information, associated to birth, achievements and death. Segment 19 exemplifies not only the language input (Simple Past) used, but also a signal that the video may be coming to its end.

By observing the information presented with all the segments, it can be argued that their objective (more specifically segments 3 through 19, as presented in Chart 9) is to provide learners with information concerning the life journey of Nelson Mandela. By presenting the information in segments, according to Plass and Jones (2005), learners are provided with the language input used to report past events in a fragmented way due to the complex intrinsic nature of the information provided in which learners need to engage in essential processing for comprehension, as extensively discussed by Mayer and Moreno (2010) and Sweller (2010). The information presented in each one of the segments may demand a high level of language proficiency on the part of learners in order to avoid overload their WM resources and, as argued by Mayer (2005a) and (2009), be able to build coherent mental representations. Due to the complex nature of the information, one factor which might hurt the process of sense-making is the pace in which the narration provides the information as well as the reduced time from one segment to the next, possibly resulting in no or little time for learners to process the information, thus, hurting learning.

As regards to the use of visual aids in the segmentation structure, it is possible to argue that the video does not present significant visual effects on its segmentation design that can have a negative effect on learning. Each of the segments start by showing the background and, as the narrator starts providing the information, the subtitles start appearing on screen, followed

by a visual aid that also appears. When the narrator finishes producing the oral utterances, the subtitles and the imagery information fade out. Such a design, in which the information fades in and fades out, can provide learners with a sense of having the information in blocks so to grant them with time to process the information in a way that makes it clear for them the moment a segment starts and ends. Based on discussions by Ayers and Paas (2007) and Mayer (2009), on the segmentation visual effects, it is possible to defend that learners' WM resources are apparently not exceeded resulting in extraneous processing due to the simple nature of the visual effects implemented in the segmentation design.

Coherence

Concerning the coherence principle, which deals with analyzing features that can be considered essential or irrelevant in the video design, some aspects can be observed. Regarding the written language in the video, the title and the subtitles can be considered essential pieces of written information. The title, as would normally be expected in an educational video, can be considered essential for they work as the first aid to locate learners within the context and to draw an initial main idea of what the video will present.

The subtitles presented throughout the video accompany the information provided by the narrator, thus, being considered essential in aiding the comprehension process. Once the video is aimed at presenting learners with content information about the life of Nelson Mandela, as the title itself suggests, the amount of information (KONING *et al.*, 2007; IBRAHIM, *et al.*, 2012) and the high level of element interactivity, in other words, according to Mayer and Moreno (2010) and Sweller (2010), the high intrinsic load the information may have can result in difficulties for learners to build coherent mental representations of the information. Due to the complexity of the information and language presented, the subtitles can be considered necessary as a comprehension aid.

The oral language presented in the video is limited to the narration of the information. Once, as afore discussed, the video is not primarily aimed at providing learners with grammatical structures *per se*, but to provide them with contextualized information reporting lifetime events of a famous figure, it is possible to conclude that the oral language that is part of the video design can be considered essential. Moreover, the oral language presented does not seem to present irrelevant pieces of information, once all the utterances presented report the past events and contextualize the information being presented. Considering that the video reports past events, especially in a contextualized way, it can be defended that it would become impossible to limit the oral language to reduced, and perhaps, unconnected sentences.

Referring to the visual and aural aids that compose the video design, as the information is provided, images and motion pictures are presented. As can be observed, the imagery information has the purpose of working as integrating the written and oral information being presented, and can be argued not to be essential for the comprehension of the information presented for only depicting the information presented orally and in the written forms.

The visual effects can be characterized concerning two features: the subtitles and the imagery information. The subtitles are presented on screen as the narration moves on, in which each letter fades in revealing each word to accompany the pace with which the narration provides the information. One reason for implementing such design in the subtitles could be that learners would read them as they appear on screen. Such visual effect might have positive and negative effects on learning: the positive effect is that for the fact that as the written information is presented accompanying the oral information, learners are aided in the process of identifying of the corresponding written/oral language. As a negative effect, learners could be distracted, not directing their attentional resources to the necessary information, resulting in extraneous processing (AYERS; PAAS, 2007; MAYER, 2009; MAYER; MORENO, 2010; SWELLER, 2010). Another visual effect that can be identified is the fade-in and fade-out effects of the imagery information, which do so with a dynamic movement coming onto the screen. Although a reason for implementing visual effects might be to provide the video with a more dynamic and appealing appearance, and the effects considered quite simple, only in part could they aid learning, as aforementioned. The case, indeed, is that the positive effect of the subtitles design may compensate for the negative one when it comes to the process of building coherent mental representations of the information presented.

The soundtrack that is part of the video design concerns music, which can be described as a light instrumental music played throughout the video. The volume of the soundtrack is rather low, so to possibly grant learners with the possibility of having a clear understanding of the oral language presented with the narration. The insertion of a music can have two reasons. The first may be to provide the video a more appealing characteristic, in which the soundtrack would work as an embellishment feature of the video in order to make it more attractive to learners. The second one is to work as a way to put all the segments in a rather continuous format similar to (most) audiovisual materials, that is, although being organized in segments, they are all observed as a whole part of the audiovisual material. Without the soundtrack, they would possibly appear more independent from each other, becoming similar to a slide presentation. Although the music can be characterized as light, it may compete for learners' WM memory resources once, as explained by Kalyuga (2010) and Mayer (2009), their auditory

WM would need to process both the oral language and the music at the same time, with the possibility of lacking the necessary WM resources to build coherent mental representations, thus hurting learning. In fact, an issue that raises concerns the fact that the second reason the music was inserted, as previously discussed, can have positive effect in aiding learners in the comprehension process, compensating for the negative effects it might impose.

4.1.1.5 Video from the Textbook Time to Share – 8th Grade

The video from the textbook Time to Share 8th grade is entitled *Usando Going to e Will para Falar de Futuro* has the length of five minutes and seven seconds, and similarly to the previous videos analyzed, involves a narrator presenting explanations about a certain topic. In the case this video, grammatical explanations about the structures used to refer to future as shown in the analysis.

Signaling

Regarding the typographical cues used in the video design, that is, the first written language presented in the video is the name of the textbook (Time to Share), the title of the video (*Usando Going to e Will para Falar de Futuro*), and the name of the material publisher's logo with its name (*Editora Saraiva*). Even though the three pieces of information are presented, the most important piece of information presented is the title. According to Schnotz (2005) and Ayres and Paas (2007) it can have the function of activating what learners already know or the function of signaling to what is to be learned. However, there is an observable hierarchical typographical design which would visually suggest otherwise, as presented in Figure 17.

Figure 17 – Typographical Hierarchical Structure of the Title in Time To Share 8th Grade Video Usando Going to e Will para Falar de Futuro

Source: Time to Share Website (2019)

Concerning the way information is presented, the name of the textbook is presented in a big font size, while the title of the video is presented in a medium font size. One reason for such design of the typographical cues can be associated to the presentation of the title observed as structured in a hierarchy, in which the degree of importance comes from the top downwards. This suggests that, once the video is part of the Time to Share textbook, the name of the book is made more salient and the name of the video in a secondary position, for its font size being smaller than the name of the textbook font size as explained concerning typographical choices by Lorch (1989) and Koning *et al.* (2010). Yet concerning saliency, the colors applied to the name of the book and the title of the video can also suggest different signaling effects: the name of the textbook is presented with the font in color orange and the title is presented in black with a background in green. Visually, it is possible to defend that the name of the textbook becomes more salient by being positioned with a white background compared to the title of video due to the green background concerning the issue of color contrast.

Needless to say, learners may probably not have difficulties in identifying which piece of information is the name of the textbook and which is the title of the book. The issue is that there can be a high probability that learners first observe the name of the textbook due to its saliency, and then turn their attention to the title of the video. In order to prevent learners from directing their attentional resources to irrelevant information for the purpose of learning, an alternative would be to present only the title of the video.

Other typographical feature observable in the video is the insertion of subtitles in its design in the form of written language, according to Plass and Jones (2005), working as language input. Although the oral language in the video is presented in Portuguese (when the narrator explains the concepts) and English (when he reads the examples), the subtitles follow an intralingual (LAVAUR; BAIRSTOW, 2011; DANAN, 2004; HAYATI; MOHMEDI, 2010) design, for they accompany the language being spoken in the video. The oral language used by the narrator in order to provide explanations is Portuguese, as can be observed in Figure 18 presented.

Figure 18 – Subtitles Design in Time to Share 8th Grade Video Usando Going to e Will para Falar de Futuro

Source: Time to Share Website (2019)

When learners are provided with examples, the target language structure is presented in the form of written input as on-screen target text and the narrator switches from Portuguese to English, in order to read the examples provided as on-screen target text. One observable characteristic in the written language design is that the target language structures are presented in English and with their translation into Portuguese under them, as can be observed in Figure 19.

Figure 19 – On-Screen Written Language in Time To Share 8th Grade Video Usando Going to e Will para Falar de Futuro

Source: Time to Share Website (2019)

The typographical structure choices seem to aim at presenting a hierarchical structure, and presenting learners with a sense of relevance based on the font sizes and colors (see Figure 18). The utterances in English are designed in a bigger size, in the color orange in comparison to the utterances presenting the translations, which, in turn, are smaller and in green. The choice of different sizes and colors suggest that the utterances in English present a higher level of importance than the translations. The colors can have similar effects concerning the saliency issue, as discussed by Lorch (1989) and Koning *et al.* (2010) once, the utterances in the color orange become more salient than the green utterances, which can have an effect on turning learners' attention primarily to the utterances in the target language, based on Plass and Jones (2005), in order to aid the apperception process and, secondarily, to their translation. Another important observable detail is the use of boldface in the parts of the utterances in which the expression going to is used. As can be observed, the use of this typographical design can work as a device to turn learners' attention to specific parts of the sentences in order to aid them in noticing (SCHMIDT, 1990) the insertion of used to in the sentences. Concerning the positioning of the written language on the screen, all the utterances are positioned in the center of the screen, filling all the screen space. Such design can aim at directing learners' attention directly to the utterances provided, not leaving space for other kinds of information or unfulfilled space, thus, avoiding the split-attention effect, extensively discussed by Kalyuga, Chandler and Sweller, (1999).

In respect to presenting written language to direct learners' attention, it is possible to observe that, in two moments of the video, written language input is presented aiming at directing learners' attention to the target language being worked. When the narrator beings his explanations about future plans, the expression going to is presented on the screen next to him, arguably to work as a device to turn learners' attention to the structure they need to use when referring to future plans, which is, using the expressing going to in a sentence, aiding their apperception process as argued by Plass and Jones (2005). In a later moment in the video, when the narrator begins his explanations about predictions, the modal will is presented as on-screen target text, also suggesting that the target word needs to be included in a sentence. Both the written pieces of information – going to and will – are followed by sentence examples in which they are used after being presented as on-screen text.

Concerning signaling with the oral language, the video *Usando Going to e Will para Falar de Futuro* has different signaling features used by the narrator that can work as aids to direct learners attention to the target information, as can be observed in the following utterances.

> "Mas como assim, professor? Eu estou confuso!" (0'36)
> "Calma, vamos por partes. Vamos ver isso passo a passo." (0'39)
> "Que verbo que é esse? Isso mesmo, o verbo *to be*." (1'25)
> "E se quisermos fazer uma pergunta usando *going to*? Isso mesmo: Seguimos a regra do verbo *to be*, invertendo a posição dele com o sujeito da frase." (2'23)

As can be observed, the narrator provides utterances that reproduce questions and expressions that would be expected from learners. The utterance "*Mas como assim, professor? Eu estou confuso!*" can be defined as assuming that the initial explanation provided by the narrator would not be enough for the comprehension of the grammatical topic future with going to, making extra explanations necessary. The next utterance also provides evidence for such a view, once it reproduces a teacher's line, having the objective of lowering learners' anxiety and conducting them to the step-by-step explanations. The two last examples provide further evidence of what would be a teacher-learner interaction, where the teacher asks questions to learners, and also talks to them directly, or responds to their inquiries. One explanation for the insertion of such oral language design would be to establish engagement with learners who are watching the video by foreseeing their possible doubts concerning the grammatical topic future with going to.

Another oral signaling device produced by the narrator can be observed as calling learners' attention to the information to come, as in these examples.

> "Vamos começar com planos para o futuro. Para situações assim, utilizamos *going to*." (0'48)

> "Para responder essa pergunta, vamos verificar alguns exemplos." (0'55)
> "Observe essas frases." (0'56)
> "Observe as mesmas frases que acabamos de analisar, agora na forma negativa:" (1'43)
> "Vamos voltar aos nossos exemplos pra ver isso na prática. De olho na tela!" (2'28)
> "Pra esses casos, usamos *will*, como pode ser visto nesses exemplos:" (2'57)
> "Observe os exemplos." (4'24)

Grounded on Schmidt (1990) and Lorch (1989), the examples provided may show that their aim is to direct learners' attentional resources to the information that is in the video. In the first utterance, the narrator is introducing the grammatical topic to be presented in the video by making it clear to learners that they will start working with the target language grammatical structure used to refer to future plans by using going to. The other utterances that can also be defined as signaling devices aims at inviting (or conducting) learners to move forward in the video in order to observe examples showing going to inserted in sentences or questions. Such design seems to be due to a resemblance to the classroom routine, in which a teacher would (be expected to) first provide explanations concerning the topic, followed by examples where it is applied. According to Plass and Jones (2005) and Schmidt (1990), in this scenario, the teacher has the role of helping learners by directing their attention to the relevant pieces of information to be apperceived.

Still concerning directing learners' attention, the narrator makes use of specific words and expressions as a signaling device, not only to call their attention to the information to come, as in the previous examples, but also to the information (explanations) being presented, as can be observed with the words in bold in the following examples.

> "**Repare** que em todas as frases, *going to* permanece inalterado, seguido por um verbo em sua forma básica." (1'18)
> "**Observe** que seu uso varia de acordo com o sujeito: 'I am', 'You are', 'He is'." (1'31)
> "**E olha só**, pra formar as frases na negativa usando *going to*, também trabalhamos com o verbo *to be*." (1'38)
> "**Basicamente**, inserimos a palavra *not* após o verbo *to be*, mas também podemos usar as forma contraídas de *are not* e *is not*, que são respectivamente *aren't* e *isn't*." (2'12)
> "**Percebeu** que, diferentemente do *going to*, aqui não há variações de acordo com o sujeito?" (3'18)
> "**É importante destacar** que tanto para as frases afirmativas quanto para as frases negativas, é possível utilizar contrações – que, inclusive, são muito comuns." (4'22)
> "**Vale destacar**, também, que o *will* possui outro uso: ele expressa decisões que tomamos no momento da fala." (4'40)

The utterances presented provide evidence of the target expressions used by the narrator in order to, according to Hasler, Kersten and Sweller (2007) direct learners' attentional

resources to the information being explained. The insertion of the expressions "*Repare que*" (Notice that), "*Observe que*" (Observe that), "*Percebeu que*" (Have you noticed that)", "*E olha só*" (And look), "*É importante destacar*" (It's important to mention) and "*Vale destacar*" (It's worth mentioning) can work as aids to direct learners' attention to the relevant target information and aid their apperception process.

Another aspect that can work as a basis for this assumption is that all the words and expressions in the examples are positioned in the beginning of the sentences, once having the function of a hook in order to catch learners' attention to the information to be presented.

Recapitulating previously presented information is another oral signaling device that is part of the video design, as shown in the following example.

> "Agora que já sabemos como expressar planos para o futuro em inglês, vamos verificar como fazer previsões." (2'52)

As can be observed in the utterance presented, the author uses the oral language in order to recapitulate the target language structure worked in the video by providing learners with the utterance "*Agora que já sabemos como expressar planos para o futuro em inglês […]*" (Now that we already know how to express future plans in English […]). It is possible to assume that the narrator presents this utterance in order to recall to learners what the video has presented so far, that is, the target language structure used to refer to future plans, as is literally mentioned in the utterance. Furthermore, this part of the utterance can have the function of closing the topic future plans and starting the next one with the second part of the utterance. The second part of the utterance "[...] vamos verificar como fazer previsões." ([…] let's check how to make predictions.) can have also a signaling function to call learners' attention to the next topic to be presented in the video. In addition, according to Ambrose *et al.* (2010) it can have the function of activating their prior knowledge (SCHNOTZ, 2005; AYRES; PAAS, 2007; KALYUGA, 2010) concerning the topic predictions or even a device to facilitate the process of new knowledge building as defended by Schnotz (2005), Ayres and Paas (2007) and Kalyuga (2010).

A final oral signaling device concerns with closing the video and encouraging learners to practice the language worked, as is presented in the following example.

> "E você, quais são os seus planos para o futuro? Converse com os seus amigos a respeito disso e ouça o que eles têm a dizer, colocando em prática o que acabamos de estudar." (5'03)

The utterance can provide learners with a closure of the video. The production of the first part "*E você, quais são os seus planos para o futuro?*" (And you, what are your plans for the future?), may imply that the video is coming to its end, and that no further information and grammatical explanations will be provided. Moreover, it calls learners' attention by questioning them about what their plans for the future are, having the function of recalling the language in the video. In the second part of the utterance, the narrator encourages learners to talk to their friends and interact in the target language about their future plans by using the grammatical structures they have watched in the video in order to promote practice. The narrator, thus, seems to follow the presentation-practice sequence.

Comparing to a regular L2 classroom, the expected order of events would be, first, to provide learners with the explanations and examples and, second, to provide them with practice opportunities. This is what the narrator does in the rest of the utterance "*Converse com os seus amigos a respeito disso e ouça o que eles têm a dizer, colocando em prática o que acabamos de estudar.*" (Talk to your friends about it and listen to what they have to say, putting what we have just studied into practice.) in which the explanation moments should be followed by practice moments.

Concerning the use of visual aids to work as signaling devices, the video *Usando Going to e Will para Falar de Futuro* can be argued to present a rather limited use of visual devices in its design. Regarding the use of images, the video does not present any imagery elements in its design to aid learners in the comprehension process, being limited to presenting only the written language as visual aids.

Regarding the implementation of visual effects in the design, it can be defended that the video presents a considerably limited amount of effects. The only visual effects that can be detected are presented when learners are provided with the language structure examples as written text on screen. When the narrator provides each sentence, it appears on screen with a reduced size and keeps becoming bigger for a short period of time until it reaches a bigger size. This effect is also implemented to the translation sentences, presented after each target structure sentence, similarly to the visual effects that can be applied to slideshow presentations, as presented in Figure 18. Such visual effect can have been used in order to grant a dynamic characteristic to the written language, and to work as an aid to direct their attentional resources (SCHMIDT, 1990; LORCH, 1989; KONING *et al.*, 2010; HASLER; KERSTEN; SWELLER, 2007) to the sentences being presented both orally and in the written forms and does not appear to compete for working memory (WM) resources due to its simplistic design, ultimately not hurting learning.

Segmenting

Although the video *Usando Going to e Will para Falar de* Futuro presents information in a transient fashion (LEAHY; SWELLER, 2016) due to being an audiovisual material, that is, a video, it can be defined as being designed following a segmentation structure, as defined by Sweller (2010) and IBRAHIM (2012) once the information is presented to learners in blocks, as can be observed in Chart 10.

Chart 10 – Segmentation Structure in Time to Share 8th Grade Video Usando Going To e Will para Falar de Futuro

Segments	Information Presented
Segment 1	TIME TO SHARE Usando going to e will para falar de futuro
Segment 2	**Narrator:** Olá! Tudo bem? Seja Bem-vindo a esta videoaula!
Segment 3	**Narrator:** Enquanto não é inventado um dispositivo que nos permita viajar no tempo e ver o futuro, seguimos apenas falando dele, por meio de planejamentos e suposições.
Segment 4	**Narrator:** Na língua portuguesa, não existem regras específicas para se fazer planos ou previsões. Mas em inglês, é diferente: as estruturas que utilizamos são muito importantes, uma vez que afetam diretamente o objetivo da nossa fala.
Segment 5	**Narrator:** “Mas como assim, professor? Eu estou confuso!” Calma... vamos por partes. Vamos ver isso passo a passo. E, acredite: é mais simples do que parece.
Segment 6	**Narrator:** Vamos começar com planos para o futuro. Para situações assim, utilizamos going to. Mas como empregar going to corretamente? Pra responder a essa pergunta, vamos verificar alguns exemplos. Observe essas frases:
Segment 7	**Narrator:** I am going to visit my friends next week. (Vou visitar meus amigos na semana que vem.) You are going to play soccer tomorrow. (Você vai jogar futebol amanhã.) He is going to travel to London. (Ele vai viajar para Londres.) Repare que em todas as frases, o going to permanece inalterado, seguido por um verbo em sua forma básica.
Segment 8	**Narrator:** Mas tem um outro verbo que muda em todos os exemplos. Que verbo que é esse? Isso mesmo, o verbo to be. Observe que seu uso varia de acordo com o sujeito: I am, You are, He is.
Segment 9	**Narrator:** E olha só: pra formar as frases na negativa usando going to, também trabalhamos com o verbo to be. Observe as mesmas frases que acabamos de analisar, agora na forma negativa:
Segment 10	**Narrator:** I am not going to visit my friends next week. (Eu não vou visitar meus amigos na semana que vem.)

Segments	Information Presented
	You are not going to play soccer tomorrow. (Você não vai jogar futebol amanhã.) He is not going to travel to London. (Ele não vai viajar para Londres.) Basicamente, inserimos a palavra not após o verbo to be, mas também podemos usar as formas contraídas de are not e is not, que são respectivamente aren't e isn't.
Segment 11	**Narrator:** E se quisermos fazer uma pergunta usando o going to? Isso mesmo: seguimos a regra do verbo to be, invertendo a posição dele com o sujeito da frase. Vamos voltar aos nossos exemplos para ver isso na prática. De olho na tela!
Segment 12	**Narrator:** Am I going to visit my friends next week? (Eu vou visitar meus amigos na semana que vem?) Are You going to play soccer tomorrow? (Você vai jogar futebol amanhã?) Is he going to travel to London? (Ele vai viajar para Londres?) Agora que já sabemos como expressão planos para o futuro em inglês, vamos verificar como fazer previsões. Pra esses casos, usamos will, como pode ser visto nesses exemplos:
Segment 13	**Narrator:** You will pass the exam. (Você vai passar no teste.) She will be here tomorrow. (Ela vai estar aqui amanhã.) They will start the presentation at 8 o' clock. (Eles vão começar a apresentação às 8 horas.) Percebeu que, diferentemente do going to, aqui há variações de acordo com o sujeito? Pois é! O importante, nesse caso, é lembrar da colocação do will, que fica entre o sujeito e o verbo.
Segment 14	**Narrator:** E para a forma negativa? Bem, nesse caso, incluímos a palavra not logo após o will, como mostram os exemplos na tela.
Segment 15	**Narrator:** You will not pass the exam. (Você não vai passar no teste.) She will not be here tomorrow. (Ela não vai estar aqui amanhã.) They will not start the presentation at 8 o' clock. (Eles não vão começar a apresentação às 8 horas.) Já para a forma interrogativa, usamos o will no início da frase, antes do sujeito, como mostram os exemplos:
Segment 16	**Narrator:** Will you pass the exam? (Você vai passar no teste?) Will she be here tomorrow? (Ela vai estar aqui amanhã?) Will they start the presentation at 8 o' clock? (Eles vão começar a apresentação às 8 horas?)
Segment 17	**Narrator:**

Segments	Information Presented
	É importante destacar que, tanto para as frases afirmativas quanto para as frases negativas, é possível utilizar contrações - que, inclusive, são muito comuns. Observe os exemplos:
Segment 18	**Narrator:** You'll pass the exam. Com a junção do sujeito you e do will. She won't be here tomorrow. Com a junção de will e not.
Segment 19	**Narrator:** Vale destacar, também, que o will possui outro uso: ele expressa decisões que tomamos no momento da fala. Assim, ao decidir apagar a luz para assistir a um filme, por exemplo, eu posso dizer:
Segment 20	**Narrator:** Let's watch a movie. I'll turn off the lights. Percebe que eu não tracei nenhum plano para apagar a luz? Foi algo que eu decidi naquele momento.
Segment 21	**Narrator:** E você? Quais são seus planos para o futuro? Converse com os seus amigos a respeito disso e ouça o que eles têm a dizer, colocando em prática o que acabamos de estudar. Até a próxima!

Source: The researcher (2021)

In the segmentation structure of the video, the use of written language can be observed in the first segment presented.

(Segment 1)

TIME TO SHARE
Usando *going to* e *will* para falar de future.

According to Plass and Jones (2005), the presentation of language input in the written form in segment 1 can be observed to have the function of introducing the video by presenting the name of the textbook and the title of the video, following a structure that would be regularly expected in an audiovisual production. In the case of the *Usando Going to e Will para Falar de Futuro* video, it is possible to observe that the aim of the segment, beyond introducing the segments which will be subsequently presented by opening the video, is to situate learners within the grammatical topics that will be presented. Segment introduction with written language can also be observed in segment 6. Only the utterances Going to and Will can be considered input, once they are the only verbal written information presented in the L2.When the narrator is explaining that learners will first be presented with explanations concerning the use of going to to express plans, going to is presented on the television screen positioned next to him, introducing the next segments, which are associated to providing learners with explanations about its use. A crucial issue is that although the video presents explanations in a segmented way, only the utterances that are presented in the L2 can be considered input.

Presenting segments whose aim is at proving learners with examples of oral and written language input of the target language structure can be observed in a series of segments. One example is segment 7, as presented.

(Segment 7)

Narrator:
I am going to visit my friends next week.
(Vou visitar meus amigos na semana que vem.)
You are going to play soccer tomorrow.
(Você vai jogar futebol amanhã.)
He is going to travel to London.
(Ele vai viajar para Londres.)
[...]

As can be observed, segment 7 is designed in order to provide learners with examples of the target language in the form of written input. In this case, the aim is at exemplifying the structure used to refer to future plans with the insertion of going to. Based on the design of segment 7, it is possible to say that, in order to aid learners in the comprehension of the meaning of the sentences, they are provided in the target language, that is, in English, and with the translation, in Portuguese. The same pattern can be observed in segments 10, 12, 13, 15, 16 and 18, as presented.

(Segment 10)

Narrator:
I am not going to visit my friends next week.
(Eu não vou visitar meus amigos na semana que vem.)
You are not going to play soccer tomorrow.
(Você não vai jogar futebol amanhã.)
He is not going to travel to London.
(Ele não vai viajar para Londres.)
[...]

(Segment 12)

Narrator:
Am I going to visit my friends next week?
(Eu vou visitar meus amigos na semana que vem?)
Are You going to play soccer tomorrow?
(Você vai jogar futebol amanhã?)
Is he going to travel to London?
(Ele vai viajar para Londres?)
[...]

(Segment 13)

Narrator:
You will pass the exam.
(Você vai passar no teste.)
She will be here tomorrow.
(Ela vai estar aqui amanhã.)
They will start the presentation at 8 o' clock.
(Eles vão começar a apresentação às 8 horas.)
[...]

(Segment 15)

Narrator:
[…]
Já para a forma interrogativa, usamos o will no início da frase, antes do sujeito, como mostram os exemplos:

(Segment 16)

Narrator:
Will you pass the exam?
(Você vai passar no teste?)
Will she be here tomorrow?
(Ela vai estar aqui amanhã?)
Will they start the presentation at 8 o' clock?
(Eles vão começar a apresentação às 8 horas?)

(Segment 18)

Narrator:
You'll pass the exam.
Com a junção do sujeito you e do will.
She won't be here tomorrow.
Com a junção de will e not.

All the segments presented above follow the same design pattern, in which the target language structure is presented in order to provide learners with a visual type of linguistic input, in which they can see the language, followed by the translation (both in written form). Furthermore, it is possible to conclude that the choice of providing the translated written form of the sentences aims at guaranteeing learners' understanding of the sentences and, as explained by Plass and Jones (2005), (probably) apperceiving the structure organization to prevent them from extraneous processing (AYERS; PAAS, 2007; MAYER, 2009). Moreover, segments 10 and 12 present examples of the target language structure used for negatives and questions with the use of going to.

Segments 13, 15, 16 and 18 also follow the design pattern of presenting the target language structure followed by the translation as written input. However, they use the modal "will" in order to make predictions, as can be observed in the examples presented. Segment 13 provides learners with examples of the affirmative structure using "will", segment 15 presents examples of the negative structure and segment 16 presents examples of how questions are structured with the use of the modal. Lastly, segment 18, presents the sentences previously provided with the contracted form of "will". Again, the aim of the segments that present the target language structure with the use of the modal "will" is to provide learners with examples of how affirmatives, negatives and questions are structured, followed by the translation, probably, in order not to leave room for miscomprehension of both the language structure and the meanings, of the examples provided, overloading learners' WM resources (KALYUGA, 2010; MAYER, 2009; SWELLER, 2017).

Regarding the use of the oral language in order to introduce the segments, it is possible to identify a number of segments with this feature in their design. The first feature that is possible to be observed in the oral language for instructing learners and introducing the next segments can be observed in the following examples.

(Segment 6)

> **Narrator:**
> Vamos começar com planos para o futuro. Para situações assim, utilizamos going to. Mas como empregar going to corretamente? Pra responder a essa pergunta, vamos verificar alguns exemplos. Observe essas frases:

(Segment 9)

> **Narrator:**
> E olha só: pra formar as frases na negativa usando going to, também trabalhamos com o verbo to be. Observe as mesmas frases que acabamos de analisar, agora na forma negativa:

Segments 6 and 9, although not providing learners with the target language input in any way, can be defended as having the objective of directing learners' attention and, according to Plass and Jones (2005), aiding the apperception process concerning the information that will be provided by the narrator about the use of going to to make plans (segment 6) and the structuring of the of sentence by providing explanations about verb to be. Further explanations provided concerning the structuring of the sentences with the use of the verb to be are also presented in segments 8 and 10, as can be observed.

(Segment 8)

> **Narrator:**
> Mas tem um outro verbo que muda em todos os exemplos. Que verbo que é esse? Isso mesmo, o verbo to be. Observe que seu uso varia de acordo com o sujeito: I am, You are, He is.

(Segment 10)

> **Narrator:**
> [...]
> Basicamente, inserimos a palavra not após o verbo to be, mas também podemos usar as formas contraídas de are not e is not, que são respectivamente aren't e isn't.

It is possible to defend that the aim of the segments presented is to direct learners' attention to the explanations provided by the narrator, that is, that the verb to be has to agree with the subjects of the sentences (segment 8). In addition, for the structuring of negative sentences with the verb, the insertion of the word not is inserted after the verb. A comparable structure can be identified in segments 12, 14, 15, 17 and 19, as presented.

(Segment 12)

> **Narrator:**
> […]

Agora que já sabemos como expressão planos para o futuro em inglês, vamos verificar como fazer previsões. Pra esses casos, usamos will, como pode ser visto nesses exemplos:

(Segment 14)

Narrator:
E para a forma negativa? Bem, nesse caso, incluímos a palavra not logo após o will, como mostram os exemplos na tela.

(Segment 15)

Narrator:
[...]
Já para a forma interrogativa, usamos o will no início da frase, antes do sujeito, como mostram os exemplos:

(Segment 17)

Narrator:
É importante destacar que, tanto para as frases afirmativas quanto para as frases negativas, é possível utilizar contrações - que, inclusive, são muito comuns. Observe os exemplos:

(Segment 19)

Narrator:
Vale destacar, também, que o will possui outro uso: ele expressa decisões que tomamos no momento da fala. Assim, ao decidir apagar a luz para assistir a um filme, por exemplo, eu posso dizer:

The segments presented can be observed to have two objectives: the first is to direct learners' attention to certain explanations, or instructions provided by the narrator to be apperceived and the second objective is to introduce the following segments, which present examples of the target language structure being explained. Such segmentation design can be observed as providing learners with time from one segment to the next one in order to process the information that was presented by the narrator before observing it in a sentence. Directing learners' attention can also be observed in segments 7 and 13, when the narrator calls their attention to the use of going to, as can be observed.

(Segment 7)

Narrator:
[…]
Repare que em todas as frases, o going to permanece inalterado, seguido por um verbo em sua forma básica.

(Segment 13)

Narrator:
[…]
Percebeu que, diferentemente do going to, aqui há variações de acordo com o sujeito? Pois é! O importante, nesse caso, é lembrar da colocação do will, que fica entre o sujeito e o verbo.

The narrator's speech is produced in a way that resembles a face-to-face interaction between teacher and learners, possibly to establish some degree of engagement with learners (perhaps) to make them feel more comfortable with the explanations and to lower their (possible) anxiety about the topic. The utterances produced by the narrator aim at directing learners' attention to the fact that when using going to no changes are made in all the example sentences, being followed by a verb in the base form (segment 7), and, when asking whether they have noticed how different will is from going to in the sentences regarding the agreement with the sentence subject. As can be observed, the segments present a design structure in which the narrator provides learners with grammatical explanations by directing their attention to the information being explained, followed by inviting them to move to the next segment to observe the target language structure being shown in the written form.

Concerning the use of oral language in the segmentation design to end segments, it is possible to detect such feature in segment 21, as presented.

(Segment 21)

> **Narrator:**
> E você? Quais são seus planos para o futuro? Converse com os seus amigos a respeito disso e ouça o que eles têm a dizer, colocando em prática o que acabamos de estudar. Até a próxima!

Segment 21, as presented, can be argued to have the function of ending all the segments presented in the video, as well as the video itself. When the narrator presents the question about what learners' plans for the future are, encouraging them to talk to their peers about plans using the language structure they "have studied", it can be suggested that no further information and explanations will be provided. Finally, it becomes clear to learners that the video is ending when the narrator says farewell, also ending the video itself.

Concerning the segmentation design of the video with the use of visual aids, the video *Usando Going to e Will para Falar de Futuro* does not present any kind of imagery aids, as previously discussed, thus not having this kind of segmentation design. The video, however, presents editing effects as part of the segmentation design aiming at introducing the segments. Throughout the video, it is possible to detect that editing effects are inserted before the segments that present written verbal information, such as segment 1, in which the name of the textbook and the title of the video are presented as well as the segments that present the examples of the target language structures.

The editing effect can be defined as resembling a slideshow transition effect, presenting a dynamic movement of lines in horizontal direction from left to right, and right to left. Using the colors white, black, orange and green, which can be observed as used in the segments that

present examples of the target language structure being, suggesting a color pattern used throughout the video, especially, when presenting the written verbal language. The insertion of the editing effect, however, can only be detected before the segments that present written information, since, after the segments end, there is the clear-cut editing which cuts directly back to the narrator, possibly to prevent learners from distracting with visual or embellishment effects, resulting in extraneous processing, as discussed by Mayer (2009) and Sweller (2010)

Coherence

Regarding coherence as the principle that concerns with the information presented in design that can be considered essential or irrelevant to aid learning, some features in the video *Usando Going to e Will para Falar de Futuro* can be detected. Concerning the presentation of irrelevant written language, it is observable that when the title of the video is presented, the name of the textbook is also presented unnecessarily. It is possible to argue that the presentation of the name of the textbook can be considered irrelevant, once it can be assumed that, for learners to be provided with this video, they would be using the textbook in their classes, thus, knowing its name, making it unnecessary to have the name of the textbook again presented in the video.

Another irrelevant piece of written information that may cause extraneous processing is concerning the translated form of the written examples. The presentation of the translated form of the target language structures can be defended to hurt learning once the narrator presents the oral information in learners' L1, and after reading the examples in the L2, he provides the oral translation of the utterances. Arguably, there is no need to present the translation in the written form, once it is already presented in the oral form, and in the subtitles. A possible result would be extraneous processing, once learners are provided with a massive amount of information, according to Koning, Tabbers, Rikers and Paas (2007), resulting in difficulties to direct their attentional resources to, thus hurting the apperception process.

Furthermore, the subtitles that are part of the video design can be considered irrelevant once the language used by the narrator in the video in order to provide learners with explanations concerning the use of going to and will in order to refer to future are presented in learners' L1, that is, in Portuguese. Moreover, when presenting the examples of the target language structure, not only are they presented in the written L2, that is, in English, but they are also provided in the L1, as translated form, as can be observed in Figure 18 previously presented. The subtitles presented in the video design can, thus, be argued to be irrelevant once, instead as working as an aid for learners to understand the information provided, they can work as a distractor provoking extraneous processing by the split-attention effect as discussed by

Kalyuga, Chandler and Sweller (1999), exceeding their WM resources due to the fact that the written information is presented in a redundant way.

The oral language in L2 presented in the video can be considered crucial for the comprehension of the information presented, once it works as an aid to contextualize the target language, that is, the use of going to and will to make plans and predictions. It is possible to defend that the oral language presented in the video is crucial to aid learners' comprehension of the explanations, thus, facilitating learning.

Regarding the use of visual and aural effect in the video design, aural effects such as noises or soundtrack are not detectable, making it possible to classify the video as a straight-to-the-point audiovisual material, in which the explanations are presented in a rather concise manner, with no significant visual effects. As regards to the editing effects that are inserted as introducing the segments that present the examples of the target language structure, it is possible to defend that the insertion of the effects does not contribute to the comprehension of the information presented, thus hurting learning. In addition, the editing effects could also work as a distractor to learners causing them to lose track of attention to the information, thus provoking extraneous processing. In a further view grounded on Ayers and Paas (2007), Mayer (2009) and Sweller (2010), such effects could impact learning negatively, once, ideally, the information that does not contribute to learning, or that is irrelevant, should be eliminated from the video, in order to prevent learners from engaging in extraneous processing, thus, hurting learning.

4.1.2 Videos from the Interchange Textbooks

As aforementioned, this subsection presents the analysis of the videos that accompany the Interchange textbooks. The videos analyzed are organized according to the levels of the textbooks, in which the first two videos are from the Interchange Intro textbook, one from unit 1 (subsubsection 4.1.2.1) and one from unit 16 (subsubsection 4.1.2.2). The two subsequent videos accompany unit 8 from the Interchange 1 and 2 textbooks (subsubsections 4.1.2.3 and 4.1.2.4, respectively). Finally, the last video analyzed is from the Interchange 3 textbook, unit 1 and 16 (subsubsection 4.1.2.5). These analyses are followed by section 4.2 Discussion, revisiting the research questions that guide this Doctoral Dissertation, by answering them based on the analyses.

4.1.2.1 Video from the Textbook Interchange Intro – Unit 1

The video from Interchange Intro Unit 1 has the length of two minutes and twenty-two seconds. The video is contextualized in a school situation where the characters are asking and answering personal information, such as the names, the classes and the rooms they are in and the names of the teachers. The design of the video follows a dramatized story where two students who are new at school need to find their classroom as is presented in the analysis.

Signaling

Concerning the presentation of written language in the Interchange Intro Unit 1 video design, it is possible to observe that the video does not present a title, as would (normally) be expected in a video. The only piece of information that might work as a title is the name of the file in the Interchange 5th Edition website, which, after downloaded, is displayed as Interchange_video_Intro_Unit_1. Moreover, the video does not present subtitles in its design as typographical cues that might aid learners' understanding of the information. Concerning typographical cues that may work as signaling devices to aid learning, there are two types of cues that are presented. The first one is a sign behind the characters with the utterance "welcome students" written in capital letters, as can be observed in Figure 20.

Figure 20 – Contextualization Typographical Signaling Device in Interchange Intro Unit 1 Video

Source: Interchange 5th Edition website (2021)

According to Innaci and Sam (2017) and Tuyen and Huyen (2019), the sign may work as a contextualization cue, in which students are being welcomed to school, suggesting they are returning from vacation, or from a break, and that the year or semester is beginning. The

utterance “Welcome Students” may be considered an essential comprehension device for two reasons. The first reason is that the sign is presented with eye-catching size, word font and colors, which can have the effect of attracting learners’ attention to the context of the interaction, in which the students are new to that place. The second reason is that although the sign is presented in a back plane, learners’ attention would be directed to the two characters and their oral productions and to it in order to draw the bigger picture of the situation.

Another typographical clue that is given focus to in the video, thus working as a signaling device, is the presentation of the numbers of the classrooms, suggesting that the two students are moving around looking for the correct classroom, as presented in Figure 21.

Figure 21 – Typographical Signaling Device Show the Classrooms Numbers in Interchange Intro Unit 1 Video

Source: Interchange 5th Edition website (2021)

Regarding the use of oral language, the video does not present narration due to its design that depicts a conversational situation where students and teachers interact, introducing themselves and trying to find their classroom. It is possible, however, to observe that, once there is no use of subtitles to aid learners in the process of understanding the oral language, there is a concern with the intonation used in the production of the utterances. When the students Peter and Molly provide their names to the teachers, they change not only the intonation, but also the pace in order to emphasize their names, possibly, according to Plass and Jones (2005), to aid learners in the process of apperceiving the information presented. When Peter produced the utterance “My class is at nine o’ clock with… Ms. … Taylor”. As can be observed, there is a pause (indicated by …) that may direct learners’ attention to name of the teacher. In a later

moment, they will ask the names of the teachers in order to find out who their teacher is. Based on Kalyuga (2010), Mayer (2009) and Sweller (2017), in addition to directing learners' attention, the pauses may provide them with time to process the information presented.

The repetition of utterances is another oral signaling device that can help learners in, according to Plass and Jones (2005), the apperception process. According to Schmidt (1990), Lorch (1989) and Koning, Tabbers, Rikers and Paas (2010), as learners' attentional resources are directed to the target information, the understanding of it can be fostered. This can be observed when Molly reacts to Peter's utterance by saying "Ms. Taylor? She's my teacher.", by changing its intonation in order to emphasize the word *my* (indicated by the underscore). The emphasis can be observed to aid learners infer that the two characters are in the same class once they do not mention it explicitly, providing a reason for both to search for their class together. Information repetition is also observable when the two students are talking to the teachers. When introducing themselves, Peter greets the teachers providing his first name, followed by his full name in the utterance "I'm Peter. Peter Chrome".

Not only is the repetition of the names part of the oral interactions that the video presents, but also the spelling of the names, as can be observed when Molly corrects the first teacher who misunderstood her name for Holly, producing the utterance "No, its Molly, M-O-L-L-Y.". Emphasis to the first letter is given to indicate that was the only mistake. As explained by Koning, Tabbers, Rikers and Paas (2010), Hasler, Kersten and Sweller (2007), this can work as a signaling device to direct learners' attention to the information being provided. Such emphasis can also be observed when the first teacher the two students talk to responds Peter's question about her identity answering "Oh, no, I'm not. […] I'm Mrs. Smith. Ms. Taylor is in room two-oh-three over there", directing learners' attention to the numbers of the room in order for the two students to understand they are in the wrong place.

Concerning visual signaling devices, the clock can be considered a device, once when the bell rings, the two students look up (suggesting they are looking at the clock), and it is presented on screen, showing its two-to-nine, suggesting the classes are about to begin, making them go search for their classroom.

As can be observed in the video, the signaling devices present in its design are conceived mostly in the oral language, which can be associated to the way the utterances are produced with the use of repetitions, changes in intonation and pauses within utterances. Because of the reduced signaling aids presented in the form of typographical aids, these signaling aids cannot be considered essential for the comprehension of the characters' oral productions. Because the

video is designed with a limited amount of typographical and visual cues, the oral language becomes the major means for directing learner's attention to the target language.

Segmenting

Based on Sweller (2010) and Ibrahim (2012), the Interchange Intro Unit 1 video can be argued to present a segmentation design. Although, as explained by Leahy and Sweller (2016), the video follows a transient fashion, it is possible to observe that there are different moments, or blocks where the target language is used. In the video, the target language is organized in three segments, structured by topic change, or a change of location (from the stairway to the hall) or concerning who the participants of the interactions are. In addition, each of the segments focus on the target language of making introductions to different people, as is presented in Chart 11.

Chart 11 – Segmentation Structure in Interchange Intro Unit 1video

Segments	Information Presented
Segment 1	**Molly:** Excuse me. Hum, hello. **Peter:** Hi. **Molly:** My name is Molly. What's your name? **Peter:** I'm Peter. My friends call me Pete. **Molly:** My friends call me Molly. (**Laughs**) Hi, Pete, it's nice to meet you. **Peter:** It's nice to meet you too. **Molly:** Are you a student here? **Peter:** Yes, I am. My class is at nine o' clock with Ms. Taylor. **Molly:** Ms. Taylor? She's my teacher. You're in my class. **Peter:** Great. **Molly:** Where is our class? **Peter:** It's over there.
Segment 2	**Peter:** Hi. I'm Peter. Peter Chrome. I'm in your class. **Teacher 1:** Hi, Peter. Nice to meet you. Hello, what's your name? **Molly:** I'm Molly. I'm in your class too. **Teacher 1:** Hello, Holly. Ahm, is that H-O-L-L-Y? **Molly:**

Segments	Information Presented
	No, it's Molly. M-O-L-L-Y. **Teacher 1:** M-O-L-L-Y. Molly. What's your last name, Molly? **Molly:** Lin. L-I-N. **Teacher 1:** Peter Chrome and Molly Lin. **Peter:** Ah, wait. Excuse me. Are you Ms. Taylor? **Teacher 1:** (**Laughing**) Oh, no I'm not. I'm Mrs. Smith. Ms. Taylor is in room 203, over there. **Peter:** Oh, thanks. **Teacher 1:** Have a good day. **Peter:** Good bye. **Teacher 1:** Bye.
Segment 3	**Molly:** Hello, Ms. Taylor? **Teacher 2:** Yes. **Molly:** I'm Molly Lin. **Teacher 2:** Hi, Molly, nice to meet you. There you are. Welcome Molly. **Molly:** Thank you. **Peter:** Hi, I'm Peter. Peter Chrome. **Teacher 2:** Chrome…ah, here you are. Welcome Peter. Ok, let's begin class.

Source: The researcher (2021)

As can be observed in the chart provided, the video presents a contextualized situation, in which two students are newcomers to the school and they need to find their classroom and teacher. The target language presented in the video is associated to asking for and providing personal information, such as names, nicknames, and the names of the teachers, as can be observed.

(Segment 1)

Molly:
Excuse me. Hum, hello.
Peter:
Hi.
Molly:
My name is Molly. What's your name?
Peter:
I'm Peter. My friends call me Pete.

Molly:
My friends call me Molly. (**Laughs**) Hi, Pete, it's nice to meet you.
Peter:
It's nice to meet you too.
[…]

It is possible to argue that the main objective of segment 1 is to provide learners with examples of the target language used in introducing oneself to a stranger and exchanging personal information, as can be observed in the utterances "What's your name?", "I'm Peter. My friends call me Pete." and "My friends call me Molly.". Although the two students talk about other pieces of information, such as the time and the classes they are in, it is possible that one main objective of the video is to provide learners with the target language about asking and answering names, checking for information and spelling names.

In addition, providing information about a third person is in segment 1 with the utterance "Ms. Taylor? She's my teacher.", as can be further observed.

(Segment 1)

[…]
Peter:
Yes, I am. My class is at nine o' clock with Ms. Taylor.
Molly:
Ms. Taylor? She's my teacher. You're in my class.
[…]

The utterances above can provide learners with the target language used to attributing information to a third person she, in this case, the use of the corresponding verb to be and the person's job. Once the video concerns the target language developed in unit 1, based on Plass and Jones (2005), the video works as providing oral language input exemplifying the structure worked with, that is, asking and answering questions about personal information as well as providing information about a third person.

The checking for information and spelling names is another target language developed in unit 1, thus, presented in the video, as observed in segments 2 and 3.

(Segment 2)

[…]
Teacher 1:
Hello, Holly. Ahm, is that H-O-L-L-Y?
Molly:
No, it's Molly. M-O-L-L-Y.
Teacher 1:
M-O-L-L-Y. Molly. What's your last name, Molly?
Molly:
Lin. L-I-N.
Teacher 1:
[…]

(Segment 3)

Molly:
Hello, Ms. Taylor?
Teacher 2
Yes.:
Molly:
I'm Molly Lin.
[…]
Teacher 2:
Chrome…ah, here you are. Welcome Peter. Ok, let's begin class.

Checking the spelling of people's names as well as checking their identities can be part of the language associated to the greetings/introductions context in which people exchange personal information. Such language structure can be observed, for example, in segment 2, in which spelling is used for checking information when the teacher wants to make sure the name of the student is on her list. It can be also observed in the utterance "Ahm is that H-O-L-L-Y?" to check whether it is the right spelling of the name, followed by the utterance "No, it's Molly. M-O-L-L-Y." providing the correct information. Checking for information is further observed in segment 3, when the utterance "Hello, Ms. Taylor?" is produced, and the use of repetition in the utterance "Chrome…ah, here you are. Welcome Peter.". As can be observed, the name of the student is spelled more than once, probably, to aid learners in directing their attention (KONING *et al.*, 2010; HASLER; KERSTEN; SWELLER, 2017) to the use of the alphabet letters for spelling one's name, as this is also content presented in the textbook.

Finally, another possible function of the segments is to provide learners with examples of the target language used in saying goodbye and in welcoming people.

(Segment 2)

Teacher 1:
Have a good day.
Peter:
Good bye.
Teacher 1:
Bye.

(Segment 3)

Teacher 2:
Hi, Molly, nice to meet you. There you are. Welcome Molly.
Molly:
Thank you.
[…]
Teacher 2:
[…] Welcome Peter. Ok, let's begin class.

The examples provided show evidence that the main function of the two segments is to provide learners with examples of the target language people may use when saying goodbye with the utterances "Have a good day.", Good-bye" and "Bye". For the fact that the video is

aimed at learners with low or no language proficiency, and who are working with the language used to say goodbye, segment 2 provides three examples to aid their learning of these utterances. Furthermore, segment 3 aims at providing examples of the language used in greeting someone with the utterances “nice to meet you” (Also presented in segment 1) and to welcome a stranger to a group of people, or, in the case of the context presented in the video, to a new class with the utterance “welcome […]”.

The design of the video in segments can be argued to provide learners with the target language in different situations. In a first moment, the two characters exchange personal information, and in a second moment, they need to find their classroom and teacher, demanding them to talk to the teachers until they find the right class. Moreover, the design in segments provides learners with a sense of organization of the interactions as well as providing them with time to process the target language being presented, defended by Kalyuga (2010), Mayer (2009) and Sweller (2010) as being of high importance, and also to engage in the generative processing, that is, the processing that is essential for knowledge building as explained by Kalyuga (2010) and Mayer (2009). This processing time can also be observed when the characters make pauses and repeat the information for checking.

Coherence

Concerning the information presented in the video that, according to Ibrahim (2012), can be irrelevant for learner’ comprehension, resulting in extraneous processing and hurting learning, the video presents some that does not seem to contribute to learning.

As discussed, the written language presented in the video is rather limited (the welcome sign and the numbers of the classrooms). Although the numbers of the classrooms can be important once they suggest that the two students are walking through the hallway looking for their room, the written language presented on the sign with the words “WELCOME STUDENTS” does not appear to have an essential importance for the comprehension of the specific context and the language used. A possible reason for the sign to be considered irrelevant for learners’ comprehension is the fact that no focus is given to it, to direct learners’ attention to it, as discussed, making the written information unnecessary. Moreover, for the fact that the sign is presented in the background, one possible result might be the split-attention effect, discussed by Kalyuga, Chandler and Sweller (1999), in which learner’s attention might be divided between the written information positioned behind the two characters and the information presented in their conversation.

Regarding the oral information, as the video is aimed at beginner learners, the target language, and ultimately, the utterances presented may be considered relevant for the

comprehension of the pieces of information exchanged by the characters. That can be supported by the fact that the video provides contextualized conversational situations that resemble real-life conversations, in which new students need to situate in a new school and get to know the people, that is, their teacher and find their classroom.

Concerning the aural effects, the sounds that are part of the video are of footsteps and of people talking in the background to help contextualize the location where the conversations take place, that is, a school. Such noises, however, can distract learners while directing their attentional resources to the conversations. Yet concerning aural information, there is insertion of a music soundtrack that plays when the two students walk down the hallway in order to look for their classroom, which can also be considered irrelevant in aiding learners in the process of comprehension. The soundtrack does not seem to aid learners' comprehension of the situation, that is, that the two characters are looking for their classroom. One possible reason for its insertion is to fill the lack of conversation, once it plays while the two students are walking down the hallway, but without talking. When they stop at a classroom door to talk to the teacher, the soundtrack fades out. It is possible to examine, however, that the soundtrack might not cause extraneous processing specifically due to the fact that it is played while no oral language is produced, thus, according to Kalyuga (2010), Mayer (2009) and Sweller (2010), without competing with learners' working memory resources.

Concerning visual effects, there are no visual effects or information that may disturb learning.

As can be observed, the Interchange Intro Unit 1 video can be argued to present a reduced amount of information in its design that does not aid learning. They are associated to the presentation of written language, as is the case of the sign welcoming students that works as an aid in the contextualization of the scene, and the music soundtrack that may have been inserted to compensate the lack of conversation, or to provide the video with a cinematic fashion.

4.1.2.2 Video from the Textbook Interchange Intro – Unit 16

The video from Interchange Intro Unit 16 has the length of four minutes, and can be characterized as a conversational video, presenting a dramatized conversation between two friends who meet at school and talk about the previous weekend. Once the video is associated to the topic worked in unit 16 in the textbook, the general target language presented in the video

is about making invitations and providing excuses. The analysis of the video design based on the SSC framework is presented.

Signaling

Concerning the use of typographical cues working as a signaling device to aid learners' comprehension, the absence of written language can be observed. One characteristic of the video design is that it does not present a title, or typographical information of any type. Although videos regularly present a title, the Interchange Intro Unit 16 video does not present one. The only written information that might work as a title is the name of the file, presented on the Interchange 5th Edition website, where all extra resources are available. After the video is downloaded, the file name that is shown after the download is presented as Interchange_4_Intro_Level_Unit_16_Video[30]. Furthermore, the video design also does not include subtitles, or on-screen text, which might work as an extra aid to help learners in the process of understanding not only the context of the video, but also about the information the two characters are talking about, as can be observed in Figure 22.

Figure 22 – Absence of Topographical Signaling Device on Interchange Intro Unit 16 Video

Source: Interchange 5th Edition webpage (2021)

The video design seems to be restricted to providing learners with the target language information by means of the oral language. Once the video is designed in a fashion that presents a conversational situation in which two friends meet at school and talk about activities one of them did the previous weekend (resembling a real-life conversation between two friends), two

[30] Although the video files present the name Interchange 4 (Interchange 4th Edition) on their names, they are aimed at accompanying the textbooks from the 5th edition as well.

features can be observed as the characters have their conversation. The first is associated to the use of different intonations while they talk to aid the process of comprehension, and the second is the use of conversational marks possibly to provide the interaction a real-life-like conversation aspect.

Concerning the use of different intonation, it is possible to observe that such feature is used throughout the video, as can be observed in the following examples.

> "Oh… Well, that's kind of a long story." (0'16)
> "And? What did he say?" (0'40)
> "He asked me to go to the movies with him […] that night." (0'43)
> "No, I couldn't. Ahhm, because my little sister, I had to babysit. So then, he asked if I want to go to the basketball game the next night…" (0'54)
> "What? Are you crazy? Why? (**Girl laughs**) What was your excuse?" (1'05)
> "No excuse. I told him that I like him but I really do not like sports… But, hum, he was really nice about it, and he even asked me to go to the art fair with him on Sunday." (1'20)
> "On Sunday. I asked him to go to the amusement park with us, and he accepted my invitation, he said "I'd love to go!"" (1'48)
> **"**So you all went to the amusement park on Sunday." (1'51)
> "Well, I'm not sure…" (1'56)
> "Wohh, We were on this big roller coaster." (2'01)
> "I felt terrible about it. But then my sister came over and she had Greg's wallet. He left it at the restaurant where we had lunch before we went on the roller coaster. She picked it up and kept it for him." (2'32).
> "Yeah, she got two (XXX)* chocolate ice cream. She loved it. But he didn't." (2'46)
> "She accidentally dropped her ice cream on him. She got chocolate ice cream all over Greg's new shirt, she ruined it." (2'57)
> "[…] have this huge test the next day and I'm gonna study with all my classmates…" (3'50)

Based on Hasler, Kersten and Sweller (2007), the utterances provided above seem to work as an aid to help direct learners' attention to the information being presented. As the video presents a conversational scene in which no written language is used to help learners in the process of understanding the information, the only aid they have while watching is the oral information provided while the two characters are talking, more precisely, while the girl is reporting the events that took place on the previous Sunday. The use of intonation (the words that are produced with an emphasized intonation are indicated by underscore) in order to emphasize the content words in the utterances produced by them can be defined as the only device that helps learners in the process of apperceiving the information provided, as explained by Plass and Jones (2005), and building coherent mental representations (MAYER, 2005a; 2009) of the sequence of events described. When the girl produced the utterance "No excuse. I told him that I like him but I really do not like sports… But, hum, he was really nice about it, and he even asked me to go to the art fair with him on Sunday.", it is possible to observe that the intonation is applied to the parts "really do not", "art fair" and "Sunday" in order to help

learners build an association between basketball, which had already been emphasized in a previous utterance, and the girl expressing her dislike about sports, followed by the expressions art fair and Sunday, indicating a different activity done.

The intonation design, in which learners' attention is directed to specific parts of the utterance can also be observed when the girl produces "On Sunday. I asked him to go to the amusement park with us, and he accepted my invitation, he said "I'd love to go!"". The emphasis is applied to the words Sunday, asked, accepted and love, indicating that these words can be considered key words in the context of invitations, once they are associated to the day of the week, the invitation and the response. The objective of this video may be for learners to understand the main ideas and the events that took place as the girl reports them. Such objective may also be a reason why emphasis is applied to key words, according to Mayer (2005a) and (2009), in order to direct learners' attention to them and help them build coherent mental representations not only of the language use, but (probably) mostly of the events described. Such strategy may also have been adopted in order to aid learners in the process of focusing on meaning, in order to picture the events and, as discussed by Mayer and Moreno (2010) and Sweller (2010), to prevent them from extraneous processing probably resulting from focusing of the language structure.

The use of conversational marks is also observable in the oral language design of the video. The insertion of laughs (sometimes overlapping Peter's utterances) while they are talking, and conversation marks, such as "Oh…, Ahhm, Hum" and "Wohh" can be argued to be inserted in the video oral language design in order to resemble a real-life conversation, which would be easily encountered outside the classroom. The video design which provide real-life-like conversation may also aid learners in developing their listening strategies and listening for the gist, that is, even though they might have difficulties in understanding certain pieces of information, they (probably) would be able to understand the main gist of the conversation. Moreover, their understanding might be aided because the video is in consonance with the language topic worked in unit 16 in the textbook, and also because it is the last unit of the level Intro, that is, learners would have enough linguistic proficiency for the comprehension of the information presented.

Concerning the use of imagery features to work as aiding learners in the process of understanding the information provided in the video, it does not present images or graphs on its design. Once the target information is presented to learners in the form of a conversational situation, in which two friends are talking, the video from Interchange Intro Unit 16 can be considered to use only motion picture to provide the information to learners. Moreover,

concerning the visual structure, the video can be considered of a simplistic design due to the fact that there is no use of visual or editing effects to make specific pieces of information salient to learners, that is, to direct their attention to the language being presented, as extensively discussed by Schmidt (1990), Lorch (1989), Koning *et al.* (2010), and Hasler, Kersten and Sweller (2017).

Segmenting

Although the Interchange Intro Unit 16 video presents the target language in a conversational situation, with the information in a transient flow (LEAHY; SWELLER, 2016), based on Ibrahim *et al.* (2012), it is possible to argue that it is designed in segmented structure, as presented in Chart 12.

Chart 12 – Segmentation Structure in Interchange Intro Unit 16 Video

Segments	Information Presented
Segment 1	**Peter:** Hey, how's it going? **Girl:** (Laughs) I'm good, Peter, how are you? **Peter:** I'm fine. Oh, hey, did you see Greg last weekend? **Girl:** Oh… Well, that's kind of a long story. **Peter:** Well, we have plenty of time before class… Come on, let's go outside! **Girl:** Okay.
Segment 2	**Peter:** I'm all ears! **Girl:** Okay. Well, I saw Greg here at school on Friday. **Peter:** Yeah? Did you talk to him? **Girl:** Yeah, but he was in a hurry, so he said "Can I call you later?" and I said "Okay" and he did. **Peter:** And? What did he say? **Girl:** He asked me to go to the movies with him that night. **Peter:** Great, so you went to the movies together. **Girl:** No, I couldn't. Ahhm, because my little sister, I had to babysit. So then, he asked if I want to go to the basketball game the next night… **Peter:** Oh, so you went to the game Saturday night. **Girl:** No, I turned him down…

Segments	Information Presented
	Peter: What? Are you crazy? Why? (**Girl laughs**) What was your excuse? **Girl:** No excuse. I told him that I like him but I really do not like sports… But, hum, he was really nice about it, and he even asked me to go to the art fair with him on Sunday. **Peter:** So you went to the art fair… I was there but I didn't see you. **Girl:** No, we didn't go. I couldn't. On Sunday I had planned with my family to go to the amusement park. **Peter:** So you didn't see him. **Girl:** Yes, I did. **Peter:** Ok, wait a minute. (**Girl laughs**) I'm confused. When did you see him? **Girl:** On Sunday. I asked him to go to the amusement park with us, and he accepted my invitation, he said "I'd love to go!"
Segment 3	**Peter:** So you all went to the amusement park on Sunday. **Girl:** Uhum. **Peter:** That's cool. How was it? **Girl:** Well, I'm not sure… **Peter:** What do you mean you're not sure, what happened? **Girl:** Wohh, We were on this big roller coaster. **Peter:** Was it fun? **Girl:** Yes, but, I got sick and I had to sit down for a while. **Peter:** What did Greg do? **Girl:** He offered to buy me a soda to help settle my stomach, then he realized that his wallet was gone, it fell out of his pocket on the roller coaster. **Peter:** Oh, that's too bad. **Girl:** I felt terrible about it. But then my sister came over and she had Greg's wallet. He left it at the restaurant where we had lunch before we went on the roller coaster. She picked it up and kept it for him. **Peter:** What a relief! Was he happy? **Girl:** Oh, yeah! He even offered to buy her ice cream to say thank you. **Peter:**

Segments	Information Presented
	Did she like that? **Girl:** Yeah, she got two (XXX)* chocolate ice cream. She loved it. But he didn't. **Peter:** What do you mean? **Girl:** She accidentally dropped her ice cream on him. She got chocolate ice cream all over Greg's new shirt, she ruined it. **Peter:** Oh, no! **Girl:** He was cool about it, but I'm just afraid Greg didn't have a very good time. **Peter:** That's too bad. **Girl:** Yeah…
Segment 4	(**Girl'sTelephone rings**) **Peter:** Maybe it's Greg. **Girl:** Very funny! Aahh! (Showing the cellphone to Peter) **Peter:** I told you! Answer it! **Girl:** Greg? Hi! Oh, I'm fine, thank… You did? Oh, Oh, I had a very good time too. You're welcome! What? Really? Yes, sure I'd love to… No no, tomorrow's bad, yeah, I have dance class… haum… no my grandparents are coming over Wednesday night it's their party of anniversary… and ah, no, not on Thursday because I have this huge test the next day and I'm gonna study with all my classmates… Friday? Let me see… Hum, no, I have a dentist appointment in the morning and then I…

*(XXX) – Unintelligible spoken expression.
Source: The researcher (2021)

As aforementioned, once the video from Interchange Intro Unit 16 does not present any written language information in its design, the segmentation structure is restricted only to the oral language and to the visual aids, and is organized in 4 segments.

For being a video presenting a conversational situation, in which two friends interact, the first segment may have the function of opening the conversation, introducing the topic they will talk about. In addition, it may also have the function of introducing the next segment, as can be observed.

(Segment 1)

Peter:
Hey, how's it going?
Girl:

(Laughs) I'm good, Peter, how are you?
Peter:
I'm fine. Oh, hey, did you see Greg last weekend?
Girl:
Oh… Well, that's kind of a long story.
Peter:
Well, we have plenty of time before class… Come on, let's go outside!
Girl:
Okay.

Based on the language presented, segment 1 may have the function of opening the conversation when the characters encounter in the school hallway and greet each other using the utterances "Hey, how's it going?" and "I'm good, Peter, how are you?", which expectedly are part of a conversation when two acquainted people encounter. Moreover, the utterances "[…] did you see Greg last weekend?" and "well, that's kind of a long story." suggest that the two friends are going to exchange information about the events that took place the previous weekend. Furthermore, the first utterance provides learners with the target language structure used to make questions concerning the past with the use of the auxiliary did and the time expression last weekend, possibly, as discussed by Schnotz (2005), and Ayres and Paas (2007), having the function of activating their prior knowledge concerning the language used to make questions about part events. The utterance "[…] we have plenty of time before class… Come on, let's go outside!" suggests that, as they are changing from inside the school to outside it, there is also a change in segments, that is, they will start talking about the events that took place the previous weekend.

The oral language in segment 2 has the function of opening the segment, and also providing learners with information about the events being narrated by the girl, as can be observed.

(Segment 2)

Peter:
I'm all ears!
[…]

The first utterance by Peter "I'm all ears!" provides evidence that he is encouraging the girl to start talking about the events of the previous weekend, in this case, about whether she saw Greg, as this information was introduced in the previous segment. As segment 2 proceeds, it is possible to observe that it has two main functions. The first is to provide learners with examples of the target language grammatical structure to ask about and report events of the past, as can be observed.

(Segment 2)

[…]

Girl:
Okay. Well, I saw Greg here at school on Friday.
Peter:
Yeah? Did you talk to him?
Girl:
Yeah, but he was in a hurry, so he said "Can I call you later?" and I said "Okay" and he did.
Peter:
And? What did he say?
[…]
Peter:
Great, so you went to the movies together.
[…]
Peter:
Oh, so you went to the game Saturday night.
[…]
Peter:
So you didn't see him.
Girl:
Yes, I did.
Peter:
Ok, wait a minute. (**Girl laughs**) I'm confused. When did you see him?

(Segment 3)

Peter:
So you all went to the amusement park on Sunday.
[…]
That's cool. How was it?
[…]
Peter:
What do you mean you're not sure, what happened?
Girl:
Wohh, We were on this big roller coaster.
Peter:
Was it fun?
Girl:
Yes, but, I got sick and I had to sit down for a while.
Peter:
What did Greg do?
Girl:
He offered to buy me a soda to help settle my stomach, then he realized that his wallet was gone, it fell out of his pocket on the roller coaster.
[…]
Girl:
I felt terrible about it. But then my sister came over and she had Greg's wallet. He left it at the restaurant where we had lunch before we went on the roller coaster. She picked it up and kept it for him.
Peter:
What a relief! Was he happy?
Girl:
Oh, yeah! He even offered to buy her ice cream to say thank you.
Peter:
Did she like that?
Girl:
Yeah, she got two […] chocolate ice cream. She loved it. But he didn't.
[…]
Girl:
She accidentally dropped her ice cream on him. She got chocolate ice cream all over Greg's new shirt, she ruined it.

[…]
Girl:
He was cool about it, but I'm just afraid Greg didn't have a very good time.
[…]

The utterances presented above show evidence that the focus of segments 2 and 3 is to refer to past events when with the use of main verbs saw, was, were, said, went, got, offered, fell and left in the simple past tense in affirmative sentences, and the use of the auxiliary did in questions and didn't in negatives. Moreover, by providing examples of the target language used to report past events, the segments can be examined as aiming at supporting learners in engaging in generative processing, according to Mayer and Moreno (2010), and Sweller,(2010), the processing that leads to learning, of the target language (simple past, vocabulary and the context itself), more specifically suing the verbs in the past in order to refer to experiences one has had.

The second function that can be attributed to segment 2 is to provide learners with information about invitations and excuses, as can be observed.

(Segment 2)

[…]
Girl:
He asked me to go to the movies with him that night.
[…]
Girl:
No, I couldn't. Ahhm, because my little sister, I had to babysit. So then, he asked if I want to go to the basketball game the next night…
[…]
Peter:
What? Are you crazy? Why? (**Girl laughs**) What was your excuse?
Girl:
No excuse. I told him that I like him but I really do not like sports… But, hum, he was really nice about it, and he even asked me to go to the art fair with him on Sunday.
[…]
Girl:
No, we didn't go. I couldn't. On Sunday I […] with my family to go to the amusement park.
[…]
Girl:
On Sunday. I asked him to go to the amusement park with us, and he accepted my invitation, he said "I'd love to go!"

As can be observed, providing learners with information about invitations is also part of the oral language design in segment 2. The utterances "He asked me to go to the movies with him that night.", "So then, he asked if I want to go to the basketball game the next night…", "[…] he even asked me to go to the art fair with him on Sunday.", "I couldn't" and "I'd love to go!" provide examples of the language used to make, accept and decline to invitations.

Because the video presents a conversational situation in which the two characters are talking about past events one of them went through, instead of presenting the invitations as

direct speech, they are presented as reported accounts, which, in turn, may result more difficult for learners to comprehend or to attribute meaning to. Moreover, the girl also reports the excuses she provided for declining the invitations with the utterances “No, I couldn’t. Ahhm, because my little sister, I had to babysit.”, “I told him that I like him but I really do not like sports…” and “On Sunday I had planned with my family to go to the amusement park”.

The main function of segment 2, thus, seems to be concerned primarily with providing learners with information about the events that took place by means of using the language in the past for being a report about the previous weekend. Another evidence that suggests such interpretation is that a part of the information is not presented in direct speech, but as reported accounts. Instead, the interaction between the two characters unfolds in a manner that (expectedly) assists the understanding of the information as a whole to help learners in building coherent mental representations of the activities mentioned in past and the excuses provided, as explained by Mayer (2005a) and (2009).

The last segment, segment 4, has the function of ending the video. For such, it provides further examples of excuses that can be given to invitations. As the character is talking on the phone and it is not possible to listen to her interlocutor, only her answers can be heard, which are presented as a whole, as can be observed.

(Segment 4)

> […]
> Greg? Hi! Oh, I’m fine, thank… You did? Oh, Oh, I had a very good time too. You’re welcome! What? Really? Yes, sure I’d love to… No no, tomorrow’s bad, yeah, I have dance class… haum… no my grandparents are coming over Wednesday night it’s their party of anniversary… and ah, no, not on Thursday because I have this huge test the next day and I’m gonna study with all my classmates… Friday? Let me see… Hum, no, I have a dentist appointment in the morning and then I…

Once the main objective of the Interchange Intro Unit 16 video can be to provide learners with information about the events a person is narrating concerning making invitations, accepting, declining and providing excuses, it ends with the girl receiving a phone call, suggesting she is receiving an invitation and declining due to the excuses she provides. Such ending design can have the function of recapitulating different ways of declining invitations and providing excuses. Another function of segment 4 might be to end the video with a rather comic situation, in which the girl gives a massive amount of excuses.

Concerning the use of visual aids in the segmentation structure, the video does not present elaborated visual effects. The only effects detected in the video design are seen in the beginning and the ending. The beginning of the video is designed with the fading in effect and the ending with the fading out effect. The transition between the scenes are also designed with

the use of clear-cut effects, making the video a considerably simplistic piece of audiovisual material in its visual design.

Coherence

Once coherence deals with analyzing features that are part of the video design which can be considered irrelevant for aiding learners' process of understanding, thus hurting learning as a result of causing extraneous processing (AYERS; PAAS, 2007; MAYER, 2009; MAYER; MORENO, 2010; SWELLER, 2010), it is possible to analyze that in the Interchange Into Unit 16 video, the amount of features that can cause learners to engage in extraneous processing can be considered of a limited amount.

Regarding the use of typographical information that might be considered irrelevant, the video does not present irrelevant written language that might hurt learning due to the amount of information provided.

Since the video presents (verbal) information only orally, it is possible to observe that the language presented might be difficult for learners to build coherent mental representations of the events being narrated. Although the video is aimed at presenting learners with a conversational situation resembling a real-life interaction between two friends, it is important to consider that it is aimed at learners that are in the Intro level, that is, that may be considered to have a low proficiency level. Based on Mayer and Moreno (2010) and Sweller (2010), the high level of intrinsic load that the information of the video may present to learners might hurt their comprehension of the information, thus, hurting learning by exceeding their working memory (WM) resources in trying to understand and make sense of all the information, also discussed by Kalyuga (2010) and Sweller (2010).

Providing information about the invitations the girl received and the excuses for declining to those may be considered difficult due to the fact that the target language used to making invitations, accepting/declining and providing excuses are not presented as direct speech between the two characters, but as reported information. Such oral design can be considered of a high level of intrinsic load due to the element interactivity of the information presented. Grounded on Ambrose *et al.* (2010), for a proper understanding and building coherent mental representations of the events being reported, learners need to have (and activate) prior knowledge about the language structure used to refer to past events, to making invitations, accepting or declining invitations and providing excuses. Moreover, learners would be expected to discern between the direct speech and the reported information, once the two characters (Peter and girl) are talking about a third character (Greg). Extraneous processing could be a possible result due to the high level of information processing demands from

learners' WM, as extensively discussed by Kalyuga (2010), Mayer (2009), and Sweller (2010) in which they would have to listen to the information and convert it into meaningful pictures in order to make sense of the events narrated. Depending on learners' prior knowledge, their understanding could be low or even none considering that they are in the introductory level (AMBROSE *et al.*, 2010).

In regard to the visual aids used that might be considered irrelevant, the video does not present visual effects. Based on Ayers and Paas (2007), the editing effects applied to the segments transition may not be considered irrelevant once the fading in and out (in the beginning and ending of the video) do not seem to cause extraneous processing, as well as the clear-cut transition between segments 2 through 4.

4.1.2.3 Video from the Textbook Interchange 1 – Unit 8

The video from Interchange 1 Unit 8 is two minutes and fifty-three seconds long, presenting a conversational situation in which a couple and a realtor are talking about an apartment and the neighborhood where it is located. As the video is intended to accompany unit 8 from the textbook, the target language focus is on the description of the places of a neighborhood, or city as is shown in the video analysis.

Signaling

Concerning the use of typographical cues to work as devices to aid learners in the comprehension of the information presented, some aspects can be observed. The first is the fact that the Interchange 1 Unit 8 video does not present a title. Considering that a video title can work, as explained by Plass and Jones (2005) as an initial written language input and as a contextualization device and prior knowledge activator based on Schnotz (2005), and Ayres and Paas (2007), the lack of a title can fail in preparing learners for the information they will watch. The only observable title (whether it becomes possible to refer to it as a title) is the name of the video file as Interchange_4_Level_1_Unit_8_Video after being downloaded. The issue, however, is that no specific information concerning the video is provided on the file name, making it questionable whether it can be referred to as a title.

Concerning other types of typographical cues, the video does not present subtitles or on-screen written language. Alternatively, the typographical cues that can be observed in the video design are the names of places, as presented in Figure 23.

Figure 23 – Typographical Cues in Interchange 1 Unit 8 Video Design

Source: Interchange 5th Edition website (2021)

It is observable in the figure that throughout the video different places are shown such as restaurants, coffee shops, bookstores and grocery stores. It is possible that the places are presented in the video for learners to establish an association between them and the description provided by the realtor, once they were previously mentioned in the video, when the couple is talking to her and she describes the neighborhood. In addition, learners' understanding of the places they are seeing is further aided by the target lexical elements associated to them being presented, such as the words restaurant, bookstore, or grocery store. As for directing learners' attention to them, according to Lorch (1989) and Koning *et al.* (2010), the typographical cues presented are positioned in a strategic location on screen, or on a specific angle, so to make them salient and learners' attentional resources can be directed to them, making it possible for them to identify the images as being storefronts, possibly aiming at creating a integration between the written language input and physical places, as discussed by Plass and Jones (2005).

Regarding the use of oral language input to work as a device to direct learners' attention some aspects can be detected in the video design. As the Interchange 1 Unit 8 video presents a conversational situation, and the target language is presented to learners in the form of a conversation between the three characters (the realtor and the couple), two features can be observed. Based on Hasler, Kersten and Sweller (2007), and Koning *et al.* (2010), the first feature is associated to the intonation used in order to direct learners' attention to specific oral information, as can be observed in the following examples (the emphasized words are underlined).

> "We really like the apartment!" (0'12)
> "So, it's a safe neighborhood." (0'28)
> "Oh, yes, it's very safe and quiet. […]" (0'31)
> "What about… public transportation?" (0'46)
> "The public transportation is excellent. […]" (0'48)
> "What a great neighborhood!" (1'50)
> "Look, Luiz, there's a really nice grocery store." (2'10)

As can be observed, emphasis is given to adjectives used to qualify the places or services that are part of the neighborhood, such as safe, quiet, excellent, great and really nice. Based on Lorch (1989) and Koning *et al.* (2010), the oral language design, in which the positive adjectives are made salient, may have been chosen in order to direct learners' attentional resources to the target language used to emphasize a feeling, describe a neighborhood and ask about the services available, as is the case of the question about public transportation. In addition, not only is learners' attention directed to the adjectives *per se*, but also to the target language used to describe a neighborhood and aid their understanding of the interaction, in which the three characters use positive adjectives to characterize not only the places and services offered, but also the neighborhood as a whole and how they feel about it. In short, emphasizing certain words in the utterances can work as aiding learners in building a coherent mental representation (MAYER, 2005a; 2009) of the complete situation, that is, as a whole.

The use of intonation for emphasizing specific pieces of information is also detectable when the realtor exemplifies the types of restaurants there are in the neighborhood, as can be observed.

> "Oh, yeah, there are a lot of restaurants. Just take a walk down the street to the end, and you'll see there are lots of Greek and Italian restaurants. There's Indian, Chinese, everything." (1'05)

The oral language in the utterances presented is designed in a fashion that emphasizes the types of restaurants that can be seen in the neighborhood following an (oral) enumeration style. In addition, the saliency provided can also be argued to aim at providing learners with examples of the target language aspects used when describing restaurants concerning the kinds of food they serve, in this case, by their nationalities, or when presenting a list of items.

Another function of the oral language design concerns with conversational strategies used, such as showing delight and the use of repetition.

> "Hum, sounds great." (1'07)
> "Ok, let's take a look. […]) (1'10)
> "I love that bookstore!" (1'48)
> "[…] I really like this neighborhood." (2'19)
> "There's a movie theater, and restaurants…" (1'55)
> "Movie theater, restaurants…" (1'56)

The examples provided can have the function of directing learners' attention (SCHMIDT, 1990; LORCH, 1989) to understanding the fact that the woman is happy with the information being provided by realtor about the neighborhood. Her enthusiasm can be further observed as she constantly uses positive words to describe the places they see and to express her feelings concerning the places and the neighborhood as the couple walks around to get to know the places nearby.

The use of repetition can also be observed to work as a device to direct learners' attention to the information. When the woman describes the places they can see nearby as being part of the neighborhood (movie theater, restaurant), the man immediately repeats names of the places. The oral language design with use of repetition can have the function of clarifying to learners examples of the places found not only in their neighborhood, but, as discussed by Zarei and Khazaie (2011), presenting vocabulary that can (usually) be used when describing a neighborhood regarding the places found on it.

Concerning the use of visual aids to work as signaling devices, the only detectable use of imagery aids to facilitate the process of seeing the information and building coherent mental representations of the places is the insertion of images of store fronts in the design.

Figure 24 – Imagery Signaling Devices in Interchange 1 Unit 8 Video Design

Source: Interchange 5th Edition website (2021)

The insertion of the imagery information in the video design is conceived by the use of images store fronts of the kinds of places which were talked about in the interaction between the realtor and the couple. Such imagery aids can have the function of helping contextualize the conversational situation presented in the video. Since the video provides learners with a

contextualized situation in which people describe the places and characteristics of a neighborhood, the video displays images showing different places found in a neighborhood (restaurants, coffee shops, grocery stores and jewelry stores). The insertion of the imagery information showing the places may be in consonance with the names of places previously mentioned by the realtor and the woman. According to Mayer (2005a) and (2009), this may have the function of integrating the visual information (store fronts in the written forms, see Figure 24) to the names of the places previously mentioned, thus aiding learners in building coherent mental representations of the information.

As regards to the use of visual effects to work as signaling devices, it is possible to argue that no visual effects are used in the video design.

Segmenting

The Interchange 1 Unit 8 video, although presenting information in a transient flow (LEAHY; SWELLER, 2016), can be defined being designed in segmented structure (SWELLER, 2010; IBRAHIM, 2012; IBRAHIM *et al.*, 2012), organized in four segments, as is presented in Chart 13.

Chart 13 – Segmentation Structure in Interchange 1 Unit 8 Video

Segments	Information Presented
Segment 1	**Woman:** We really like the apartment! **Man:** Yeah, it's good. Very roomy. **Realtor:** Good! **Woman:** What's the neighborhood like? **Realtor:** Oh, it's a very special neighborhood. There's a real mix of people here. There are older people, young families and students. There's a lot of different cultures. **Woman:** So, it's a safe neighborhood. **Realtor:** Oh, yes, it's very safe and quiet. There isn't much noise. Wuh, usually (**Laughs**). **Woman:** (**Laughing**) Is there much crime? **Realtor:** Oh, no, there isn't much now. **Woman:** Now? **Realtor:** Well, there were some problems, but that was ten years ago. **Woman:** Ok…

Segments	Information Presented
	Man: What about… public transportation? **Realtor:** The public transportation is excellent. It's just a few minutes to downtown. **Woman:** We like to eat out. Are there many restaurants and coffee shops? **Realtor:** Oh, yeah, there are a lot of restaurants. Just take a walk down the street to the end, and you'll see there are lots of Greek and Italian restaurants. There's Indian, Chinese, everything. **Woman:** Hum, sounds great. **Man:** Ok, let's take a look. Thank you for your help. **Realtor:** It was my pleasure. Give a call this afternoon cause this apartment won't last long. **Woman:** Ok. Thank you. Bye. **Man:** Bye. **Realtor:** Bye!
Segment 2	**Woman:** I love that bookstore! **Man:** Yeah? **Woman:** What a great neighborhood! **Man:** Yeah. **Woman:** There's a movie theater, and restaurants… **Man:** Movie theater, restaurants… **Woman:** There's a furniture store… **Man:** Uhum, yeah… **Woman:** And, there's a jewelry store! My birthday is coming next month. **Man:** Next month, yeah, I know (**laughing along with woman**)
Segment 3	**Woman:** Look, Luiz, there's a really nice grocery store. **Man:** Yeah, it's a nice grocery store. **Woman:** And there are a lot of really good coffee shops! I really like this neighborhood. **Man:** Yeah, it's… really great… **Woman:**

Segments	Information Presented
	Yes! **Man:** Wow! (**Looking a guitar on a shop window**) That's an amazing guitar!
Segment 4	**Woman:** So, can we take the apartment? **Man:** (**looking at her, and back at the guitar, then back to her**) sure. **Woman:** (**Laughing**) **Man:** Why not?! Let's go for it!

Source: The researcher (2021)

Concerning the segmentation structure, it is possible to say that, once the video is structured in reduced number of segments, the amount of language presented is concise so to aid learners in the processing of essential pieces of information, that is, based on Mayer and Moreno (2010) and Sweller (2010), to aid them in engaging in generative processing of the target language of both the contextual information presented for the understanding of the situation as a whole and the language elements provided, in his case, how to make descriptions of a neighborhood. Furthermore, there is no use of written language input presented as on-screen text or to introduce and end the segments, being the written language shown in the video presented as part of imagery information.

The information in the video is presented by means of the oral language produced by the three characters who engage in contextualized interactions. Furthermore, it is possible to observe that the segments have the function of proving learners with examples of the target language used to make questions and describe a neighborhood, as can be observed below.

(Segment 1)

[…]
Good!
Woman
What's the neighborhood like?
Realtor
Oh, it's a very special neighborhood. There's a real mix of people here. There are older people, young families and students. There's a lot of different cultures.
Woman
So, it's a safe neighborhood.
Realtor
Oh, yes, it's very safe and quiet. There isn't much noise. Wuh, usually (**Laughs**).
Woman
(**Laughing**) Is there much crime?
Realtor
Oh, no, there isn't much now.
Woman

Now?
Realtor
Well, there were some problems, but that was ten years ago.
[…]

The examples presented show evidence that the aim of the segment is to provide learners with examples of the target language used to making a question about one's neighborhood and to answering it with structure of questions and sentences with there to be. In addition, once the context of the situation presented in the video is about a couple looking for a new apartment, they are interested in learning information about the place where it is located. When asked about the neighborhood, the realtor provides the couple with an amount of information concerning the people and their origin as well as about issues associated to safety with the utterances "[…] There are older people, young families and students. There's a lot of different cultures." and "Oh, no, there isn't much now."

Another function that can be observed in segments 1 and 2 is the use of examples of making questions and providing information concerning the services that are available in the neighborhood.

(Segment 1)

[…]
Man
What about… public transportation?
Realtor
The public transportation is excellent. It's just a few minutes to downtown.
Woman
We like to eat out. Are there many restaurants and coffee shops?
Realtor
Oh, yeah, there are a lot of restaurants. Just take a walk down the street to the end, and you'll see there are lots of Greek and Italian restaurants. There's Indian, Chinese, everything.
[…]

(Segment 2)

Woman
I love that bookstore!
[…]
Woman
What a great neighborhood!
[…]
There's a movie theater, and restaurants…
[…]
Woman
There's a furniture store…
[…]
Woman
And, there's a jewelry store! My birthday is coming next month.
[…]

The excerpts provided show evidence that the segmentation structure also has the function of providing learners with the target lexis (restaurant, bookstore) associated to places that are (normally) used when describing a neighborhood. Moreover, it is possible to examine that the function is to further present the target language structure used when describing the existence of services in a neighborhood with the use of there is/are as can be observed in the utterances "There's a movie theater, and restaurants…" and "And, there's a jewelry store! […]". Grounded on Mayer and Moreno (2010), and Sweller (2010), it can be argued that the repeated exposure of learners to the language structure can have the function of helping them in the processing of the essential information of both the situation (concerned about the neighborhood) and the language structure, (probably) for later use.

The oral language design can also be observed to have the function of ending segments as can be observed.

(Segment 1)

[…]
Man
Ok, let's take a look. Thank you for your help.
Realtor
It was my pleasure. Give a call this afternoon cause this apartment won't last long.
Woman
Ok. Thank you. Bye.
Man
Bye.
Realtor
Bye!

(Segment 3)

Man
Wow!

Segment 1 is ended with the oral language used in farewells when the couple is ending the conversation with the realtor by thanking her with the utterances "Ok. Thank you. Bye." and "Bye!". Moreover, it is possible that the utterance "Ok, let's take a look. […]" also works as introducing the next segment, in which, after the realtor describes the neighborhood concerning the services available, the couple walks around to get to know the place. Such structure provides evidence that first presenting the places orally can function as to prevent learners from engaging in extraneous processing (MAYER, 2009; MAYER; MORENO, 2010) for presenting information in small amounts, each portion in a different segment.

The structure in which the information is presented orally, as the realtor describes the nearby places, followed by the visual information showing the places, can be argued to present the target language to learners in different portions, first orally, second, visually, probably, to

first present (or revisit) the oral names of the places, and later to provide visual information in order to aid the proper integration of the two types of information (oral + visual), as explained by Plass and Jones (2005) concerning multimedia materials for L2 learning.

The last segment presents the closing of the situation, in which the couple decides to take the apartment.

(Segment 4)

Woman
So, can we take the apartment?
[…]
Man
Why not?! Let's go for it!

Segment 4 can be defined as having the focus of closing the situation, then, according to Plass and Jones (2005), providing learners with oral language input. After seeing the apartment and being provided with an amount of information of the characteristics of the neighborhood concerning the people, their neighbors and the services, they finally decide to move in.

As regards the segmentation structure with the use of visual aids, it is possible to observe that segment 2 is introduced with imagery aids. Images of the fronts of restaurants, stores and a number of books on a table outside of bookstore are presented before the couple comes out of the bookstore and engages in a conversation. Another use of imagery information can be observed in the last segment, when a guitar on a shop window is shown and the man becomes extremely excited about seeing it, suggesting that he likes to play the instrument, or is fond of music. After the image is shown, he runs pulling the woman to the window to see the guitar before they engage in the last part of the conversation, thus, ending the video. The design of the last segment may have the aim of evoking a comic situation, once the man had been apathetic throughout the whole video, showing little interest on the pieces of information provided, or no excitement concerning the apartment and the neighborhood, being the guitar the only element that he seemed to have liked or being interested in, unlike the woman, who showed excitement about everything.

As aforementioned, the Interchange 1 Unit 8 video seems to be structured in a segmentation fashion in order to, based on Mayer and Moreno (2010) and Sweller (2010), aid learners' essential processing in two different ways: the first is concerned with their comprehension of the communicative situation as a whole. According to Plass and Jones (2005), with the integration of the language input in the written and oral forms and the imagery information they can understand the situation and what is happening, such as a couple visiting

an apartment and their appreciation of the surroundings. The second is concerned with proving learners with the target language structure and lexis used when describing the places and services that are part of a neighborhood or city (*e.g.* restaurants, stores, public transportation), and the structure used to describe the existence of these (*e.g.* there are restaurants, grocery stores).

Coherence

Considering the pieces of information that are part of the video design that can be considered irrelevant for they do not aid learning, regarding written language, it is possible to say that the video design does not present irrelevant written information of any kinds once the information in the video is provided in mostly in the oral language.

Concerning the oral language, it can be argued to be fully relevant in the video design for two reasons. The first is that it provides learners with the necessary information for them to understand the conversational situation that is taking place as a whole – people describing a neighborhood. The second is that the language is concisely focused on providing the pieces of information – lexical and structures.

As for the use of visual and aural aids, two musical soundtracks are part of the video design. The first soundtrack is a joyful instrumental music that starts playing at the end of segment 1, when the couple says goodbye to the realtor and walks away, and at the beginning of the second segment, when a number of storefronts appears on screen. This soundtrack can be considered irrelevant once it does not seem to aid learners in understanding the information being provided. The second soundtrack is presented when the man sees the guitar on the shop as the sound of guitar chords probably to establish the idea of surprise. Differently from the first soundtrack, the guitar chords soundtrack can be considered relevant once, as the image of a guitar is shown, learners can clearly understand that the man became interested in it, also being integrated to the image of the guitar. The issue that rises is to what extend such soundtrack would aid learning. Although the two music soundtracks can be considered not to aid learning, thus becoming unnecessary in the video, they are presented when no oral information is provided, thus possibly not having a negative effect on learners' working memory (WM) resources provoking extraneous processing (AYERS; PAAS, 2007; KALYUGA, 2010; MAYER, 2009; MAYER; MORENO, 2010; SWELLER, 2010). Also, they were instrumental music, with no words to compete for language input.

4.1.2.4 Video from the Textbook Interchange 2 – Unit 8

The video from Interchange 2 Unit 8 is four minutes and thirty-three second long. It is designed in a fashion presenting a TV show host interviewing people in a supermarket about thanksgiving, and is in consonance with the main topic of unit 8 in the textbook, which is concerned with celebrating holidays, as can be observed in the video analysis.

Signaling

Concerning the use of written language to work as signaling devices used in the video design, it is possible to observe that the video does not present a formal title. The only piece of written information that might be considered as title is the name that is displayed on the file name when it is downloaded from the Interchange 5th Edition website, that is Interchange_4_Level_2_Unit_8_Video. The insertion of a title in the video design could work as a prior knowledge activator. Furthermore, grounded on Schnotz (2005), and Ayres and Paas (2007), as being a complementary material to unit 8 from the textbook, a title could also activate their prior knowledge concerning the topic in case they had already worked the topic in class as well as aid the knowledge building process.

Still regarding the use of written language as signaling device, subtitles are another feature that is not part of its design. The absence of subtitles could be debatable for the fact that the oral information presented could be considered dense to some degree, or even the pace with which the interviewees provide information could be considered beyond learners' working memory (WM) resources, resulting in extraneous processing (AYERS; PAAS, 2007; MAYER, 2009; MAYER; MORENO, 2010; SWELLER, 2010). Considering that the video depicts a number of interactions in which a TV host talks to a number of people to learn information about the way they celebrate thanksgiving and the foods they eat on that holiday, the information presented is provided entirely as oral language in the form of conversations, thus, being the oral language the only linguistic input, as discussed by Plass and Jones (2005).

As regards the oral language that is part of the video design, some features can be seen as having the function of signaling devices to direct learners' attention to the information. The first feature refers to the opening of the video by contextualizing the situation and the language that will be presented.

> "Hello! And welcome to this mix episode of Dinner Table! The show about food. My name is Anthony Russo, and today we're gonna be talking about thanksgiving." (0'16)

The first oral language information provided in the video can be argued to have two main functions. The first is concerned with opening the video once the host welcomes viewers to the show; the second, with situating learners within the context or type of show they are watching by providing its name accompanied by a definition about it in "[…] episode of Dinner Table! The show about food. […]". In addition, according to Ayres and Paas (2007), the opening oral language may work as activating learners' prior knowledge about the topic (foods) as to the information they will be exposed to. Furthermore, other function of the host's oral language design is to activate learners' prior knowledge with questions directed at them and inviting them to keep on watching.

> "[…] You think everyone celebrates thanksgiving in the same way? Do you think they eat the same foods? Let's ask some people and find out!" (0'39)

The utterances provided show evidence that when the host makes questions about how people possibly celebrate thanksgiving and about the foods they may eat, the main function may be to activate learners' prior knowledge about the different ways people may celebrate a holiday or a special date, being in consonance with the topic discussed. It is important to clarify, however, that in order to facilitate the activation of learners' prior knowledge, they would be expected to already have worked with this topic in class as well as having a proper understanding of it, as argued by Ambrose *et al.* (2010). Inviting learners to continue watching in order to find out the information about people's celebrating habits also has the function of directing learners' attentional resources to the information presented (SCHMIDT, 1990; LORCH, 1989; KONING *et al.*, 2010; HASLER; KERSTEN; SWELLER, 2007).

Based on Plass and Jones (2005), the use of intonation to direct learners' attention in the process of apperceiving the information presented can also be observed in the oral language design to emphasize specific pieces of information, concerning the kinds of foods that are served and prepared, as can be observed - the words with a higher intonation are represented as underlined.

> "Oh, yeah, we have the big turkey. We usually have two of them cause it's so many of us. Cranberry sauce… ah, mashed potatoes, gravy… […]." (1'15)
> "Well, we have a turkey. We'll also have stuffing, and macaroni and cheese, and mashed potatoes, and sweet potatoes, […]." (2'13)
> "Well, we definitely have the traditional thanksgiving food, ah, turkey, stuffing, mashed potatoes… but I'm Korean, so we also usually have so sort of rice with kimchi. Kimchi is a definite must, cause we're Korean…[…]" (4'50)
> "My friends always cook the traditional thanksgiving dinner. However, I bring also a dish from my home country Venezuela." (3'46)
> "Maracuchitos, which is plantain with cheese, fry…" (3'54)

> "I usually also make a sweet from pineapple and papaya. We serve that with ice cream." (4'09)

Grounded on the study conducted by Zarei and Khazaie (2011), the function of the use of intonation can be observed to give emphasis to the food vocabulary items that are provided as interviewees answer the TV host's questions and talk about the foods that are prepared and served for thanksgiving. As discussed by Hasler, Kersten and Sweller (2007), emphasizing the lexical elements may have the function of directing learners' attentional resources to them and establish a connection between the information that the characters are providing and the topic of the textbook. Foods such as turkey, mashed potatoes, stuffing are mentioned by more than one participant, possibly for two reasons. The first is that the repeated exposure to the vocabulary, as discussed by Zarei and Khazaie (2011), throughout the video can aid learners in the process of learning the food vocabulary, and the second can be associated to the fact that learners may infer that these foods are traditional to thanksgiving due to the information the interviewees provided.

Furthermore, intonation is also used for emphasizing people's place of origin, as can be observed in the third and the forth examples afore presented, when the characters say "I'm Korean" and "[…] my home country Venezuela.". According to Innaci and Sam (2017), and Tuyen and Huyen (2019), the information about the characters' origins can have the function of contextualizing (or justifying) their choices of typical food from their home countries to be inserted in their menus. Moreover, the addition of these typical foods can also have the function of raising learners' cultural awareness concerning the foods from different countries, as is the examples of *kimchi* as a Korean dish, and *maracuchitos* as a Venezuelan dish.

Intonation is also detectable in the utterances produced by the interviewees when they talk about the probable leftovers and the cleaning as part of celebrations, as can be observed.

> "There are not a whole lot of leftovers by the time thanksgiving is done." (2'23)
> "Cleaning is definitely a part of it. You wash dishes, you clean the table, you bring it back to the kitchen." (1'28)

As can be observed, the first example provided concerns with providing information about the food leftovers that may result from preparing a large amount of food, which in turn is negatively emphasized by the interviewee. The second example provides information about the organization, or cleaning, that is a subsequent part of having any kind of celebration, in this case, of having meals. When the interviewee lists the activities involved in cleaning, he gives emphasis to them, possibly, to direct learners' attention to the steps that are part of the

organization that takes place after a celebration ends (SCHMIDT, 1990; LORCH, 1989; KONING *et al.*, 2010; HASLER; KERSTEN; SWELLER, 2007).

Another feature that can be examined in the oral language design is the use of repetition, grounded on Plass and Jones (2005), aiding in the apperception process as can be observed in these examples.

> "A lot of food! A lot of food!"
> "Lot of food, lot of food!" (1'18)
> "and, ah, very, very delicious. […]" (3'57)
> "Hum, sounds good, sounds good! […]" (4'00)

The insertion of repetition in the oral language design may also have the function of emphasizing or intensifying the information provided. For example, in the first and second utterances presented, the characters state that, in the thanksgiving meal there, the amount of food is substantial (first example), and when describing the food they prepare, they intensify the fact of being delicious. The fourth utterance can be observed as a conversational resource used by the TV host in order to show interest in the information provided by the interviewees and possibly to encourage him to provide extra information.

Recapitulating the main topic of the video is another feature observed as a signaling device in the oral language design.

> "All this talk of food is making me hungry. I'm Anthony Russo from Dinner Table, wishing you and yours a happy thanksgiving!" (4'27)

Although the last utterance presented works as signaling that the show is ending, it does not make an explicit recapitulation of all the pieces of information presented, but in a considerably concise way only by recalling the main topic of the video, that is, food.

Another function that the oral language presented can have is to (probably) establish some degree of social engagement with learners. When the host produces the utterance "All this talk of food is making me hungry. […]" he is providing personal information about him – being hungry – and also possibly granting a comic attribute to his speech. Moreover, when he repeats his name, the name of the show and the wish of a happy thanksgiving to the audience, he is making it clear to learners that the video is coming to its end, thus signaling again.

As regards the use of visual aids in order to signal the relevant information, it is possible to observe that, in the Interchange 2 Unit 8 video, the use of imagery information is presented throughout the whole video, as in Figure 25.

Figure 25 – Imagery Signaling Devices in Interchange 2 Unit 8 Video Design

Source: Interchange 5th Edition website (2021)

The imagery information may have the function of proving learners with visual examples of the foods being described aiding them in the apperceiving (PLASS; JONES, 2005) process and in integrating the oral and imagery pieces of information. These examples show a varied number of foods being prepared and served as the TV host provides information to the audience and as the interviewees provide information about the foods they prepare and eat on thanksgiving celebration. The choice of such imagery design can be to aid learners in the process of understanding the information about the foods described and to possibly to facilitate the building of coherent mental representations by integrating the oral and visual information about the foods being talked about, as discussed by Mayer (2005a, 2000), and Plass and Jones (2005) .

Segmenting

Once, based on Leahy and Sweller (2016) the information on the Interchange 2 Unit 8 video is presented in a transient fashion, it is designed in a segmented structure (SWELLER, 2010; IBRAHIM, 2012; IBRAHIM *et al.*, 2012), presenting information in eight segments as can observed in Chart 14.

Chart 14 – Segmentation Structure in Interchange 2 Unit 8 Video

Segments	Information Presented
Segment 1	**TV Show Host:** Hello! And welcome to this mix episode of Dinner Table! The show about food. My name is Anthony Russo, and today we're gonna be talking about thanksgiving.
Segment 2	**TV Show Host:**

Segments	Information Presented
	As you know, North America is a huge melting pot. You think everyone celebrates thanksgiving in the same way? Do you think they eat the same foods? Let's ask some people and find out!
Segment 3	**TV Show Host:** How would you explain thanksgiving to someone outside of North America? **Elderly Man:** Well, y'know thanksgiving's a unique holiday. It's kind of a time to sit back, relax and little reflection on how you should be thankful for what you have.
Segment 4	**TV Show Host:** Hi, what's your name? **Man:** Ah, my name is Joe. **TV Show Host:** What do your folks serve for thanksgiving day? Traditional thanksgiving day meal? **Man:** Oh, yeah, we have the big turkey. We usually have two of them cause it's so many of us. Cranberry sauce… ah, mashed potatoes, gravy… yeah it's great. **TV Show Host:** A lot of food! A lot of food! **Man:** Lot of food, lot of food! **TV Show Host:** Now, Joe you have a lot of food… lot of dessert… there must be a lot of dishes and cleanup. Do you pitch in? **Man:** Cleaning is definitely a part of it. You wash dishes, you clean the table, you bring it back to the kitchen. **TV Show Host:** So, there's gotta be a lot of leftovers. **Man:** Nothing, huh, nothing beats, leftovers after thanksgiving, mashed potatoes, turkey… **TV Show Host:** And it lasts through the weekend? **Man:** Oh, for the whole week. It's so much food.
Segment 5	**TV Show Host:** Do you have a favorite part of thanksgiving? **Woman:** My favorite part of thanksgiving is coming home… and seeing everybody. And seeing my grandmother. Well I love her very dearly. **TV Show Host:** That's always nice. Any family traditions that you've carried on throughout the years? **Woman:** The main family tradition is probably football. (**Laughs**) Because the Detroit Lions always play on thanksgiving, so my uncle always has to watch the Detroit Lions play. **TV Show Host:** And what food is served on thanksgiving?

Segments	Information Presented
	Woman: Well, we have a turkey. We'll also have stuffing, and macaroni and cheese, and mashed potatoes, and sweet potatoes, but I don't really like sweet potatoes (**Laughs**). **TV Show Host:** (**Laughs**) You got a lot of food there. Must be some leftovers. **Woman:** There are not a whole lot of leftovers by the time thanksgiving is done. **TV Show Host:** Think grandma gets all of them if there are any. **Woman:** Grandma gets of them. Grandma gets whatever she wants (**Laughs**). **TV Show Host:** That's great!
Segment 6	**TV Show Host:** Any, ah, specific dishes and a family tradition? **Woman:** Well, we definitely have the traditional thanksgiving food, ah, turkey, stuffing, mashed potatoes… but I'm Korean, so we also usually have so sort of rice with kimchi. Kimchi is a definite must, cause we're Korean… Ahm, As well as various other, ahm, Korean food we usually have. Our relatives play potluck and bring something with them, so… whatever they decide to bring… **TV Show Host:** What do you do after the meal? **Woman:** Ahm, we are usually… pretty comatosed about the food we've eaten. So we sit around and we usually talk. Sometimes, though, ah, grown ups play, ah, old Korean card games. Ahm… and the kids usually sit around and talk or we go out to a movie. That's usually we do. **TV Show Host:** Well, have a happy thanksgiving. **Woman:** Thank you!
Segment 7	**TV Show Host:** How do you spend thanksgiving alone Carlos? **Man:** I spend thanksgiving with friends. **TV Show Host:** Sounds nice. With all those nice friends, there must be some great food. What food is on the menu for thanksgiving? **Man:** My friends always cook the traditional thanksgiving dinner. However, I bring also a dish from my home country Venezuela. **TV Show Host:** And what dish is that? **Man:** Maracuchitos, which is plantain with cheese, fry… and, ah, very, very delicious. It's always a hit. **TV Show Host:** Hum, sounds good, sounds good! And how about for dessert? **Man:**

Segments	Information Presented
	I usually also make a sweet from pineapple and papaya. We serve that with ice cream. It is another hit. **TV Show Host:** Wow, sounds like a fantastic meal! **Man:** It is a big celebration.
Segment 8	**TV Show Host:** All this talk of food is making me hungry. I'm Anthony Russo from Dinner Table, wishing you and yours a happy thanksgiving!

Source: The researcher (2021)

The Interchange 2 Unit 8 video is structured in segments in order to present learners with the target language used to talk about foods, as can be observed in the examples.

(Segment 4)

[…]
TV Show Host:
What do your folks serve for Thanksgiving Day? Traditional Thanksgiving Day meal?
Man:
Oh, yeah, we have the big turkey. We usually have two of them cause it's so many of us. Cranberry sauce… ah, mashed potatoes, gravy… yeah it's great.
[…]

(Segment 5)

[…]
Woman:
My favorite part of thanksgiving is coming home… and seeing everybody. And seeing my grandmother. Well I love her very dearly.
[…]
TV Show Host:
And what food is served on thanksgiving?
Woman:
Well, we have a turkey. We'll also have stuffing, and macaroni and cheese, and mashed potatoes, and sweet potatoes, but I don't really like sweet potatoes (**Laughs**).
[…]

(Segment 6)

TV Show Host:
Any, ah, specific dishes and a family tradition?
Woman:
Well, we definitely have the traditional thanksgiving food, ah, turkey, stuffing, mashed potatoes… but I'm Korean, so we also usually have so sort of rice with kimchi. Kimchi is a definite must, cause we're Korean… Ahm, As well as various other, ahm, Korean food we usually have. Our relatives play potluck and bring something with them, so… whatever they decide to bring…
[…]

(Segment 7)

[…]
Man:
I spend thanksgiving with friends.

> **TV Show Host:**
> Sounds nice. With all those nice friends, there must be some great food. What food is on the menu for thanksgiving?
> **Man:**
> My friends always cook the traditional thanksgiving dinner. However, I bring also a dish from my home country Venezuela.
> […]
> **Man:**
> Maracuchitos, which is plantain with cheese, fry… and, ah, very, very delicious. It's always a hit.
> **TV Show Host:**
> Hum, sounds good, sounds good! And how about for dessert?
> **Man:**
> I usually also make a sweet from pineapple and papaya. We serve that with ice cream. It is another hit.
> […]

As can be observed, the main function of segments 4, 5 and 6 seems to be to provide learners with information about the foods that people traditionally prepare and eat when celebrating thanksgiving, in which the host first asks about what is served, and the interviewees provide the information asked. It is observable as the interviewees mention turkey (segments 4, 5 and 6), mashed potatoes (segments 4, 5 and 6) and stuffing (segments 6 and 7) as being traditionally served in thanksgiving meals. Segments 6 and 7 present information about the foods served, establishing an association between the foods prepared and eaten with the home country of the characters (as is the example of *kimchi*) and (as is the example of *maracuchitos*) providing the information in the two segments respectively.

It is also possible to argue that the structure of videos in segments, each one with a different person talking about thanksgiving and the foods, is not aimed at teaching a specific language structure. Alternatively, the focus may be to present the learners with the target food vocabulary associated to thanksgiving. Moreover, as argued by Ayers and Paas (2007), and Mayer (2009), there seems to be concern with providing learners with a reduced amount of information that might be new to them, preventing them from engaging in extraneous processing due to the amount of information, each interviewee provides, as also discussed by Koning *et al.* (2007), and Ibrahim, *et al.* (2012), focusing specifically on the food vocabulary.

Presenting the information in a fashion that resembles a TV show in which a host interviews a number of people on the foods they have in celebrating thanksgiving can be argued, based on Ayers and Paas (2007), to have the function of preventing learners from extraneous processing, once, the information presented is contextualized in the celebrations topic in segments, each person at a time. Furthermore, the segmented structure, in which the amount of information is reduced can aid learners in the understanding of each interviewee's information

as a whole, instead of focusing on isolated elements, as would be in a grammatical explanation, for example.

Segments 1, 2 and 8 may have the function of opening the video, contextualizing the information presented and closing it.

(Segment 1)

> **TV Show Host**
> Hello! And welcome to this mix episode of Dinner Table! The show about food. My name is Anthony Russo, and today we're gonna be talking about thanksgiving.

(Segment 2)

> **TV Show Host**
> As you know, North America is a huge melting pot. You think everyone celebrates thanksgiving in the same way? Do you think they eat the same foods? Let's ask some people and find out!

(Segment 8)

> **TV Show Host**
> All this talk of food is making me hungry. I'm Anthony Russo from Dinner Table, wishing you and yours a happy thanksgiving!

The function of segment 1 may be to open the video by contextualizing the type of video learners are watching, in this case, resembling a TV show. Moreover, the TV show-like video is made clear when the host produces the utterance "[…] welcome to this mix episode of Dinner Table! The show about food" also establishing information about the topic the video will deal with, that is, food. Segment 2 may have the function of narrowing the context of the video to the topic thanksgiving, especially when the host provides the utterances "You think everyone celebrates thanksgiving in the same way? Do you think they eat the same foods? Let's ask some people and find out!", encouraging learners to reflect about the way people celebrate and the foods they eat, and inviting them to keep watching the video to find the information. Grounded on Schnotz (2005), and Ayres and Paas (2007), the contextualization made by segments 1 and 2 can aid learners in activating their prior knowledge on the topic foods and celebrations (whether they already worked with it in the textbook) in order to be prepared for the information and language they may be provided with in the video.

Segment 8, the last one, seems to have the function of ending the video. When the host produces the utterance "wishing you and yours a happy thanksgiving!", it is shown that the video has come to its end, especially, because, as the main topic is thanksgiving, he wishes a happy thanksgiving to the audience, ultimately, to learners.

As regards visual and aural effects in the segmentation design, it is possible to see that imagery information is not used to introduce or end the segments, and their function is to signal

the information presented. Concerning the visual design, the video can be considered simplistic, once the only observable effects are the fading in and fading out of the transitional effects in the first and the last segments. The transitional effects applied to the segments 2-8 throughout the video are designed in a clear-cut effect, that is, when each segment ends, the screen is cut so the next one can start.

Coherence

Concerning the presentation of information aiding learners in the process of understanding and knowledge construction in the Interchange 2 Unit 8 video design, some aspects can be observed. Concerning irrelevant written language, as aforementioned, there is no insertion of written language in the video design, and the oral language is the main means through which information is presented to learners, especially, due to the fact that the video depicts conversational situations in which a TV host is interviewing people.

The oral language design can be considered largely relevant. Some pieces of oral information, however, can be questioned. The oral language used when the TV host asks people about the amount of leftovers can be considered irrelevant for not being in consonance with the topic "how do people celebrate?" and "what do they eat?". The insertion of such irrelevant oral language design may result in learners engaging in extraneous processing, once they would use precious WM resources for trying to understand it, as extensively discussed by Ayers and Paas (2007), Kalyuga (2010), Mayer (2009), Mayer and Moreno (2010), and Sweller (2010). Consequently, the irrelevant oral information could be considered not to contribute for achieving the main objective of the video, which is to build coherent mental representations of the foods by understanding the target language structure and vocabulary presented.

Regarding the use of visual and aural aids, the video does not present visual and transitional effects in the segments, as aforementioned. The imagery information design can be considered relevant in order to present learners with a visual aid for the understanding of the target vocabulary presented (ZAREI; KHAZAIE, 2011), thus, facilitating the process of building coherent mental representations of the information as a whole (MAYER, 2005a; 2009), that is, the foods eaten at specific celebrations. However, one feature that can be considered irrelevant is the insertion of a smooth instrumental music as soundtrack throughout the video. The issue that arises, however, is whether the insertion of soundtrack may hurt learning due to the fact that it may compete with the oral language input for learners' WM resources (KALYUGA, 2010; MAYER, 2009; SWELLER, 2010).

4.1.2.5 Video from the Textbook Interchange 3 – Unit 1

The video from Interchange 3 Unit 1 has the length of six minutes and five seconds. Its design depicts a TV date show, in which one woman is expected to choose one out of three bachelors to have a date based on their personality characteristics and preferences, as can be seen in the analysis.

Signaling

Concerning the insertion of typographical cues in the video design to work as a signaling device, it is possible to say that the written language presented in the video design is of a minimum amount, at the beginning and the end of the video. Once the use of subtitles to work as a signaling device that might aid learners in the process of understanding is absent from the video design, the only observable written language input that may considered to aid learners in understanding is the use of a title on its design. The written language, however, does not work as a title for the video itself. The use of the name Dream Date in a dynamic stylish fashion shows that it is the name of the TV show learners are watching, as in Figure 26.

Figure 26 – Typographical Cues as Title in Interchange 3 Unit 1 Video Design

Source: Interchange 5th Edition website (2021)

Moreover, another observable factor that supports the fact that the written language works as the name of the TV show is that, after the opening of the show, the name is also shown on the back where the three bachelors are placed, as in Figure 27.

Figure 27 – Typographical Cues in Interchange 3 Unit 1 Video Design

Source: Interchange 5th Edition website (2021)

Grounded on Plass and Jones (2005), it is possible to say that the presentation of the typographical cues in the form of title, that is, the name of the show and, later, as the background of the studio set, may function to aid learners in the apperception process, establishing a link between its name and the structure of the show, suggesting that one of the three men sitting with their back to the screen may be a dream date. In addition, based on Schnotz (2005), and Ayres and Paas (2007), another function is to aid learners in the activation of their prior knowledge. Once the video accompanies unit 1 of the textbook, whose topics are concerned with people's personality traits and describing ideal mates, when learners are provided with the name of the TV show, they can draw connections between the name Dream Date and what would be expected for a dream date, concerning the characteristics an ideal date should have.

Concerning the typographical choices for the presentation of the TV show's name, it is designed in a stylish font in the color red positioned within two circles, one purple and one light yellow, granting the name a saliency effect due to the color contrast. The saliency feature can also be attributed to the TV show name, possibly having the function of aiding learners in apperceiving the topic and associating it to the textbook. The title Dream Date is positioned in the center of the screen, possibly to aid learners in directing their attentional resources to the name (SCHMIDT, 1990; LORCH, 1989; KONING *et al.*, 2010; HASLER; KERSTEN; SWELLER, 2007). By directing learners' attention to the necessary information presented on the screen, extraneous processing can be avoided, as defended by Mayer (2009), Mayer and Moreno (2010), and Sweller (2010).

The font (style) and the color (red) used in the title can be associated to the main topic the show is about, which is dating, also being associated to love and romance as well as comprehend interpersonal meaning, suggesting the kinds of relationships individuals to whom they are associated can have. As is referred by Kress and Van Leeuwen (2002) as "colour coordination" (p. 349), textual (color) coherence is associated to the matching of the predominant color. Such visual design, as further explained by the authors, as can be coupled with other kinds of sources, such as typography, that is, written language, or visual elements presented. In the case of the video, the color red matches the curtains and the flowers in the studio and in the title presentation.

Regarding the use of oral language to work as signaling devices in the video design, some features can be observed to work as devices to signal, that is, according to Hasler, Kersten and Sweller (2007), to direct learners' attentional resources to the information being presented.

The first feature concerns the oral language design used to introduce and end the video, as observed in the examples provided.

> "Hello, everyone and welcome to Dream Date! The show where one lucky lady gets to choose her perfect date. I'm your host Richard Darian and now… let's meet our contestants." (0'26)
> "Well, that takes the mystery out of the show! I'm your host Richard Darian, thanks for joining and see you next time on Dream Date!" (5'52)

The first example has the function of opening the video by welcoming learners to the show and providing further information to contextualize what it is about. When the host produces "The show where one lucky lady gets to choose her perfect date", it is possible to say that he is providing extra cues for learners to draw connections between the idea that a female participant can choose a male participant for a date, and the name of the show Dream Date. Moreover, after he introduces himself, the utterance "[…]let's meet our contestants" also provide cues for learners to understand that more than one person is taking part on the show, and that, the word contestants may suggest that the show is some type of competition. Such signaled cues can have the function of situating learners in the context of the video, and aid them in activating their prior knowledge.

The second example provided has the function of ending the video. By producing the utterance "I'm your host Richard Darian, thanks for joining and see you next time on Dream Date!" it is possible to say that the host is signaling the end of the TV show. Such oral design choice can be due to grant the video an oral language structure resembling a TV show learners may encounter in the real world. In addition, when the host thanks the audience for watching the show and says a farewell "see you next time" he is inviting them to watch the next episode.

The utterances provide learners with cues signaling that the video has come to its end and no further information will be presented.

Another oral language signaling device refers to emphasizing certain words, when, for example, the characters talk about preferences concerning people's personality traits with the use of emphasis possibly, as argued by Plass and Jones (2005), aiding learners in the process of apperceiving specific language items, as in the examples below.

> "I like a guy who is… honest and straightforward. Someone who is… considerate, and interested in other people." (0'58)
> "Is there anything you don't like in a guy?" (1'01)
> "Yes. I don't like when a guy is egotistical. When he talks about himself all the time. That is something that really bothers me." (1'13)
> "It really bothers me when… people argue over unimportant things. I just think people should be less temperamental and more relaxed." (5'11)

The underlined words represent the words produced with a stronger emphasis. The choice for such oral language design may be associated to the main topic of the textbook, in which learners work and discuss about people's personality traits. In order to be congruent with the topic, the vocabulary items about personality are signaled. Based on Koning *et al.* (2010), and Hasler, Kersten and Sweller (2007), this oral design is aimed at directing learners' attentional resources to them and build a coherent mental representations (MAYER, 2005a; 2009) of the situation, that is, to understand not only the vocabulary items *per se*, but about attributing them to the context where the woman is describing what she considers good and bad traits, according to her preferences. The last example provides information about what is appreciated and not. When the Bachelor Contestant 2 explains what bothers him, he emphasizes the utterances argue, unimportant things, temperamental and relaxed, making them salient so as to aid learners in apperceiving (PLASS; JONES, 2005) his contrasting ideas about how people should react or not react. The same can be observed in the following examples, when the contestants describe positive and negating characteristics about themselves.

> "Well I think I'm a pretty good friend, very reliable. Something negative… I'm very direct. And… that sometimes upsets people." (3'42)
> "Well, I'm actually a pretty good guy. As for something not so good, I, aahh… I can't think of anything!" (4'03)

As the contestants are asked to say what they consider good and bad qualities about themselves, emphasis is given to the personality characteristics vocabulary, possibly to make them salient to learners. Furthermore, the utterances provide examples of the language structure used when one is expected to describe personality traits, either positive or negative.

Intonation changes can also be observed in the oral language design when the host introduces the three male contestants, as in providing personal information about them.

> "Bachelor number one is a writer and a former college soccer star who loves watching almost any kind of sports. Bachelor number two is a director of educational programs at a science museum and in his free time he enjoys reading, running and camping. Bachelor number three is a model who also studies marketing." (1'41)

It is possible to observe, in the utterances provided by the host, two types of information. The first type is concerned with occupations, for example writer, soccer star, educational programs and marketing, made salient to direct learners' attention to them. The second type is concerned with the three contestants' leisure activities, in which the words sports, reading, running and camping are also orally emphasized as well as preceded by the utterances "loves" and "in his free time he enjoys", based on Plass and Jones (2005), aiding learners in the process of apperceiving the information made salient and building coherent mental representations of what the three contestants do for work and for leisure activities (MAYER, 2005a; 2009).

Finally, once the video depicts a TV show, in which a conversational situation takes place, a conversational strategy of repetition is also observable.

> "Elisabeth, meet Adam York." (5'31)
> "Adam York?" (5'33)
> "Betty? Betty Campbell?" (5'36)

Repetition used as a conversational strategy may be part of the oral language design. Such design may have been chosen in order to resemble a real life situation in which people can make use of repetition for information checking and signaling that they already know each other.

In the case of using visual aids to signal information, it is possible to observe that the only imagery devices used are the TV show's logo, which presents its name with a number of flowers dynamically moving from the right and left sides towards the center of the screen, that is, in direction of the show's name. The choice for such imagery elements may be due to the topic of the show, which is concerned with getting two people to have a date, romance and love.

Segmenting

The Interchange 3 Unit 1 video is designed in a manner that, despite providing learners with information in a transient fashion, as defined by Leahy and Sweller (2016), presents the information load in a segmentation structure (SWELLER, 2010; IBRAHIM, 2012; IBRAHIM *et al.*, 2012), organized in eight segments, as presented in Chart 15.

Chart 15 – Segmentation Structure in Interchange 3 Unit 1 Video

Segments	Information Presented
Segment 1	Dream Date
Segment 2	**Host:** Hello, everyone and welcome to Dream Date! The show where one lucky lady gets to choose her perfect date. I'm your host Richard Darian and now… let's meet our contestants.
Segment 3	**Host:** She an English teacher from Chicago, Illinoi. She's twenty-five years old, interested in American history, literature and the arts. Ladies and gentlemen meet Elisabeth Campbell. Hello Elisabeth. And welcome to Dream Date. **Elisabeth:** Hi, Richard. I'm delighted to be here! **Host:** So, tell us, Elisabeth, what qualities you look for in a guy? **Elisabeth:** I like a guy who is… honest and straightforward. Someone who is… considerate, and interested in other people. **Host:** Is there anything you don't like in a guy? **Elisabeth:** Yes. I don't like when a guy is egotistical. When he talks about himself all the time. That is something that really bothers me. **Host:** (**Laughing**) I know what you mean. Well, it's time to meet our bachelors.
Segment 4	**Host:** Bachelor number one is a writer and a former college soccer star who loves watching almost any kind of sports. Bachelor number two is a director of educational programs at a science museum and in his free time he enjoys reading, running and camping. Bachelor number three is a model who also studies marketing. He says everyone should recognize him from his latest running shoe add on billboards and magazines. Welcome gentlemen! It's good to have you with us! Ok, let's get started.
Segment 5	**Host:** What is your first question? **Elisabeth:** Bachelor number one. If you're with your friends, and we're late to our first date, what excuse would you give? **Bachelor Contestant 1:** Well… I'd be too embarrassed to tell you the truth. So, I'd probably say I had car trouble, or got stuck in traffic, or something. **Elisabeth:** Ok. Bachelor number two. The same question. **Bachelor Contestant 2:** Well, it bothers me when people lie. So, I would tell you the truth and hope you were generous enough to forgive me. **Elisabeth:** Bachelor number three. **Bachelor Contestant 3:** Well, I'd probably tell you what I tell everyone else. It's hard to split my time among so many people. **Elisabeth:**

Segments	Information Presented
	Ok. Next question.
Segment 6	**Elisabeth:** Bachelor number one, describe to me your ideal date. **Bachelor Contestant 1:** Getting together with the bunch of my friends, having a barbecue, and watching sports on TV. Yeah, that'd be a great date! **Elisabeth:** Ok, Bachelor number 2? **Bachelor Contestant 2:** I'd take you out for a nice dinner. Let's just relax and enjoy the evening and get to know each other. **Elisabeth:** Bachelor number 2, tell two things about yourself. One positive and one negative. **Bachelor Contestant 2:** Well I think I'm a pretty good friend, very reliable. Something negative… I'm very direct. And… that sometimes upsets people. **Elisabeth:** No, I think that's a good thing! Bachelor number 3. Something good and something not so good about yourself. **Bachelor Contestant 3:** Well, I'm actually a pretty good guy. As for something not so good, I, aahh… I can't think of anything! **Host:** (Laughs) ok. Elisabeth, final question and then you're gonna have to make your choice.
Segment 7	**Elisabeth:** Right. Bachelor number one, finish this sentence: "I can't stand it when…" **Bachelor Contestant 1:** I can't stand it when… people talk when I'm trying to watch the soccer match on TV. **Elisabeth:** Bachelor number three, finish this sentence: "I think it's disgusting when…" **Bachelor Contestant 3:** I think it's disgusting when… when I go to a fancy restaurant and I don't get the service I deserve. **Elisabeth:** Bachelor number two. Finish this sentence: "it really bothers me when…" **Bachelor Contestant 2:** It really bothers me when… people argue over unimportant things. I just think people should be less temperamental and more relaxed. **Host:** We are out of time.
Segment 8	**Host:** Now, I'm going to ask you to make that choice. **Elisabeth:** I think I'm going to choose bachelor number two. **Host:** Are you sure? **Elisabeth:** Yes!

Segments	Information Presented
	Host: All right! Well, let's meet this guy, bachelor number two comes from Chicago, your own home town. Elisabeth, meet Adam York. **Elisabeth:** Adam York? **Bachelor Contestant 2:** Betty? Betty Campbell? **Elisabeth:** (Laughing) Adam! **Host:** Wait a minute, you two know each other? **Elisabeth:** We went to high school together! (**Laughs**) **Host:** Well, that takes the mystery out of the show! I'm your host Richard Darian, thanks for joining and see you next time on Dream Date!

Source: The researcher (2021)

As has already been discussed, the aim of the segmentation structure is to provide learners with information in a structured fashion to control the amount of information presented each time and prevent them from engaging in overload processing as argued by Ayers and Paas (2007), Mayer (2009), Mayer and Moreno (2010), and Sweller (2010). In the case of the Interchange 3 Unit 1 video, the segmentation structure seems to have the function of providing learners with information of topic of the video and of the target language used when providing personal information.

In the case of written language used, only one segment is observable to be started with such design with reduced amount of information, (segment 1), in which the TV show's name is presented. The name of the TV show also works as introducing the segments to come, once it can also be examined work as a contextualization cue to aid learners in understanding the context of the information presented in the subsequent segments. In the case of written language used to end segments, only one example can be detected (segment 8), when the credits of the TV show roll up the screen. The presentation of the credits can be considered commonplace and only have a contextual function to the type of video. As the video depicts a TV show, as is regularly expected, when the show comes to its end, the credits are presented. Such design may have been chosen in order to resemble a real life TV show that learners might see on television.

The oral language segmentation design of the video, in turn, may have the function of providing learners with examples of the target oral language structures used to provide different kinds of personal information about the TV show participants, as in the examples below.

(Segment 3)

[…]

> **Host:**
> She is an English teacher from Chicago, Illinoi. She's twenty-five years old, interested in American history, literature and the arts. […]

(Segment 4)

> **Host:**
> Bachelor number one is a writer and a former college soccer star who loves watching almost any kind of sports. Bachelor number two is a director of educational programs at a science museum and in his free time he enjoys reading, running and camping. Bachelor number three is a model who also studies marketing. He says everyone should recognize him from his latest running shoe add on billboards and magazines. […]

(Segment 8)

> **Host:**
> All right! Well, let's meet this guy, bachelor number two comes from Chicago, your own home town. Elisabeth meet Adam York.
> **Elisabeth:**
> Adam York?
> **Bachelor Contestant 2:**
> Betty? Betty Campbell?

It is possible to argue that the segments presented aim at providing learners with the target language structure used to talk about people's personal information. Segment 3 presents examples of the target language used to provide personal information such as one's name, age, place of origin and what they are interested in. Providing name and place of origin can also be observed in segment 8, when the female participant finally chooses one of the three Bachelor Contestants and they meet. Segment 4 aims at providing learners with examples of the target language used to talk about people's occupations and leisure activities when the utterances writer, former college soccer star and loves watching almost any kind of sports are presented about Bachelor Contestant 1, and "director of educational programs at a science museum and he enjoys reading, running and camping" about Bachelor Contestant 2.

Concerning Bachelor Contestant 3, only information concerning his occupation is provided when the utterance "model who also studies marketing" is presented. It is possible to say that, since the TV show host provides information about the three Bachelor Contestants in segment 4, the information is of a reduced amount, probably, to prevent them from engaging in a cognitive overload due to excessing their working memory resources, resulting in extraneous processing (KALYUGA, 2010; MAYER, 2009; SWELLER, 2017).

Another type of personal information that the segments aim at presenting refers to personality traits, as in the following examples.

(Segment 3)

> […]
> **Elisabeth:**

> I like a guy who is… honest and straightforward. Someone who is… considerate, and interested in other people.
> […]
> **Elisabeth:**
> Yes. I don't like when a guy is egotistical. When he talks about himself all the time. That is something that really bothers me.

(Segment 6)

> […]
> **Bachelor Contestant 2:**
> Well I think I'm a pretty good friend, very reliable. Something negative… I'm very direct. And… that sometimes upsets people.
> […]
> **Bachelor Contestant 3:**
> Well, I'm actually a pretty good guy. As for something not so good, I, aahh… I can't think of anything!

Segments 3 and 6 provide examples of the target language used when describing the personality traits appreciated in people when the utterance "I like […]" is produced, followed by the personality vocabulary honest, straightforward, considerate and interested in other people, as well as the traits that are not appreciated with the utterance "I don't like […]" followed by the vocabulary egotistical and the utterance "when he talks about himself all the time". Segment 6 further provides examples of the target language structure used when one describes what they consider to be positive about themselves with the utterances "I'm a pretty good friend, very reliable" and "I'm actually a pretty good guy" and what they consider negative with the utterances "Something negative… I'm very direct." and "As for something not so good, I, aahh… I can't think of anything!". Although the last utterance does not effectively present a personality trait, it can be examined to aim at assisting learners in inferring and attributing personality traits to Bachelor Contestant 3 (*e.g.* self-centered). In the case of the two segments, the amount of personality traits presented is limited, possibly, in order not to overload learners' working memory (WM) resources by exceeding their processing resources, resulting in extraneous processing (AYERS; PAAS, 2007; KALYUGA, 2010; MAYER, 2009; MAYER; MORENO, 2010; SWELLER, 2010).

Another feature in the segmentation structure is that the oral language is used to introduce the segment to come, as the examples below.

(Segment 2)

> **Host:**
> […] and now… let's meet our contestants.

(Segment 3)

> **Host:**
> […] Well, it's time to meet our bachelors.

(Segment 6)

Host:
Elisabeth, final question and then you're gonna have to make your choice.

Segments 2 and 3 have the function of introducing the segments that follow them. In segment 2, after the host introduces the TV show by providing its name and what it is about, the utterance "and now… let's meet our contestants" may have the function of inviting learners to continue watching in order to meet the participants. The same function may be seen in segment 3, with the utterance "[…] it's time to meet our bachelors" in which learners are also invited to continue watching to know who the three male participants of the show are.

The oral language design used to introduce the next segment seems to have the function of aiding learners in preparing for the information that is to be presented. Furthermore, grounded on Schnotz (2005), and Ayres and Paas (2007), the oral language used in the design to introduce the next segments may also have the function of activating learners' prior knowledge concerning the topic of the video, that is, people's personal information to be provided as they are introduced.

Finally, segment 6 can be examined to establish that the TV show is moving towards the end when the utterance "[…] final question and then you're gonna have to make your choice" is produced, clarifying that the female participant is able to make only one more question.

A final feature in the segmentation design is the use of the oral language to end the segments, as follows.

(Segment 7)

[…]
Host:
We are out of time.

(Segment 8)

[…]
Host:
Well, that takes the mystery out of the show! I'm your host Richard Darian, thanks for joining and see you next time on Dream Date!

The utterance in segment 7 seems to have the function of ending the segment in which the three Bachelor Contestants were asked to finish sentences started by the female participant. In addition, by observing the TV show as a whole, the utterance "We are out of time" also provides learners with evidence there is no time for extra questions. The utterance in segment 8 "I'm your host Richard Darian, thanks for joining and see you next time on Dream Date!" may have the function of clarifying to learners that the TV show has come to its end and no

further information will be provided. Such oral language design seems to grant the video a sense of structure and, in the case of resembling a real life TV show, that it is finishing and that the audience is invited to watch the next episode.

Concerning the use of visual effects in the segmentation design, one is used when the name of the TV is presented in order to establish a link between segments 1 and 2. In segment 1, the title Dream Date is presented in a dynamic way, enlarged on the center of the screen. As the written utterance enlarges, the background, which is initially of a light yellow color, changes for the presentation of the TV set. The name of the show, then, fades out, showing the same show's logo portrayed on the back wall of the TV set. Apart from the visual effects applied to the show's name, the transition between each segment can be observed to be made by clear-cut effects.

Coherence

Concerning the information that is part of the Interchange 3 Unit 8 video design that can be considered irrelevant due to the fact it does not aid the process of learning, some aspects can be observed. As afore discussed, the only pieces of written information that are presented in the video are its title, that is, the name of the TV show and the credits that roll up the screen in the end of the video. The TV show can be considered relevant in the process of aiding learners for the fact that it is associated to the main topic of the video, and aiding the process of prior knowledge activation (SCHNOTZ, 2005; AYRES; PAAS, 2007; KALYUGA, 2010). The written information on the credits, on the other hand, does not seem to assist learners' understanding, thus, considered irrelevant in the video design. Although the written information in the form of the show credits does not aid learning, it can be assumed as being a piece of information that is part of the type of the video learners are provided with, that is, a TV show.

Regarding the oral language, part of the video design, it is possible to say that the video does not present irrelevant oral language to any extent. As the video depicts a TV show in which the participants resemble conversational interactions, all the oral utterances appear to be relevant for learners to be able to picture the information by building coherent mental representations (MAYER, 2005a; 2009) of it.

When the name of the TV show is presented, it is accompanied by lively music as soundtrack, probably, to afford it a dynamic sense and also to resemble a real life show. Such music soundtrack is also presented as the host ends the show and the credits roll up the screen. The insertion of the music soundtrack may have an aesthetic function, which is to grant the TV show an enjoyable character. Considering the cognitive effects of the music soundtrack, it is presented only when no relevant information is being presented, more specifically, in segment

1 with the name of the show, and in segment 8, when the credits roll up, possibly not hurting learning due to overloading learners' WM resources (KALYUGA, 2010; MAYER, 2009; SWELLER, 2010).

In addition to the aural information, throughout the video the sounds of applause are inserted in its design, probably, to suggest learners that this is a show to a live audience. However, the sounds seem to be assisting the composition of the context of the video, which is a TV show in which the audience may manifest at different moments. Although grounded on the coherence principle the sounds of applause do not influence learners' understanding of the information, thus be considered an irrelevant piece of information due to the fact that they do not aid the learning of the target information, these sounds may work as contextualization clues aiding learners in understanding that the type of show watched is an auditorium show.

The next section presents a discussion of the analyses presented concering the three research questions that guided this research study followed by a brief description of the nine principles proposed by Mayer (2009) for the design of educational multimedia materials aside from the three ones that were part of the video analysis framework that could be identified in the videos.

4.2 DISCUSSION

This section presents a discussion of the analyses of the ten videos, establishing an association between the features that were examined in their design to the three research questions that guided this Doctoral Dissertation. The discussions are organized in four subsections: 4.2.1 Discussing Signaling, 4.2.2 Discussing Segmenting, 4.2.3 Discussing Coherence, and 4.2.4 A Few Words on Other Principles. The analyses show evidence that the videos were mostly designed with features that aim at facilitating the process of L2 learning. Grounded on the three major principles, Signaling, Segmenting and Coherence, which supported the development of the video analysis framework, it is possible to argue that the analyses of the ten videos provided important information about how videos can be designed based on the knowledge of how the human mind processes information, ultimately to aid the process of learning an L2.

4.2.1 Discussing Signaling

The research question 1 (Do the educational videos incorporate the principle of signaling, with features enhancing the L2 input? If so, how?) refers to the signaling principle, which, as is extensively discussed by Mayer and Moreno (2010) and Sweller (2010), is aimed at preventing learners from engaging in extraneous processing, specifically due to directing their attentional resources (SCHMIDT, 1990; LORCH, 1989; KONING *et al.*, 2010; HASLER; KERSTEN; SWELLER, 2007), to specific pieces of information shown. As is argued by Schmidt (1990), Koning *et al.* (2010) and Hasler, Kersten and Sweller (2007) directing learners' attention is crucial for effective learning to take place. In the case of the videos analyzed, a range of different features for directing learners' attentional resources to relevant information was detected in their design. Such features were: a) the use of written language, presented as titles of the videos, subtitles or as on-screen text; b) oral language designed to help learners not lose track of the information; and c) the insertion of imagery information in the video design so that the information can be fully integrated.

According to Schnotz (2005) and Ayres and Paas (2007) the use of written language input can have the function of activating learners' prior knowledge, which, in the case of the title of the videos, can activate the knowledge they have about the topic of the videos. As afore discussed, when learners' prior knowledge is activated, the information concerning the topics they already have stored in long-term memory as superordinate and subordinate structures are retrieved back to working memory and processed with new incoming information so to turn it into new strings of knowledge. The titles can, then, aid learners in preparing for the information that is to be presented in the video, especially, when the video is used as a means for supporting the target language already worked in class. The insertion of titles to the videos design, however, were restricted to the PNLD videos, that is, to 50% of the total amount of videos analyzed – 5 videos. In the case of subtitles, similarly to the case of titles, they were restrained to 50% of the total amount of videos analyzed (the PNLD videos – 5 videos). The use of subtitles can be observed to have the function of signaling the information that is being presented as a means to aid learners in possibly clarifying any pieces of information that might not become clear only by listening to the oral language. The function of the subtitles, then, can be as integrating the written language to the oral language being provided by the narrators/characters, possibly to free working memory (WM) resources from applying excessive effort in trying to understand the information being watched. According to Kalyuga (2010), Mayer (2009), and Sweller (2010), when WM resources do not exceed the processing

capacity, learning is benefitted. It is important to mention that although all the 5 PNLD videos present subtitles in their design, in the guidelines for the textbooks to participate in the program, there is no specific information establishing subtitles as a mandatory design feature of the audiovisual materials accompanying the textbooks.

The presentation of written language as on-screen target text can also be examined as having a signaling function, especially in the videos presenting grammatical explanations, as is the case of the PNLD video from the textbook Time to Share, that concerned entirely with providing learners with explanations about talking about future plans. In addition, on-screen target text was also examined in the video from the textbook It Fits, which presented learners with the target language used to describe activities people are able to do. As can be observed in the case of the two videos, presenting specific pieces of information as on-screen target text with examples of the target language can have the function of directing learners' attentional resources to them, as the narrator provides explanations about their use.

Still concerning the signaling feature, the oral language design may have the essential importance in directing learners' attention. Firstly, for the videos designed for learning an L2, presenting the target oral language in their design seems to be necessary. Furthermore, features as the use of different voice intonations may work as a hook that in a way can make some specific pieces of information salient for learners. The saliency issue can be seen more clearly in the videos with grammatical explanations in which the narrator explicitly called their attention by producing utterances such as "observe that […]" to clearly direct the attention to the elements being explained. By directing learners' attentional resources (SCHMIDT, 1990; LORCH, 1989; KONING *et al.*, 2010; HASLER; KERSTEN; SWELLER, 2007) to specific pieces of information, the saliency feature working as a signaling device in the videos design can be argued to have the function of preventing learners from directing their attention to unnecessary pieces of information, resulting in engaging extraneous processing, which, as argued by Ayers and Paas (2007), Mayer (2009), Mayer and Moreno (2010), and Sweller (2010) can hurt learning.

This is not case for the Interchange videos. Once they are not designed aiming at explicitly teaching grammatical aspects (thus, not presenting a narrator providing explanations), no oral language directing learners' attentional resources (SCHMIDT, 1990; LORCH, 1989; KONING *et al.*, 2010; HASLER; KERSTEN; SWELLER, 2007) to the information is part of the design. The only oral language design aiming at directing their attention (not explicitly, however) to the information is the use of different intonations in order to make pieces of information more salient than others throughout the interactions depicted. One issue that arises

is concerning whether the amount and flow of information may be in consonance with learners' language proficiency, and not exceeding their WM processing capacity. According to Kalyuga (2010) and Mayer (2009), in the case learners are provided with an amount of information (and language) that may be beyond their WM processing capacity, the result might be contrary to the expected, that is, learning would be hurt, failing to achieve the objectives of the videos, which is to aid learners in the process of developing the L2.

Finally, the use of imagery information was also observed as having signaling function, once, as discussed in chapter 2, the presentation of imagery information can have a crucial function at the service of learning. Once videos present information to learners by means of images and words, the imagery information is integrated to the oral information in learners' WM, aiding them in the process of building coherent mental representations of the information they are exposed to. As represented in the diagram in Figure 3 in chapter 2 based on the Cognitive Theory of Multimedia Learning extensively discussed by Mayer (2005a, 2009), when learners are exposed to any educational material that makes use of images, the processing of the information can be facilitated. When the visual information enters learner's cognitive system through their SM (the eyes), their WM is responsible for building mental representations of the image as well as converting it into sounds. In addition, learners' prior knowledge is also activated and retrieved from their long-term memory back to the working memory for it to be re-processed with the new incoming information.

Based on the analyses of the ten videos, the function of the imagery information can be to facilitate learners' comprehension of the target language by working as contextualization cues, especially due to the fact that they present visual aids for the integration of the verbal and imagery information.

4.2.2 Discussing Segmenting

The research question 2 (Do the educational videos incorporate the principle of segmenting facilitating the comprehension of the target language? If so, how?) refers to the segmenting principle as presented and discussed based on the triarchic model of cognitive load brought by Mayer and Moreno (2010) (Figure 6) and in Chart 1, both in chapter 2. The objective of the segmenting principle is to manage the essential processing of the information learners are exposed to, in the case, to the videos. In the case of the videos aimed at the learning of an L2, the essential processing refers to the processing of the target language necessary for the development of grammatical and lexical competencies.

Based on the definition provided, designing a video in segments, that is, in blocks, can be argued to assist learners in the process they engage in in making sense of the information being processed precisely by presenting it in small portions. Depending on the amount of information provided in the video, or the complexity of it (the level of element interactivity, in which the higher the level is, the more intricate the information is), the more complex it may become for learners based on their level of proficiency. Segmenting complex information can, thus, benefit learning. As explained by Mayer (2005a, 2009), when information with high level of complexity is broken down to smaller pieces, learners are afforded with time to be able to digest the information and organize the pieces of information in coherent mental representations. Once videos provide information and the target language in a continuous flow, their design needs to (or at best should) aid learners in processing the information being presented.

The way in which the segmentation structure is designed in the videos analyzed seems to be conceived by means of visual effects, that is, the transition between the segments when one segment is cut to the next one, or the use of imagery information. Another segmentation feature is the use of oral language in ending one segment and introducing the next one, especially in the videos presenting grammatical explanations, in which the narrator explicitly introduced the segments to come with examples.

In the case of the PNLD videos, such design could be observed in the video *Usando Going To e Will Para Falar de Futuro* from the Time to Share textbook. It can benefit learners in the process of making sense of the information due to the fact that, when the information is presented in a step-by-step fashion, learning can be benefitted (especially) due to the reduced amounts of information presented at each time, and the (probably) impression of unfolding, in which one information leads to the next.

According to Kalyuga (2010), Mayer (2009) and Sweller (2010), there is the necessity for educational materials to be organized in smaller portions so that learners can be able to clearly make sense of the information received and not exceed their working memory resources. Learners can be benefitted by the segmentation structure in videos that provide grammatical explanations due to the complexity the information being presented may have. Moreover, the more complex the information, the more segmented the videos (ideally) need to be. One example is the case of explaining the use of ‘going to’ to refer to future. While watching the video, learners are expected to understand that, in order to indicate the future, they need to insert “going to” followed by an action verb in the sentence, and (re)activate their knowledge of how the verb “be” works, that is, that it has three forms (am, is and are), depending on the person

being referred to. Depending on learners' language proficiency, the information presented in the video can be complex due to the load of element interactivity, that is, the level of complexity, based on learner's prior knowledge. Because such complexity cannot be reduced or eliminated due to the fact of being inherent to the grammatical topic (it is not possible to learn the use of future with "going to" without the use of the verb "be"), it can be managed, precisely by being organized in blocks to provide learners with step-by-step-like explanations so that their WM resources may not be exceeded, which, according to Kalyuga (2010), Mayer (2009) and Sweller (2010) may result in poor or no learning.

Providing learners with examples of the target language to be used in easily-observable (by part of learners) chunks was also examined in the video Things They Can Or Can't Do! from the PNLD textbook It Fits, in which each segment presents the target information associated to the use of the modals "can/can't" so that learners can be aided in making sense of the information and of the structure of the sentences. Although the video aims at providing learners with grammatical explanations, there is no narrator (or teacher) explaining the structure to be used. The target language with the modals can/can't is presented by a narration in the form of questions and answers. The video Things They Can Or Can't Do! can be considered a clear example of video organized in segments especially because learners may not have difficulties in noticing that each segment is aimed at presenting one example of the target language structure being studied.

The PNLD videos, however, are not restricted to the provision of grammatical explanations. The video entitled Introducing Yourself that accompanies the textbook It Fits is designed in a fashion that presents a conversational situation depicting a computer-based video call in which two children are talking about their families and schools, thus, being its focus on content (personal information). Similarly, the videos How PSAs Are Made and The Life of Nelson Mandela do not have the focus on grammatical structure. The aim of the two videos from the Become textbook is to provide learners with content information about what public service announcements are and how to develop one, and information about the history of Nelson Mandela, respectively.

In the case of the videos that accompany the Interchange textbooks, on the other hand, the videos have a different segmentation structure, which can be considered not to be so evident, or eye-popping. Although the videos have been examined as structured in segments, depending on learners' language proficiency, these videos could fail in aiding them in engaging in the generative processing, that is, according to Mayer (2009), the processing that may result in actual learning due to the transient nature of the videos and the high amount of information

provided at each time. It is important to acknowledge that videos that follow such design tend to present information in a less visually segmented fashion (differently from the grammatical videos), probably, in order to grant a considerably natural-like information flow (not appearing to be scripted), and perhaps a more film-like characteristic.

The analyses of the 10 videos provided examples of videos designed in a segmented fashion. According to Mayer and Moreno (2010) and Sweller (2010), depending on the intrinsic load of the information in the video, that is, on the level of element interactivity, the more segmented the fashion in which the information is presented, the more learners may succeed in understanding the information with the limits of their WM capacity. As put by the authors, when working memory resources are freed, they allow room for the essential processing of the information. An ultimate result may be that learners will succeed in engaging in the processing of the information watched, which, according to Mayer (2009), and Mayer and Moreno (2010) is the aim of the segmentation design.

4.2.3 Discussing Coherence

The third principle from the video analysis framework is associated to research question 3 (Do the educational videos incorporate the principle of coherence, maintaining relevant information to assist L2 learners? If so, how?). As discussed in chapter 2, (subsection 2.2), grounded on arguments brought by Mayer (2005a, 2005b, 2009), the major premise of the coherence principle is that any information presented in the video that can be considered irrelevant, or not crucial for the understanding (learning) of the information presented on it, should be removed in order to prevent learners from engaging in extraneous processing.

As regards the videos form the PNLD, every piece of information presented can be considered relevant for the understanding of the content. The videos were examined as being designed aiming at facilitating learners' understanding and learning of the target language by means of the insertion of imagery information in their design, verbal language in the written and oral forms for they presented subtitles and visual and aural aids such as noises and the transitional effects. Such features were examined as aiding the learning not only of the language structure, as is the case of the video from the Time to Share textbook (*Usando Going To e Will para Falar de Futuro*), but also the content as a whole as is the case of the videos from the Become textbooks (How Psas Are Made and The Life Of Nelson Mandela) and examples of the target language learners would (probably) be expected to use in later moments, as is the

case of the videos from the It Fits textbooks (Introducing Yourself and Things they can or can't do!).

The insertion of music and sounds as soundtrack in the videos design, however, was examined as the feature that falls into the coherence principle, being considered irrelevant for not having direct influence in facilitating learning. Such feature was mostly detected in the videos that accompany the Interchange textbooks, as for example Interchange 1 Unit 8 video, in which the insertion of instrumental music in some moments of the video in order to create a link between segments, or to grant an emotional sense, such as surprise. In the Interchange 2 Unit 8 video, soundtrack is part of the video design and is presented throughout the whole video. Finally, the Interchange 3 Unit 1 video, for depicting a TV show, presents instrumental music as soundtrack for the opening of the show. The insertion of sounds of applause were also detected in the video design, probably to grant a more realistic impression to learners while they watch.

The soundtrack inserted in the videos design can play the role of aiding learning for the fact that they can provide contextualization cues. However, it can be considered somewhat dubious. Whereas on the one hand the soundtrack aids the process of contextualizing the conversational situations depicted in the videos, on the other hand, based on Mayer's (2009) coherence principle, and based on arguments by Kalyuga (2010) and Sweller (2010), it might be considered irrelevant and further harmful for learners due to occupying working memory resources for the processing and integration of the visual and aural information all together at the same moment. The issue that further comes with such concept is that, for the fact that as the soundtrack is an instrumental music presenting no spoken words and only sounds, it may not compete, or even compete in a reduced way, with the processing of the verbal oral information.

The notion that soundtrack should be removed for it does not present significant relevance for the understanding, and ultimately, learning of the target language, thus, poses a controversial matter. Perhaps the most solid claim that might incite a deeper reflection on such assertion is the fact that when watching videos of any kind (educational or not), one might expect to hear soundtrack in the form of noises, or music in order to aid the process of building the information and (perhaps) establishing emotional engagement by part of learners. Moreover, depending on the types of videos being watched, the lack of noises might result in an uncomfortable experience to learners, in which they may feel something is missing as is the case of the Dream Date video, which demands the sounds of an audience supposedly watching the show.

4.2.4 A Few Words on Other Principles

Once the video analysis framework, which grounded the analyses of the videos, was developed based on three major design principles Signaling, Segmenting and Coherence, this subsection is aimed at briefly signalizing which of the nine multimedia principles brought by Mayer (2009) can be seen in each of the videos analyzed.

Concerning the nine other multimedia principles presented by Mayer (2009), five principles could be observed in the design of the video Introducing Yourself from the It Fits 6th Grade textbook. The first, The Redundancy Principle, is part of the video design in that the verbal information is presented in a redundant way, that is, in the oral form in spoken language and in the written form with the presentation of subtitles. Although Mayer (2009) defends that presenting information in a redundant way might result in extraneous processing, overloading learners' working memory capacity, thus, hurting learning, it is important to consider that in the context of learning English as an L2, studies have shown that the use of subtitles can work as an important feature aiding comprehension and ultimately, learning the target language (LAVAUR; BAIRSTOW, 2011; DANAN, 2004; HAYATI; MOHMEDI, 2010). The second principle, The Temporal Contiguity Principle, can be observed in the video design in that the visual pieces of information (pictures) are accompanied by the oral information. This can be examined when characters start providing information about their friends, teachers and relatives, the pictures of those people are immediately shown on screen so to make it clear to learners that the people they are talking about are the ones presented. The third principle, The Multimedia Principle, is clearly observable once the video presents information by means of providing learners with the use of visual aids and verbal aids, instead of presenting only verbal aids to promote the learning of the target language. Finally, the fourth and fifth principles, The Personalization Principle and The Voice Principle in which the spoken language produced by the characters are presented in a conversational way and in a human voice, can be observed in the video design, once it resemble real-life informal interactions between two children where they are introducing themselves and the people from their personal circles.

In the video Things they can or can't do! from the It Fits 7th Grade textbook eight of Mayer's (2009) principles can be identified in its design. The first principle, The Redundancy Principle, can be observed once the information is presented in a redundant way, by means of the oral language by the narrator and the subtitles. Although Mayer (2009) suggests that providing learners with redundant information can cause extraneous processing, in the context of learning English as an L2, subtitles can aid the comprehension process, as already discussed.

The second principle, The Spatial Contiguity Principle, is part of the video design once target language questions and answers are positioned next to the characters they are referring to as well as the questions being positioned on the top of the screen and the answers below them, suggesting a hierarchical structure. The third principle, The Temporal Contiguity Principle, is also observed, once the imagery information is presented synchronously with the narration, allowing learners to possibly build clear connections between the oral and visual information. The fourth principle, The Modality Principle, is part of the video design once the video is organized in a structure providing learners with the use of imagery information accompanied by narration. The fifth principle, The Multimedia Principle, once the video is designed to present learners with images and words, in this case, the use of images and the oral verbal language and the written verbal language. The sixth principle, The Personalization Principle, can be observed in that the narration presented in the video with a rather informal conversational style. The seventh principle, The Voice Principle, is also part of the video design, once the narration is made by a human child voice. Finally, the eighth principle, The Image Principle, can also be observed, once the absence of the narrator's figure may not have a negative effect on learners' understanding of the information.

In the How PSAs are made video from the Become 7th Grade textbook, eight principles proposed by Mayer (2009) can be seen its design. The first principle, The Redundancy Principle, can be observed as the video presents the information that is narrated in the forms of subtitles and also as written text presented on screen. Although according to Mayer (2009) this principle may hurt learning, in the case of learning English as an L2, presenting language in a redundant design may aid learning. The second principle, The Spatial Contiguity Principle, can be seen once the written information and the images are presented on the center of the screen, aiming at directing learners attention to them in a straightforward way. The third principle, The Temporal Contiguity Principle, is also part of the video design, once the written text and the imagery information are shown on screen only when the narrator starts presenting information specifically about them. The fourth principle, The Modality Principle, can also be argued to be part of the video design due to the fact that the information is presented in the forms of images and narration. The fifth principle, The Multimedia Principle, can also be defended, once the video instead of presenting only words (in both written and oral forms), presents words and images in order to aid learners' comprehension. The sixth and the seventh principles detected in the video design, The Personalization Principle and The Voice Principle, are associated to presenting the information in a rather conversational form and in a human-like voice rather than a mechanical computer voice in order to have learners socially engaged to the video. Finally,

the eighth principle examined in the video design, The Image Principle, concerns that the fact the narrator's figure not being presented on screen may not have a negative impact in learner's understanding.

The Life of Nelson Mandela video from the Become 8th Grade textbook, has eight of Mayer's (2009) principles in its design. The first, The Redundancy Principle, can be observed as the video presents information, according to Plass and Jones (2005), in the forms of oral and written language input, with subtitling, which, according to Mayer (2009), could hurt learning. As this Doctoral Dissertation is concerned about the context of learning English as an L2, providing learning with redundant information can be beneficial for the comprehension process, as already discussed by Danan (2004), Hayati and Mohmedi (2010), and Lavaur and Bairstow (2011). The second, The Spatial Contiguity Principle, in which the imagery and the written information are positioned in a way next to each other, in the case of the Life of Nelson Mandela video the imagery information is positioned under the written information (subtitles) in order to present a hierarchal structure. The third, The Temporal Contiguity Principle, can be observed as both the subtitles and the imagery information are presented only when the narration mentions them. The fourth, The Modality Principle, in which learners are provided with the information in the form of narration and imagery information can also be observed in the video design. The fifth, The Multimedia Principle, are also part of the video design, once the information is presented by means of imagery information and verbal information. The sixth and the seventh, The Personalization Principle and The Voice Principle are part of the video design, once the narration is designed in a conversational style presenting a human voice, rather than a formal, machine voice. Finally, the eighth principle, The Image Principle, that can be detected in the video design concerns with the fact that the absence of the image of the narrator may not have an effect on learners' understanding.

Six of the multimedia design principles suggested by Mayer (2009) can be detected in the video *Usando Going to e Will para Falar de Futuro* from the Time to Share 8th Grade textbook. The first principle, The Redundancy Principle, can be observed as information is presented in the written form, that is, in the form of subtitles as discussed by Danan (2004), Hayati and Mohmedi (2010), and Lavaur and Bairstow (2011), and as on-screen written information and orally by the narrator. The second, The Spatial Contiguity Principle, can be observed as the written information is presented in a way that the sentences presenting the translation are positioned under the sentences presetting the target language structure examples so to help learners making connections between the information in their L1 and L2. The third principle that can be observed, The Temporal Contiguity Principle, is part of the video design,

once the written information (both subtitles and on-screen written text) are presented simultaneously as the narrator calls learners' attention to them and produces them. The fourth principle, The Personalization Principle, can be defined as being part of the video design, once information is presented to learners in a rather information and conversational fashion, resembling a teacher-student interaction. Finally, the fifth and the sixth principles, The Voice Principle and The Image Principle respectively, can be observed as the narrator's image is presented as an on-screen, probably in order to establish social engagement by part of the learners, who presents the information with his own human voice.

In the Interchange Intro – Unit 1 video, from the multimedia principles proposed by Mayer (2009) six principles can be observed in the design of the video. The Temporal Contiguity Principle can be observed when the school bell rings, suggesting that it is time for the classes to begin, and the two students look up and a clock is showing two minutes remaining to nine o'clock, The Modality Principle is associated to the video presenting the target language by means of visual information (the characters) and the aural information (the spoken language). The Multimedia Principle can be seen once the information in the video is presented by means of images and words, instead of presenting only words. The Personalization Principle related to the oral language produced in an informal conversational manner. Finally, The Voice Principle and The Image Principle can be examined in the video design as the oral information is presented by means of human voices, for the characters are all human beings, and their figures can be seen on screen throughout the video.

The Interchange Intro – Unit 16 has a reduced number of principles that can be seen in its design, that is, three. The first, The Personalization Principle, can be observed as the information in the video is presented in a conversational way, due to the fact that it depicts two school who encounter at school at exchange information about the previous weekend. The second, The Voice Principle, concerns about the narration of the fact provided about the previous weekend is done by the girl, ultimately provided to learners in a human voice, alternatively to a voice produced by a machine. Finally, The Image Principle, is the third of Mayer's (2009) principles detected in the video design, once learners can see the figures of the participants, especially the girl who is narrating the events that took place the previous weekend.

Regarding the Interchange 1 – Unit 8 video, six principles proposed by Mayer (2009) can be observed in its design. The first, The Temporal Contiguity Principle, can be examined. Although it should be part of the whole video design, it can be observed only once when the image of the storefront of a grocery store is shown on screen as the woman talks about the grocery store so that learners can integrate the words and the image presented. The second, The

Modality Principle, can also be observed as the video is designed in a fashion that provides learners with information by means of images and words. The third, The Multimedia Principle, is part of the video design once the information is presented to learners in the form of words (spoken utterances produced by the characters) and images (the motion pictures). The fourth, The Personalization Principle, can be examined as the video presents an informal conversational situation in which the characters are providing information about the places in the neighborhood. Lastly, the fifth and the sixth, The Voice Principle and The Image Principle can be argued as being part of the video design for the fact that, as the video presents a conversational interaction, the information is ultimately provided in a human voice to be in accordance with the characters whose figures are shown on screen.

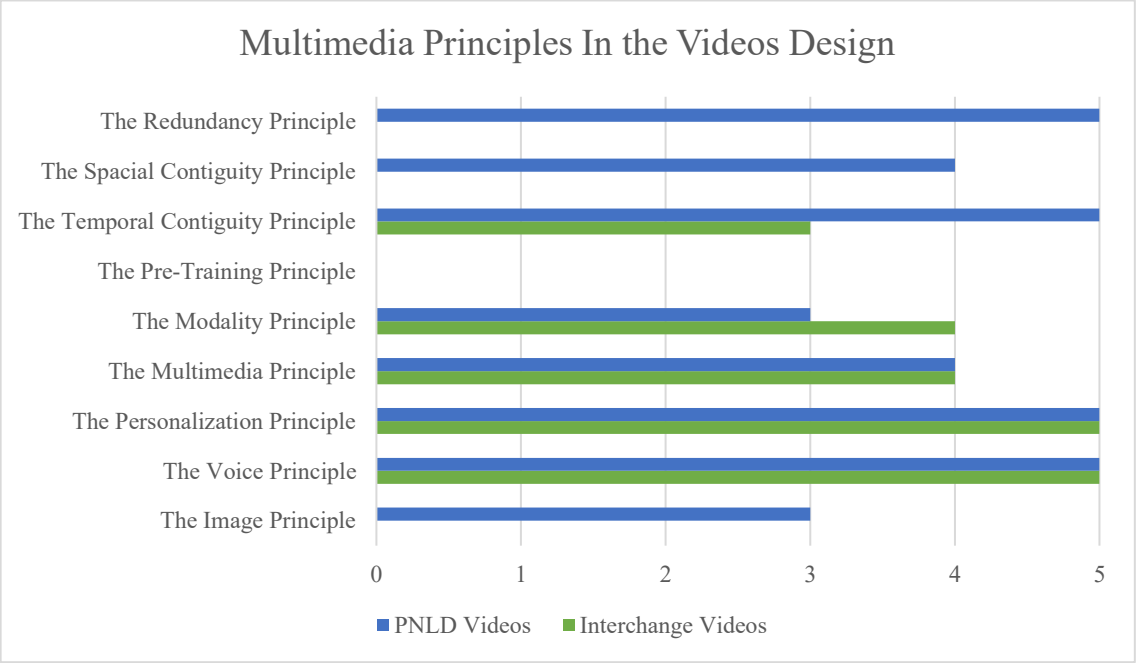
Six principles can also be observed in the design of the Interchange 2 – Unit 8 video. The first, The Temporal Contiguity Principle, is observable in the video design once the imagery information is presented simultaneously when the TV host or the interviewees provide specific pieces of information in order to depict the foods being described. The second, The Modality Principle, can be seen in the video design once the information is presented by means of imagery and words. The third, The Multimedia Principle, is part of the video design once the information is presented to learners in the form of words (spoken utterances produced by the characters) and images (the motion pictures). The fourth, The Personalization Principle, is part of the oral language design in the video is presented in a conversational informal style, depicting a real-life interaction. Lastly, the fifth and sixth principles that can be examined in the video design, The Voice Principle and The Image Principle, concern the fact that, as the information is provided by people, the spoken information is presented by human voices, also to be in accordance with the figures shown of the TV host and the people he is interviewing.

Finally, five principles proposed by Mayer (2009) can be identified in the design of the Interchange 3 – Unit 1 video. The first, The Modality Principle, can be seen as the information in the video is presented to learners in the form images, that is, with motion pictures and the use of words, in the case of the video, oral language. The second, The Multimedia Principle, is part of the video design for the fact that in order to aid learners in the process of knowledge construction, information is presented in the form of motion pictures and words, in a combined fashion. The third principle examined, The Personalization Principle, is observable as the information in the video is done so by means of an informal conversational style, probably in order to establish a social engagement with learners. Finally, the fourth and the fifth principles, The Voice Principle and The Image Principle, can be argued to be part of the video design once the whole information is presented in the form of human voices and the images of all the

characters depicted in the video are presented on screen, again, possibly to establish social engagement with learners to some extend.

A visual summary of the principles as to the number of videos from the two sources is presented in the following graph.

Graph 1 – Multimedia Principles Proposed by Mayer (2009) Detected in the Videos Aside from Signaling, Segmenting and Coherence



Source: The researcher (2022)

Based on the description and the graph presented above concerning the number of principles versus the number of videos, it is possible to observe that the nine principles may fall into 3 different categories: The first concerns with principles that were examined as being part of the design of the PNLD and the Interchange videos. The second concerns with principles that were detected only in one of the two sources (PNLD or Interchange) and, the third type, principles that were not detected to any extent.

From the principles that were identified in both the PNLD videos and in the Interchange videos, it is possible to say that the two principles that were part of the design of the ten videos were The Personalization Principle and The Voice Principle. One reason for the two principles in the design of all videos may be associated to presenting the target information by means of providing learners with images of real people, or characters that resemble real people to a minimum extend and the voices of people, instead of mechanized ones.

Concerning the principles that were not identified in all of the videos of the two sources, it is possible to say that The Multimedia Principle was examined in 4 out of the 5 Interchange

videos and in 4 out of the 5 PNLD videos. Such difference was identified due to the fact that, based on the principle, videos should make use of words and images in their design so to aid learners in the process of understanding the information being presented, which was not the case in all videos as presented in the analyses. Also, the Temporal Contiguity Principle was identified in the 5 PNLD videos while in 3 out of the 5 Interchange videos, for the oral information and the imagery information were presented simultaneously to be mentally integrated by learners, however, not in all videos.

The principles that were identified only in one of the sources (the PNLD videos) were three. The first is The Redundancy Principle for the fact that all the 5 videos presented subtitles in their design, that is, the verbal language was presented both in the oral and written forms. The second principle, The Spatial Contiguity Principle is part of 4 out of the 5 videos for the imagery information was positioned near the written information they were associated to and, the third, The Image Principle is part of 3 out of the 5 videos, due to the fact that the narrator's figure is absent from the screen in only three videos.

Finally, The Pre-Training Principle is the only one that was not detected in any of the 10 videos analyzed. One possible reason is that, as videos accompany the textbooks, a pre-training could be part of a task-cycle to which the videos are part. Thus, the pre-training would not be inserted in the video design, but possibly as a pre-watching stage.

The analyses of the 10 videos based on the framework whose three major principles for the design of effective educational videos are Signaling, Segmenting and Coherence, can be argued to invite the necessity not only for video (and materials) designers, but also for language teachers and instructors to be acquainted with the human cognitive architecture (KALYUGA, 2015; SWELLER, 2010). The knowledge of the foundations of how information is processed and turned into knowledge in humans' minds may be at the service of teachers and materials designers so that their choices for when developing their own materials, or when choosing from a number of already-designed materials available, be informed.

5 FINAL REMARKS

The incorporation of videos in the L2 classroom can be of paramount importance (especially due to the fact that digital technology is evolving each day), but a critical view on how this medium can aid learners in the process of making sense not only of the language, but also of the information watched is fundamental.

The analyses of the videos were grounded on the discussions presented in chapter 2, which take a neuroscientifical perspective of the processes involved in learning. The discussions concerned the features that videos should present in their design, such as the presentation of subtitles in order to aid learners in the process of understanding the oral information by integrating it to the oral language, the provision of information as contextualized input so that they can correlate the target language to real-life situations they may encounter outside the classroom, and the possibility of understanding the information by means of oral and visual input. Furthermore, the structure of the human cognitive architecture (SWELLER; AYRES; KALYUGA, 2011) and how the sensory memory, working memory, and long-term memory play a crucial role in the way information is processed and integrated as new strings of knowledge as well as the principles proposed by Mayer (2009) for the design of effective educational materials grounded on the Cognitive Theory of Multimedia Learning were discussed.

Grounded on the discussions, the ten videos were analyzed as to what extent their design features could be considered to facilitate or hurt learning. The videos were taken from two different sources, that is, the videos accompanying the textbooks approved by the PNLD program, and the videos that accompany the Interchange textbooks.

The five videos from the PNLD program, for being complementary materials to the three different textbooks (thus, from three different publishers), present variety in their design, ranging from providing learners with cartoon-like designs, content information about people or a certain type of medium, to providing learners with explanations and examples of a specific target language structure. When considering the videos according to the texbooks, however, they show similar characteristics among each other. Such similarities can be examined, for example, in the two first videos analyzed from the textbooks It Fits, in which they are designed in a fashion presenting characters that are not humans, but cartoons interacting and exchanging personal information (6^{th} Grade) and presenting the target language about abilities (7^{th} Grade). The third and fourth videos, from the textbooks Become, are also designed in a similar fashion, providing narrations concerning what PSAs are and their aim (7^{th} Grade), and reporting the life

events of Nelson Mandela (8th Grade). Finally, the last video, from the textbook Time to Share, is also designed in a specific structure, presenting a teacher (narrator) providing grammatical explanations about the target-language structure used to refer to future (8th Grade).

Although the videos from the PNLD can be considered of a different design from the Interchange videos, the first video is the one that poses a considerably similar design, for it also presents a conversational interaction. The only difference, compared to the Interchange videos, is the fact that it is designed as a cartoon, whereas the latter are designed presenting real people.

The five videos that accompany the Interchange textbooks, in turn, for being from the same publisher, can be examined as all being designed following the same fashion. The aim of the Interchange videos is to provide learners with conversational situations in which two or more characters interact using the target language in order to convey meaning in contextualized situations. These situations ranged from school encounters between strangers in their first day of study (Interchange Intro – Unit 1) or between friends talking about past events (Interchange Intro – Unit 16), describing a neighborhood (Interchange 1 – Unit 8) and talking about holidays celebrations (Interchange 2 – Unit 8), and TV shows about dating (Interchange 3 – Unit 1).

By having an understanding of how learners' minds process information, and about how the design of the videos may impact learning, either positively, or negatively, their choices of materials may become informed. Informed choices can be defined as teachers choosing videos with a critical view, in which their choices will not only be regarding whether the video may be considered aesthetically appropriate, or even appealing, but mostly based on their understanding about videos whose design may support learners in the processing of the information during a video viewing and in the process of turning that into knowledge. This is the case, for example, when videos from YouTube are used in the L2 classroom. As it is known, YouTube can provide a countless variety of videos that can be incorporated to the teaching practice. However, there seems to be a countless amount of audiovisual materials aimed at aiding the learning process, thus, demanding choices of the videos to be grounded on the understanding of how humans process information, instead of choosing a specific video only for considering it appealing or that learners might like it, or find it interesting. The fact that learners might like a video does not mean that they will learn from it. Instead, videos should be chosen based on their design, considering how they can be cognitively suitable for learners in aiding them in processing the information watched and constructing knowledge from it.

The discussions previously presented allows us to advocate that attention is needed to the process involved in the development of videos that are aimed at assisting the learning of an L2. That means that when educational videos are designed, or chosen, such process needs to be

planned considering te human cognitive architecture due to the fact that their design can have direct impact on learning.

This Doctoral Dissertation, thus, aimed at providing insights not only to material designers, but also to teachers and language instructors concerning the effects videos can have on learning, based on their design so that their view concerning audiovisual materials become more critical, especially, for example, according to Almurashi (2016) and Shoufan (2019), when teachers wish to use videos from YouTube in their classes.

5.1 PEDAGOGICAL IMPLICATIONS

This Doctoral research study can contribute to the field of applied linguistics, more specifically in the area of CALL as extensively discussed by Martins and Moreira (2012), and Walker and White (2017). The analyses provided insights about the importance videos can have in aiding the process of learning an L2. The most remarkable issue dealt with in this research study is the fact that, when the pedagogical actions are based on the understanding of how humans process information, as is discussed by Weinstein, Sumeracki and Caviglioli (2019), and, thus, how learning takes place, the design not only of new materials, but the choices of already-made materials available deserve reflection. However, there seems to be a need for teachers to be acquainted with the theories and concepts about how the human mind processes information so that their choices for materials become informed choices. Such discussions are of utmost importance in every teacher education undergraduate and graduate courses, and need to be carried out due to the fact that teachers and instructors need to, at least to some degree, have the notion of how learning takes place.

That said, it is possible to estate that the understanding of how the human cognitive architecture is structured as well as the complex processes that take place within humans' minds while learning occurs was highly important and meaningful for the researcher. By having such understanding on what learning comprises, teachers and material designers can make their choices in a rather informed way. In the case they develop materials to be used in their classes, they assume the role of designers and can apply the principles proposed by Mayer (2009) in designing materials that can benefit learning. Such design, could follow the structure slides presentation have, in which the information to be transmitted is broken into blocks, or segments, so to present learners with a small portion at a time. Moreover, each block should direct, or signalize, learners to the parts of information that are crucial to their comprehension and, finally, as is also explained by Carr (2010), any pieces of information that might cause

distractions to learners should be avoided, in order to keep the material coherent, so that their working memory limits are not exceeded. Distracting elements also encompass embellishment features, so much common in most (if not all) slides presentations.

Now, when the decision is made to use materials that are ready for use, they can have their choices based on how the materials are designed, using as choice criteria their knowledge concerning Mayer's (2009) twelve design principles. The most important issue is that every use of materials aimed at aiding the process of knowledge construction should be based on the impact they can have on learners' cognitive system. In the case of incorporating videos in the L2 learning context, by having a deeper understanding of the way learning can be influenced by well-designed videos, teachers' practice can be improved to assist learners in the process of knowledge construction, and provide them with better means to achieve their goals.

Even though this Doctoral study was concerned with the analyses of educational videos, such notion can go beyond such resource. When teachers and instructors intend to provide their learners with sources such as websites, encyclopedia-like worksheets, or slides (to name some resources) to present explanations and the target language in the forms of imagery and oral input, the use of the signaling principle can aid learners in directing their attention to the crucial pieces of information. In addition, the segmenting principle can be of high assistance once the information can be administered in the proper amount in order to prevent learners from engaging in extraneous processing, that is, processing that exceeds their working memory capacities. Finally, by applying the coherence principle, when educational materials are designed, information that does not contribute to learning should be left out, also to prevent extraneous processing, ultimately hurting learning. This study can, thus, provide teachers, instructors and materials designers with valuable insights to be used when producing educational materials, and concerning their choices of the materials to be used in their classes.

5.2 LIMITATIONS OF THE STUDY AND SUGGESTIONS FOR FURTHER RESEARCH

The limitation of the study refers to the fact that videos were analyzed based on theoretical background presented in chapter 2. It would be extremely important to conduct further research concerning whether the videos could, indeed, benefit learning by means of applying them to educational contexts with participants in order to generate data concerning learners' understanding of the information and language watched.

Another suggestion for future research is the fact that the research study in this Doctoral Dissertation was concerned primarily with the discussion of Mayer's (2009) twelve principles

for the design of multimedia educational materials, which narrowed for the use of only three in the video analysis framework. Perhaps, with the development of research around the globe, the number of principles proposed for the development of multimedia materials may go beyond the twelve, which would also require further investigation as to how they can benefit learners in the context of learning an L2.

Finally, another important issue for further investigation is the use of music as soundtrack. Some of the videos presented soundtrack which, according to the Coherence Principle proposed by Mayer (2009), could be considered irrelevant, resulting in possibly causing extraneous processing which, as discussed by Ayers and Paas (2007), Mayer (2009), Mayer and Moreno (2010), and Sweller (2010), due to exceeding learners' WM resources (KALYUGA, 2010; MAYER, 2009; SWELLER, 2010). Deeper research could be conducted, however, in order to investigate whether and how the soundtrack might influence learners concerning their emotional or affective engagement (YADAV, *et al.*, 2011) to the video content. Such study might provide interesting data not only for materials designers, but also for teachers and provide insights about the psychological aspects of learning.

REFERENCES

ALMURASHI, W. A. The Effective Use of You Tube Videos for Teaching English Language in Classrooms as Supplementary Material at Taibah University in Alula. **International Journal of English Language and Linguistics Research**, [*s.l.*], v. 4, n. 3, p. 32-47, 2016.

AL-SEGHAYER, K. The Effect of Multimedia Annotation Modes on L2 Vocabulary Acquisition: A Comparative Study**. Language Learning & Technology**, [*s.l.*], v. 5, p. 202-232, jan, 2001.

AJIDEH, P. Schema Theory-Based Pre-Reading Tasks: A Neglected Essential in The ESL Reading Class. **The Reading Matrix**, [*s.l.*], v. 3, n. 1, p. 1-14, April, 2003.

AMBROSE, S. A. *et al.* **How Learning Works**: 7 Research-Based Principles for Smart Teaching. San Francisco: Jossey-Bass, 2010.

AN, S. Schema Theory in Reading. **Theory and Practice in Language Studies**, [*s.l.*], v. 3, n. 1, p. 130-134, January, 2013.

ARDESHIR, J. Schema Theory: How is it Activated in L2 Readers? **International Journal of Current Research** [*s.l.*], v. 7, n. 2, p. 12.987-12.990, February, 2015.

AYRES, P.; PAAS, F. Making Instructional Animations More Effective: A Cognitive Load Approach. **Applied Cognitive Psychology**, [*s.l.*], v. 21, p. 695-700, 2007.

BAJRAMI, L.; ISMAILI, M. The Role of Video Materials in EFL Classrooms. **Procedia – Social and Behavioral Sciences**, [*s.l.*], v. 232, p. 502-505, 2016.

BENKADA, C.; MOCCOZET, L. Enriched Interactive Videos For Teaching and Learning. *In*: IEEE – INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS 21ST INTERNATIONAL CONFERENCE INFORMATION VISUALISATION (IV*)*, p. 344-349, July, 2017. **Anais** […]. 2017.

BLAKE, R. J. Technologies for Teaching and Learning L2 Speaking. *In:* CHAPELLE, C. A.; SAURO, S. (ed.). **The Handbook of Technology and Second Language Teaching and Learnin***g*. Oxford, UK: Wiley Blackwell, 2017. p. 107-117.

BRAME, C. J. Effective Educational Videos: Principles and Guidelines for Maximizing Student Learning from Video Content. **CBE – Life Sciences Education**, [*s.l.*], v. 15, n. 6, p. 1-6, May, 2016.

BRASIL. Ministério da Educação. **Fundo Nacional de Desenvolvimento da Educação**. Brasília, MEC/FNDE, 2022. Available in: http://www.fnde.gov.br/component/k2/item/518-hist%C3%B3rico. Access in: April 31st, 2022.

BRASIL. Ministério da Educação. **PNLD**. Brasília, MEC, 2022. Available in: http://portal.mec.gov.br/component/content/article?id=12391:pnld. Access in: April 31st, 2022.

BUCHNER, J. How to Create Educational Videos: From Watching Passively to Learning Actively. **Open Online Journal for Research and Education**, [*s.l.*], v. 12, September, 2018. Available in: https://www.researchgate.net/publication/327423659_How_to_create_Educational_Videos_From_watching_passively_to_learning_actively. Access in: July 16th, 2021.

BULL, G. L.; BELL, L. Introduction: Teaching with Digital Video. *In:* BULL, G. L.; BELL, L. (ed.). **Teaching with Digital Video**. Washington, DC: International Society for Technology in Education, 2010. p. 1-11.

CARR, N. **The Shallows**: What the Internet is Doing to our Brains. New York: W.W. Norton & Company, 2010.

CHAPELLE, C. A.; SAURO, S. Introduction to the Handbook of Technology and Second Language Teaching and Learning. *In:* CHAPELLE, C. A.; SAURO, S. (ed.). **The Handbook of Technology and Second Language Teaching and Learning.** Oxford, UK: Wiley Blackwell, 2017. p. 1-9.

CHI, D. N. A Review of Richard Schmidt's Noticing Hypothesis. **International Journal on Studies in English Language and Literature (IJSELL)**, [*s.l.*], v. 6, n. 3, p. 5-10, March, 2018.

CHINNERY, G. M. Going to the MALL: Mobile Assisted Language Learning. **Language Learning & Technology**, [*s.l.*], v. 10, n. 1, p. 9-16, jan. 2006. Available in: https://www.learntechlib.org/p/74432/. Access in: 17 jan. 2022.

CLARK, J. M.; PAIVIO A. Dual coding theory and education. **Educational Psychology Review**, [*s.l.*], v. 3, p. 149-210, 1991.

DANAN, M. Captioning and Subtitling: Undervalued Language Learning Strategies. **Meta**, [*s.l.*], v. 49, n. 1, p. 67–77, April, 2004.

DAVIES, G.; OTTO, S. E. K; RÜSCHOFF, B. Historical Perspectives on CALL. *In:* THOMAS, M.; REINDERS, H.; WARSCHAUER, M. (ed.). **Contemporary Computer-Assisted Language Learning.** Bloomsburry Publishing, London, UK, 2013. p. 20-38.

DECI, E. L.; RYAN, R. M. **Intrinsic Motivation and Self – Determination in Human Behavior**. New York: Plenum, 1985.

DÖRNYEI, Z. **Research Methods in Applied Linguistics**: Quantitative, Qualitative and Mixed Methodologies. Oxford: Oxford University Press, 2007.

DUNKEL, P. A. Developing Listening Fluency in L2: Theoretical Principles and Pedagogical Considerations. **The Modern Language Journal**, [*s.l.*], v. 70, n. 2, p. 99-106, 1986.

ELLIS, R. **The study of Second language acquisition**. Oxford: Oxford University Press, 1994.

FARÍAS, M. *et al.* Evaluating Types and Combinations of Multimodal Presentations in the Retention and Transfer of Concrete Vocabulary in EFL Learning. **Revista Signos Estudos de Linguística**, Viña del Mar, v. 47, n. 84, 2013.

FERREIRA, O. M. de C.; JÚNIOR, P. D. da S. **Recursos Audiovisuais no Processo Ensino-Aprendizagem.** São Paulo: EPU, 1986.

FINARDI, K. R.; PORCINO, M. C. Tecnologia e Metodologia no Ensino de Inglês: Impactos da Globalização e da Internacionalização. **Ilha do Desterro**, [*s.l.*], n. 66, p. 239-282, jan.-jun. 2014.

GASS, S. Integration Research Areas: A framework for Second Language Studies. **Applied Linguistics**, [*s.l.*], v. 9, p. 198-217, 1988.

GOMES, L. F. Vídeos Didáticos: Uma Proposta de Critérios para Análise. **Travessias Pesquisas em Educação, Cultura, Linguagem e Arte**., Paraná, v. 2, n. 3, 2008. Available

in: http://e-revista.unioeste.br/index.php/travessias/article/view/3128/2463. Access in: September 16th, 2020.

GREENFIELD, P. M. **O Desenvolvimento do Raciocínio na Era da Eletrônica**: os Efeitos da TV, Computadores e Videogames. São Paulo: Summus, 1988.

GRIFFITHS, J.; STUBBS, G.; WATKINS, M. From Course Notes to Granules: A Guide to Deriving Learning Object Components. **Computers in Human Behavior**, [*s.l.*], v. 23, p. 2.696-2.720, 2007.

HAMBROOK, J. Video in Foreign Language Teaching. **Per Linguam**, [*s.l.*], v. 2, n. 1, p. 25-28, 1986.

HAMMOND, T. C.; LEE, J. K. Digital Video in Social Studies Education. *In*: BULL, G. L.; BELL L. (ed.). **Teaching with Digital Video**. Washington, DC: International Society for Technology in Education, 2010. p. 15-38.

HASLER, B. S.; KERSTEN, B.; SWELLER, J. Learner Control, Cognitive Load and Instructional Animation. **Applied Cognitive Psychology**, [*s.l.*], v. 21, p. 713-729, 2007.

HAYATI, A. M.; MOHMEDI, F. The Effect of Films With and Without Subtitles on Listening Comprehension of EFL Intermediate Students. **Int J. of Instructional Media**, [*s.l.*], v. 37, n. 3, p. 301-313, 2010.

HOQUE, Md. E. An Introduction to the Second Language Acquisition. **Education and Development Research Council (EDRC)**, [*s.l.*], p. 1-23, 2017.

IBRAHIM, M. Implications of Designing Instructional Video Using Cognitive Theory of Multimedia Learning. **Critical Questions in Education**, [*s.l.*], v. 3, n. 2, p. 83-104, 2012.

IBRAHIM, M. *et al.* Effects of Segmenting, Signaling, and Weeding on Learning from Educational Video. **Learning, Media and Technology**, [*s.l.*], v. 37, n. 3, p. 220-235, 2012.

INNACI, D. L.; SAM, P. Using Context Clues as a Vocabulary Learning Strategy: An Experimental Study. **Journal of English and Literature (JOELL)**, [*s.l.*], v. 4, n. 3, p. 39-43, 2017.

JONES, L. C.; PLASS, J. L. Supporting Listening Comprehension and Vocabulary Acquisition in French with Multimedia Annotations. **The Modern Language Journal**, [*s.l.*], v. 86, n. 4, p. 546-561, 2002.

JUFFS, A.; HARRINGTON, M. aspects of Working Memory in L2 Learning. **Language Teaching**, [*s.l.*], v. 44, n. 22, p. 137-166, 2011.

KAJDER, S. B.; YOUNG, C. A. Digital Video in English Language Arts Education. *In* BULL, G. L.; BELL L. (ed.). **Teaching with Digital Video**. Washington, DC: International Society for Technology in Education, 2010. p. 107-130.

KALYUGA, S. Schema Acquisition and Sources of Cognitive Load. *In:* PLASS, J. L; MORENO, R.; BRÜNKEN, R. (ed.). **Cognitive Load Theory**. Cambridge, UK: Cambridge University Press, 2010. p. 48-64.

KALYUGA, S. **Instructional Guidance**: a Cognitive Load Perspective. Charlotte: Information Age Publishing, 2015.

KALYUGA, S. **Managing Cognitive Load in Adaptive Multimedia Learning**. New York: Information Science Reference, 2009.

KALYUGA, S.; CHANDLER, P.; SWELLER, J. Managing Split-Attention and Redundancy in Multimedia Instruction. **Applied Cognitive Psychology**, [*s.l.*], v. 13, p. 351-371, 1999.

KIM, D.; GILMAN, D. A. Effects of Text, Audio, and Graphic Aids in Multimedia Instruction for Vocabulary Learning. **Educational Technology & Society**, [*s.l.*], v. 11, n. 3, p. 114-126, 2008.

KONING, B. B. de *et al.* Attention Cueing as a Means to Enhance Learning From an Animation. **Applied Cognitive Psychology**, [*s.l.*], v. 21, p. 731-746, 2007.

KONING, B. B. de *et al.* Learning by Generating Vs. Receiving Instructional Explanations: Two Approaches to Enhance Attention Cueing in Animations. **Computers & Education**, [*s.l.*], v. 55, p. 681-691, 2010.

KRASHEN, S. D. **Second language acquisition and second language learning.** Oxford: Pergamon, 1981.

KRASHEN, S. D. **Principles and Practice in Second Language Acquisition**. Oxford: Pergamon, 1982.

KRESS, G.; VAN LEEUWEN, T. Colour as a Semiotic Mode. **Notes for a Grammar of Colour**, [*s.l.*], v. 1, n. 3, p. 343-368, 2002.

KRUGER, J. L. Subtitles in the Classroom: Balancing the Benefits of Dual Coding With the Cost of Increased Cognitive Load. **Journal for Language Teaching**, [*s.l.*], v. 47, n. 1, p. 55-71, 2013.

LAVAUR, J. M.; BAIRSTOW, D. Languages on the screen: Is film comprehension related to the viewers' fluency level and to the language in the subtitles? **International Journal of Psychology**, [*s.l.*], v. 46, n. 6, p. 455-462, 2011.

LEAHY, W.; SWELLER, J. Cognitive Load Theory and the Effects of Transient Information on the Modality Effect. **Instructional Science**, [*s.l.*], v. 44, p. 107-23, 2016.

LEE, H.; MAYER, R. E. Fostering Learning from Instructional Video in a Second Language. **Applied Cognitive Psychology**, [*s.l.*], v. 32, p. 648-654, 2018.

LEOW, R. P. Simplification and Second Language Acquisition. **World Englishes**, [*s.l.*], v. 16, n. 2, p. 291-296, 1997.

LONERGAN, J. **Video in Language Teaching**. Cambridge: Cambridge University Press, 1984.

LORCH, R., F. Jr. Text-Signaling Devices and Their Effects on Reading and Memory Process. **Educational Psychology Review**, [*s.l.*], v. 1, n. 3, p. 209-234, 1989.

MA, Q. Technologies for Teaching and Learning Vocabulary. *In* CHAPELLE, C. A.; SAURO, S. (ed.). **The Handbook of Technology and Second Language Teaching and Learning.** Oxford, UK: Wiley Blackwell, 2017. p. 45-61, 2017.

MARTINS, C. B. M. J.; MOREIRA, H. O Campo CALL (Computer Assisted Language Learning): Definições, Escopo e Abrangência. **Calidoscópio**, [*s.l.*], v. 10, n. 3, p. 247-255, set.-dez. 2012.

MAYER, R. E.; HEISER, J.; LONN, S. Cognitive Constraints on Multimedia Learning: When Presenting More Material Results in Less Understanding. **Journal of Educational Psychology**, [*s.l.*], v. 93, n. 1, p. 187-198, July, 2001.

MAYER, R. E. **Multimedia Learning.** 2. ed. New York: Cambridge University Press, 2009.

MAYER, R. E.; SIMS, V. K. For whom is a picture worth a thousand words? Extensions of a Dual Coding Theory of Multimedia Learning. **Journal of Educational Psychology**, [*s.l.*], v. 86, p. 389-401, January, 1994.

MAYER, R. E.; LEE, H.; PEEBLES, A. Multimedia Learning in a Second Language: A Cognitive Load Perspective. **Applied Cognitive Psychology**, [*s.l.*], v. 28, p. 653-660, 2014.

MAYER, R. E.; MORENO, R. Nine Ways to Reduce Cognitive Load in Multimedia Learning. **Educacional Phsychologist**, [*s.l.*], v. 38, n. 1, p. 43-52, 2003.

MAYER, R. E.; MORENO, R. Techniques That Reduce Extraneous Cognitive Load and Manage Intrinsic Cognitive Load during Multimedia Learning. *In:* PLASS, J. L.; MORENO, R.; BRÜNKEN, R. (ed.). **Cognitive Load Theory**. Cambridge, UK: Cambridge University Press, 2010. p 131-152.

MAYER, R. E. Introduction to Multimedia Learning. *In:* MAYER, R. E. (ed.). **The Cambridge Handbook of Multimedia Learning**. New York, USA: Cambridge University Press, 2005a. p. 1-16.

MAYER, R. E. Cognitive Theory of Multimedia Learning. *In:* MAYER, R. E. (ed.). **The Cambridge Handbook of Multimedia Learning.** New York, USA: Cambridge University Press, 2005b. p. 31-48.

MITTERER, H.; MCQUEEN, J. M. Foreign Subtitles Help but Native-Language Subtitles Harm Foreign Speech Perception. **PLoS ONE**, [*s.l.*], v. 4, n. 11, p. 1-5, November, 2009.

MORAN, J. O Vídeo na Sala de Aula. **Comunicação & Educação**. São Paulo: ECA-Ed; Moderna, 1995. v. 2, p. 27-35.

MORENO, R.; MEYER, R. E. Verbal Redundancy in Multimedia Learning: When Reading Helps Listening. **Journal of Educational Psychology**, [*s.l.*], v. 94, n. 1, p. 156-163, May, 2002.

MORENO, R. Optimising Learning from Animations by Minimising Cognitive Load: Cognitive and Affective Consequences of Signaling and Segmentation Methods. **Aplied Cognitive Psychology**, [*s.l.*], v. 21, p. 765-781, 2007.

MOTTERAM, G.; SLAOUTI, D.; ONAT-STELMA, Z. Second Language Teacher Education for CALL: An Alignment of Practice and Theory. *In:* THOMAS, M.; REINDERS, H.; WARSCHAUER, M. (ed.). **Contemporary Computer-Assisted Language Learning**. Bloomsburry Publishing, London, UK, 2013. p. 56-71.

NOOR, S. F. M.; YUSOF, N.; HASHIM, S. Z. M. Creating Granular Learning Object Towards Reusability of Learning Object In E-learning Context. *In*: 2011 INTERNATIONAL CONFERENCE ON ELECTRICAL ENGINEERING AND INFORMATICS, 17-19 July, Bandung, Indonesia, p. 1-5, 2011. **Anais** […]. Indonesia, 2011.

PAIVIO, A. **Mental Representations**: A Dual Coding Approach. New York: Oxford University Press, 1990.

PLASS, J. L. *et al.* Supporting Visual and Verbal Learning Preferences in a Second-Language Multimedia Learning Environment. **Journal of Educational Psychology**, [*s.l.*], v. 90, p. 25-36, August, 1998.

PLASS, J. L.; JONES, L. C. Multimedia Learning in Second Language Acquisition. *In:* MAYER, R. E. (ed.). **The Cambridge Handbook of Multimedia Learning**. New York: Cambridge University Press, 2005. p. 467-488.

PLASS, J. L.; MORENO, R.; BRÜNKEN, R. (ed.). **Cognitive Load Theory**. Cambridge: Cambridge University Press, 2010.

PARK, Y.; JUNG, E. Exploring the Use of Video-clips for Motivation Building in a Secondary School EFL Setting. **English Language Teaching: Canadian Center of Science and Education**, Canada, v. 9, n. 10 p. 81-89, September, 2016.

PEREZ, M. M.; NOORGATE, W. V. D.; DESMET, P. Captioned Video for L2 Listening and Vocabulary Learning: A Meta-Analysis. **System**, [*s.l.*], v. 41, p. 720-739, August, 2013.

PNLD – PROGRAMA NACIONAL DO LIVRO DIDÁTICO. **Ministério da Educação**, 2019. Available in: http://portal.mec.gov.br/component/content/article?id=12391:pnld. Access in: Dec. 10th, 2019.

POLAT, M.; ERISTI, B. The Effects of Authentic Video Materials on Foreign Language Listening Skill Development and Listening Anxiety at Different Levels of English Proficiency. *In*: INTERNATIONAL JOURNAL OF CONTEMPORARY EDUCATIONAL RESEARCH. v. 6, n. 1, p. 135-154, June, 2019. **Anais** […]. [*S.l.*], 2019.

PUJADAS, G.; MUÑOZ, C. Extensive Viewing of Captioned and Subtitled TV Series: a Study of L2 Vocabulary Learning by Adolescents. **The Language Learning Journal**, [*s.l.*], v. 47, n. 4, p. 479-496, June, 2019.

REED, S. K. Manipulating Multimedia Materials. *In:* ZHENG, R. Z. (ed.). **Cognitive Effects of Multimedia Learning**. New York: Information Science Reference, 2009. p. 51-65.

ROST, M. Listening. *In:* CARTER, R.; NUNAN, D. (ed.). **The Cambridge Guide to Teaching English to Speakers of Other Languages.** Cambridge: Cambridge University Press, 2001. p. 7-13.

SACERDOTE, H. C. D. S. Análise do Vídeo como Recurso Tecnológico Educacional. **Revista de Educação, Linguagem e Literatura da UEG, Inhumas**, Goiás, v. 2, n. 1, p. 28-37, 2010.

SCHMIDT, R. Attention, Awareness, and Individual Differences in Language Learning. *In:* CHAN, W. M. *et al.* **Proceedings of CLaSIC 2010**. [Singapore]: National University of Singapore; Centre for Language Studies, 2010. p. 721-737.

SCHMIT, R. The Role of Consciousness in Second Language Learning. **Applied Linguistics,** Oxford University Press, v. 11, n. 2, 1990.

SILVA, R. C. Schema Theory: An Introduction and its Application in EFL Reading Classroom. **Revista de Letras**, Curitiba, v. 21, n. 34, p. 157-169, set. 2019.

SHARIFF, S. B. M.; SHAH, P. M. Pupils Perception of Using YouTube and Autonomous Learning. **Creative Education**, [*s.l.*], v. 10, p. 3.509-3.520, December, 2019.

SHOUFAN, A. Estimating the Cognitive Value of YouTube's Educational Videos: a Learning Analytics Approach. **Computers in Human Behaviour**, [*s.l.*], v. 92, p. 450-458, March, 2019.

SCHNOTZ, W. An Integrated Model of Text and Picture Comprehension. *In:* MAYER, R. E. (ed.). **The Cambridge Handbook of Multimedia Learning**. New York, USA: Cambridge University Press, 2005. p. 49-69.

SORDEN, S. D. A Cognitive Approach to Instructional Design for Multimedia Learning. **Informing Science Journal**, [*s.l.*], v. 8, p. 263-279, 2005a.

SORDEN, S. D. **The Cognitive Theory of Multimedia Learning**. Northern Arizona University, 2005b.

STANLEY, G. **Language Learning with Technology**: Ideas for Integrating Technology in the Classroom. Cambridge, UK: Cambridge University Press, 2013.

SUEK, L. A. Applying Cognitive Load Theory in Teaching Tenses for Second Language Learners. **Englisia**, [*s.l.*], v. 5, n. 2, p. 66-81, 2018.

SWELLER, J. Cognitive Load Theory: Recent Theoretical Advances. *In:* PLASS, J. L; MORENO, R.; BRÜNKEN, R. (ed.). **Cognitive Load Theory**. Cambridge, UK: Cambridge University Press, 2010. p. 29-47.

SWELLER, J.; AYRES, P.; KALYUGA, S. **Cognitive Load Theory**: Explorations in the Learning Sciences, Instructional Systems and Performance Technologies. New York: Springer, 2011.

SWELLER, J. Cognitive Load Theory and Teaching English as a Second Language to Adult Learners. **TESL Ontario, English Language Learning CONTACT Magazine**, Ontario, v. 43, n. 2. August, 2017.

THOMPSON, K.; YONEKURA, F. Practical Guidelines for Learning Object Granularity from One Higher Education Setting. **Interdisciplinary Journal of Knowledge and Learning Objects**, [*s.l.*], v. 1, p. 163-179, 2005.

THOMAS, M.; REINDERS, H.; WARSCHAUER, M. Contemporary Computer-Assisted Language Learning: The Role of Digital Media and Incremental Change. *In:* THOMAS, M.; REINDERS, H.; WARSCHAUER, M. (ed.). **Contemporary Computer-Assisted Language Learning.** Bloomsburry Publishing, London, UK, 2013. p. 1-12.

TUMOLO, C. H. S. Recursos digitais para ensino/aprendizagem de inglês como língua estrangeira: o vídeo em destaque. *In:* TOMITCH, L. M. B.; HEBERLE, V. M. (org.). **Perspectivas Atuais de Aprendizagem e Ensino de Línguas**. Florianópolis: LLE/CCE/UFSC, 2017. p. 51-75.

TUMOLO, C. H. S. Recursos digitais e aprendizagem de inglês como língua estrangeira. **Ilha do Desterro**, Florinaópolis, n. 66, p. 203-238, January-June, 2014.

TUYEN, L. V.; HUYEN, V. TN. Effects of using Contextual Clues on English Vocabulary Retention and Reading Comprehension. **International Journal of English, Literature and Social Science (IJELS)**, [*s.l.*], v. 4, n. 5, p. 1.342-1.347, Steptember-October, 2019.

VIEIRA, G.; FINARDI, K. R. The Case of EFL: Does Technology Lead to Learning? **Raído**, Dourados, MS, v. 12, n. 30, p. 197-210, jul.-dez, 2018.

WALKER, A.; WHITE, G. **Technology Enhanced Language Learning**: Connecting Theory and Practice. Oxford: Oxford University Press, 2017.

WANG, Z. An Analysis on the Use of Video Materials in College English Teaching in China. **International Journal of English Language Teaching**, [*s.l.*], v. 2, n. 1, p. 23-28, September, 2015.

WATKINS, J.; WILKINS, M. Using YouTube in the EFL Classroom. **Language Education in Asia**, [*s.l.*], v. 2, n. 1, p. 113-119, 2011.

WEINSTEIN, Y.; SUMERACKI, M.; CAVIGLIOLI, O. **Understanding How We Learn**: A Visual Guide. New York: Routledge, 2019.

WOCHIT. **All You Need is Overlay – 3 Great Examples of Text Overlay Videos**, 2022. Available in: https://www.wochit.com/blog/3-great-examples-text-overlay-video/#:~:text=Text%20overlay%20or%20video%20overlay,trend%20in%20social%20media%20content. Access in: March, 2, 2022.

WOOTTIPONG, K. Effect of Using Video Materials in the Teaching of Listening Skills for University Students. **International Journal of Linguistics**, [*s.l.*], v. 6, n. 4, p. 200-212, August, 2014.

YADAV, A. *et al.* If a Picture is Worth a Thousand Words Is Video Worth a Million? Differences in Affective and Cognitive Processing of Video and Text Cases. **J. Comput High Educ. Studies,** [*s.l.*], n. 23, p. 15-37, 2011.

ZAREI, G. R.; KHAZAIE, S. L2 Vocabulary Learning Through Multimodal Representations. **Procedia Social and Behavioral Sciences**, [*s.l.*], n. 15, p. 369-375, 2011.

ZHENG, R. **Cognitive Effects of Multimedia Learning**. New York: Information Science Reference, 2009.