

THE ACQUISITION OF FORMULAIC LANGUAGE THROUGH SUBTITLES: A STUDY ACROSS GENRES

MA programme in Applied Linguistics and
Language Acquisition in Multilingual Contexts



UNIVERSITAT DE
BARCELONA

Inés Sánchez de la Viña Rodríguez
Supervised by Dr. Roger Gilabert Guerrero

Department of English and German Philology Faculty of Philology
University of Barcelona
July, 2018

TABLE OF CONTENTS

Abstract.....	3
Acknowledgments.....	4
1. Introduction.....	5
2. Literature review.....	6
2.1. The role of subtitles in multimedia learning.....	6
2.2. Subtitles in the field of second language acquisition (SLA).....	7
2.3. Subtitles and formulaic language.....	8
2.4. The issue of genre.....	10
2.5. Conceptual framework.....	11
2.5.1. Cognitive theory of multimedia learning.....	11
2.5.2. Dual coding theory.....	13
3. The current study.....	14
3.1. Justification and objectives.....	14
3.2. Research questions and hypotheses.....	15
4. Methodology.....	15
4.1. Participants.....	15
4.2. Research design.....	16
4.3. Instruments.....	17
4.3.1. Pilot-study.....	18
4.3.2. Pre-test.....	18
4.3.3. Post-test.....	19
4.3.3.1. Genre videos.....	20
4.3.3.2. Comprehension test.....	21
4.3.3.3. Word form and meaning recognition tests.....	21
4.3.4. Delayed post-test.....	22
4.3.5. Retrospective protocol analysis.....	22
4.4. Procedure.....	22
5. Statistical analysis.....	23
6. Results.....	23
6.1. Results of the retrospective protocol analysis.....	28
7. Discussion.....	29
8. Limitations.....	31
9. Conclusion and further research.....	32
10. Bibliography.....	34
Appendix.....	37
Appendix A: List of figures and tables.....	37
Appendix B: Consent form.....	39
Appendix C: Web-research questionnaire (post-test).....	40
Appendix D: Retrospective protocol analysis questionnaire.....	58

ABSTRACT

The present Master's thesis adopts a multidisciplinary approach and blends the literature related to subtitles, formulaic language, multimedia learning, and teaching and learning English as a Foreign Language (EFL). The aims of the research are, first, to examine whether genre affects meaning recall of target multi-word expressions (MWEs) by EFL learners, and second, to explore whether input enhancement of target MWEs aids learning. 40 EFL adult learners participated in this study. Participants were exposed to four short clips subtitled in English. Half of the participants watched the videos with the target MWEs enhanced. The highlighting was removed from the target MWEs for the other half. The videos used belonged to four distinct genres: documentary, public lecture, detective story, and comedy. Overall, findings revealed an effect of genre. Additionally, results showed limited but positive effects of input enhancement of the target MWEs. Results were triangulated with qualitative findings from a retrospective protocol analysis questionnaire.

Keywords: EFL, subtitles, multimedia learning, videos, genre, formulaic language, input enhancement

ACKNOWLEDGEMENTS

First and foremost, I would like to express my special appreciation and gratitude to my supervisor Dr. Gilabert for his advice, encouragement, and for sharing his expertise with me. I am also thankful to Dr. Miralpeix for providing me with the knowledge I have been able to put into practise in my thesis. In addition, I would also like to thank Dr. Peinado, who has been an unofficial dedicated mentor and a friend. It is a pleasure to thank Javier Fernández. I am gratefully indebted to him for his valuable help on this thesis. I would like to thank all my classmates, most of whom have become close friends. All my love and most earnest gratitude to Fahad Alyaqout, Aiora Enparantza, Julie-Anne McKenzie, Britt Nooij, and Judit Serra. This journey would not have been the same without you. A special mention goes to my family. Words cannot express how grateful I am to my father and sister, whose love and guidance are with me in whatever I pursue. This accomplishment would not have been possible without you. Finally, I must express my warmest gratitude to David for being able to put up with me throughout this past six months.

1. Introduction

Conversation is an essential element of social life. Phrases such as, *What's up*, *Long time no see* or *Catch you later* are frequent English greeting and farewell locutions. The competent use of such colloquial or formulaic language is an important aspect of foreign language learning. In this regards, “formulaic language” refers to multi-word expressions (MWEs) such as idioms, collocations, phrasal verbs, and proverbs. Research in the field of second language acquisition (SLA) has shown that formulaic language constitutes a vital part of the language we speak. It contributes to more fluent conversation, it brings learners to native language performance, and it helps with integration and assimilation in the native language or lingua franca communities. Yet, formulaic language is resistant to second language acquisition and it posits real challenges to learners of English as a foreign language (EFL). In this context, a question arises: how do EFL learners acquire such resistant language feature?

English is the most widely studied foreign language in the world. To learn English entails to understand real-life English beyond its more academic or formal use, a central aspect of learning the language is mastering its conversational dimension. In the last few years, more and more people have been receiving a massive amount of English input due to exposure to the media. English comes from different sources through different channels: through social media, videogames, TV series and programmes, films, songs, among many others. In most European countries, TV series, films and other multimedia products are often broadcast in their original version with L1 subtitles or sometimes L2 captions. The spoken English coming from these channels is rich in formulaic language as it makes considerable use of recurrent formulaic phrases. In this context, a relevant issue to be examined is whether subtitling plays a significant role in the acquisition of formulaic language.

This research is a preliminary effort to empirically investigate the effect of genre on the acquisition of formulaic language from subtitled video. To do so, an experiment in which subtitles will be used and manipulated in order to check if formulaic language can be ‘early’ learnt or noticed through subtitles.

In the next section, the literature review will discuss several relevant studies on multimedia learning, the role of subtitles in SLA, formulaic language, and the issue of genre

2. Literature review

2.1. *The role of subtitles in multimedia learning*

Foreign language teaching and learning has opened up significantly in the last few years. New technologies have opened new possibilities to integrate multimedia support¹ such as visual aids in the English as Foreign Language (EFL) classroom. Different forms of multimedia support are often used to assist learners of English. Broadly speaking, multimedia platforms are increasingly present in the EFL classroom, and also as out-of-class review as a leisure exercise, as autonomous or collaborative learning. In hand with the expansion of multimedia learning, content visualisation has also evolved due to the rapid growth of new devices and applications in the form of platforms, software or tools. Over the years, the benefits related to the use of multimedia support in the language classroom have been a topic of lively debate. Nowadays, it is an undeniable fact that language learners live immersed in a media world, in which most of the information is provided by visual and aural input, through different technological devices.

In this light, a growing body of research is also interested in documenting the use of media platforms as teaching and learning tools. Surprisingly though, despite the explosion of multimedia learning, there have been few studies addressing the integration of specific multimedia tools in the EFL context. This is the case of subtitles, an area that has not been sufficiently researched or implemented in language classrooms. At this point, it is worth mentioning that multimedia tools such as subtitling require teachers to work hard on the selection, design, and adaptation of these materials. This adds up to the fact that multimedia support has to be tailored to diverse learning profiles, learners' abilities, interests, or skills. As a side note we must bear in mind that technological devices are not guaranteed in all educational contexts; and some teachers have few resources at their disposal.

There are multiple benefits to using audiovisual materials in learning and teaching English as a Foreign Language (EFL) (Danan, 2004; Vanderplank, 2010; Talaván, 2013). At the level of teaching, media platforms allow practitioners to personalise media content and direct learners' attention to relevant information. At the level of learning, language learners are more at ease in dealing with audiovisuals and so multimedia learning emerges as an

¹ Multimedia support refers to different combinations of text, audio, still images, animation, and/or video content.

appealing form of instruction. Research has shown that multimedia content stimulates and engages learners if integrated into student-centred learning activities (Willmot et al., 2012). In such context, audiovisual aids encourage strong associations for retention and language use. Danan (2004: 67) also argues that subtitles aid language learning “by helping students visualize what they hear, especially if the input is not too far beyond their linguistic ability”. Other studies have indicated that watching subtitled videos facilitates reading and listening comprehension skills, leading to additional cognitive benefits, such as greater depth of processing. Furthermore, multimedia material enhanced with subtitles can increase student motivation, enhance learning experience, and develop students’ autonomy. In spite of all of this, we still face big challenges in promoting and implementing the use of subtitles in multimedia learning as will be seen in the following sections.

2.2. Subtitles in the field of second language acquisition (SLA)

Studies in the field of neuroscience have shown that “significant increases in learning can be accomplished through the informed use of visual and verbal multimodal learning” (Fadel, 2008:12). Paivio’s (1971) dual coding model holds that multimedia learning aids students to learn more deeply thanks to the combination of words and graphics (Mayer, 2003). If we acknowledge the role of Paivio’s model, we can thereby find real-life applications to assess the effects of this model at the educational level. One line of research related to this is the processing and acquisition of formulaic language. We will later return to this issue in section 2.3. so as to provide a deeper understanding on the specific qualities of this type of language.

A good starting point to discuss the importance of subtitles is to acknowledge the work of early researchers such as Robert Vanderplank. Vanderplank was one of the first authors to ever explore the value of subtitles as an aid to second language acquisition. Vanderplank’s (1988) earliest work looked at the benefits of intralingual subtitles, i.e., L2 subtitles or subtitles in the target language. In his nine-week-long treatment fifteen European exchange students watched different BBC subtitled television programmes for one hour each week. Participants were asked to give specific feedback on their experience with the experiment. Results indicated that subtitles were beneficial to the language development of these EFL learners. All participants stated to have developed a series of strategies and techniques to derive the greatest benefits from the subtitles. Furthermore, “subjects reported that they were conscious of learning a great deal of language from the programmes watched. They talked of ‘finding’ new words and phrases which they would be able to use themselves” (1988: 275).

In spite of the overall positive aspects, Vanderplank (1988) asserts that intralingual subtitles are of more value to high-intermediate and post-proficiency level. This is something quite relevant as far as educational implications are concerned.

In a follow up study Vanderplank (1990) looked into the benefits and limitations of subtitles as a tool in language learning. Based on his first experiment (as presented above), and what other researchers had found, he argues that learners' attention is not affected as the use of subtitles is not a distraction for "the double modal input appears to enhance comprehension better than simple script or sound" (1990: 223). Furthermore, he argues that subtitling increases learners' motivation. Other authors, such as Talaván (2006), also emphasize motivation as a determining factor in arising learners' interests. Talaván highlights the entertaining nature of subtitles in classroom activities as a methodological approach (2006). At this point it is relevant to note that Talaván refers to the use of subtitles as a pedagogic technique within task-based language teaching (TBLT).

The goals of more recent studies examining the role of subtitles in second language acquisition have centred on different areas. At the level of overall comprehension, research has examined the potential benefits of subtitles in language learning and comprehension skills (Vanderplank, 2010; Birulés-Muntané & Soto-Faraco, 2016; Rodgers & Webb, 2017). In this context, subtitling has been proven to be beneficial for augmenting language listening skills and fostering vocabulary learning. Some studies suggest that subtitling can also lead to additional cognitive benefits, such as greater depth of processing. At the level of vocabulary learning, studies have focused on the impact of audiovisual input on vocabulary gains (Peters et al, 2016), and on the effects of manipulating captioned video (full vs. keyword captioning) to guide learners' attention to novel words and expressions (Montero-Perez et al., 2013). Another issue of interest is the effects of different types of subtitles (L1 subtitles, L2 subtitles) in comparison to captioning (Montero-Perez et al., 2015). Despite the achievements of recent research, many questions about how formulaic language is processed in the context of subtitles and learnt remain unanswered, among them; the most central and crucial one is how EFL learners acquire such language.

2.3. Subtitles and formulaic language

In natural language, native speakers often use multi-word expressions (MWEs) such as "pull strings" or "kick the bucket". MWEs are quite frequent in languages such as English.

Numerous research on MWEs have used different terms to define the same phenomena (Sinclair, 1991; Wray, 2002; Sprenger, 2003; Schmitt, 2004; Gries, 2008). MWEs have been defined as “sequences of more than one word that behave like ‘single choices’” (Sinclair, 1991). Sprenger identify them (using different terminology) as being “fixed expressions [referring] to specific combinations of two or more words that are typically used to express a specific concept” (2003:4). According to Sprenger, “the defining feature of a fixed expression is that it is a word combination, stored in the Mental Lexicon of native speakers, that as a whole refers to a (linguistic) concept” (2003: 4). As we can see, MWEs are complex by definition. The notable disparity in the literature reflects the inclusiveness or exclusiveness of the definitions. For the purposes of the current research MWEs will be referred to as a string of formulaic language, i.e., a sequence, continuous or discontinuous, of two or more words with distinct characteristics of use. In line with Wray’s notion of formulaic sequence the string of formulaic language “appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than subject to generation or analysis by the language grammar” (Wray, 2002: 9). All things considered, formulaic language is used here as the most comprehensible term due to its multi-faceted nature and its diverse usage in different contexts of real-life situations.

Having a good command of formulaic language makes second language speakers more fluent and efficient language users. The lack of knowledge about MWEs has been highlighted as a major factor for the degree to which non-natives are perceived as native-like. Research has shown that this is particularly crucial at upper-intermediate levels of proficiency (Sprenger et al., 2016). In the case of formulaic language, using subtitles opens many opportunities for learning. TV programmes, series, and films are rich in MWEs as fictional language tries to resemble natural language. Subtitles support EFL learners when watching non-dubbed versions of TV series, documentaries, and films. In this light, subtitles have been suggested to be an efficient channel for learners to familiarise themselves with MWEs for three main reasons: i) subtitles facilitate recall and meaning recognition, ii) the combination of different modes (visual, aural aids comprehension, and iii) words and expressions appear contextualised so EFL learners can decode unknown utterances. Here it is worth noting that, as these MWEs are embedded, language learners can guess ambiguous meanings, even if a string of formulaic language is ‘opaque’ in nature. At the same time subtitles make it easier to recognise and process formulaic language, reducing learners’ understanding efforts. In this

way, the link between subtitles, formulaic language, and audiovisuals aids in the form of videos is quite straightforward.

2.4. The issue of genre

In order to delve into what genre variation might entail we need to understand first the meaning of “genre”. In particular, in the context of television, genres can be described as consisting of a “coded set of formulas and conventions which indicate a culturally accepted way or organizing material into distinct patterns. Once established, genres dictate the basic conditions of cultural production and reception” (Kellner, 1980: 7). According to Kellner, genres are articulated in the sense that their basic structure and universal nature turn up to construct archetypes. Consequently, one of the distinguishing traits of genre is its habit-forming. For example, a detective story that does not involve the solution of a mysterious crime is not a detective story. In other words, the idea of genre reflects a long-established framework of a specific social or cultural construct. Since genres build on and uses particular schematic structures, it is worth noting some of the factors that may be relevant to distinguish them. One way of making a distinction among genres is to set a number of criteria. On the basis of the work on genres by the GRAL group at the University of Barcelona, we propose a macro-genre taxonomy based on certain fundamental differences that should be noted: i) co-occurrence of action plus text, ii) presence or non-presence of action (or narrative), iii) speed of action, iv) degree of abstractness, and v) clarity and speed of delivery. An example will clarify this classification. In the detective story, for instance, we can argue that the speed of action is quicker than in an animal documentary. Further, we can argue that clarity of the delivery is a more distinctive feature of the animal documentary than of the detective story. Apart from these factors, other aspects affecting how well a genre is received have to do with learners’ concrete characters: their engagement, preferences, perception of difficulty, and anxiety.

All considered, genre can be expected to have a strong influence on learning since different genres may create different conditions for learning. For example, some genres may be better at drawing attention to both form and meaning (e.g. where viewers may be able to coordinate the processing of images with the processing of subtitles), whilst other may be better at decomplexifying the task of viewing by bringing images, soundtracks, and subtitles to reinforce the message. We can assume that focused attention (reinforced by text and image,

for example) is easier to process than divided attention (where images and subtitles do not correspond or are even in contradiction with each other). We shall come back to this specific issue on genres in the discussion section.

Having explored the state of the art, the following section will deal with a more detailed description of the main theoretical models involved in this research.

2.5. Conceptual framework

By providing a wide scope of the dual coding model and the cognitive model of multimedia learning, we will attempt to shed some light on the cognitive theories related to the use of subtitles in second or foreign language acquisition.

2.6. Cognitive theory of multimedia learning

The cognitive theory of multimedia learning (CTMML) is a cognitive-affective theory of learning popularised by an American psychology professor, Richard E. Mayer (1997). Mayer's theoretical foundation for the CTML draws from several cognitive theories. As for this context is concerned, the basic premises of the CTMML relies on three main theories or models: (a) the dual coding theory or dual processing (Paivio, 1986; Clark & Paivio, 1991), (b) the cognitive load theory or limited working memory capacity (Baddeley, 1992; Mousavi, Low, & Sweller, 1995), and (c) the notion of meaningful learning, i.e., learners are active agents in their learning and select relevant information so as to build generative learning (Mayer, 1999; Wittrock, 1990).

To demonstrate the dual modality effect (Paivio, 1986) Mayer and Moreno (1998) tested learners' understanding of a short animation. In this experiment, learners were split into two groups: one group of participants was presented a simultaneous narration (audio and images) while the other group received visual stimulation, i.e., on-screen text and images. Participants of the former group performed better on the three comprehension tests –retention, matching images and names, and transfer– than learners receiving animation and on-screen text (Mayer & Moreno, 1998). In other words, in the group receiving visual stimulation the visual working memory became overloaded whilst in the simultaneous narration group the fact that the information was delivered via two different channels (auditory and narration sounds) might have helped learners to better comprehend the information.

Another important aspect of the CTML is what is known as the multimedia principle or multimedia effect. This principle holds that the most effective learning environments are those that combine verbal and nonverbal representations of the knowledge using mixed-modality representations (a combination of illustrations, charts, animations, videos, and words or text). This is also known as the ‘tell and show me’ idea; people’s understanding is aided by adding visual representations of knowledge to complement the verbal ones. Such reading views interactivity as a design feature to be manipulated to enhance ‘deep cognitive processing’ (Mayer & Moreno, 2010).

In order to delve into Mayer’s views of learning it is vital to refer to his distinction between information acquisition and knowledge construction (2001). In the information acquisition approach knowledge, i.e., information that students are expected to assimilate, flows from the instructor to the learners. Such view is based on a behaviourist approach to learning and, therefore, such learning environments are described as non-interactive. In the knowledge construction approach learners are ‘sense-makers’ who incorporate new content knowledge (information) with prior existing knowledge. The ultimate goal of this approach is to guide learners through the process of making sense of the instructional materials (Mayer, 2001). In this perspective, teachers are seen as resources providers, prompting learners to build associations between corresponding aspects of verbal and nonverbal representations. This is a much more interactive approach as it promotes constructivism (Piaget, 1952).

Designed based on principles of the cognitive theory of multimedia learning (CTML) Moreno proposes what is known as a cognitive-affective theory of learning with media (CATLM; Moreno, 2005a). As the name of the model suggests (cognitive-affective), Moreno attempted to combine the role of affective factors with the CTML. In other words, CATLM is just an expansion of the CTML (Moreno, 2007). If we look at the CATLM model (see Figure 1) arrows from long-term memory point back to the cognitive processes of integrating, organizing, and retrieving, indicating the role of motivation, affect, and metacognition in initiating, supporting, and controlling cognitive processing during learning.

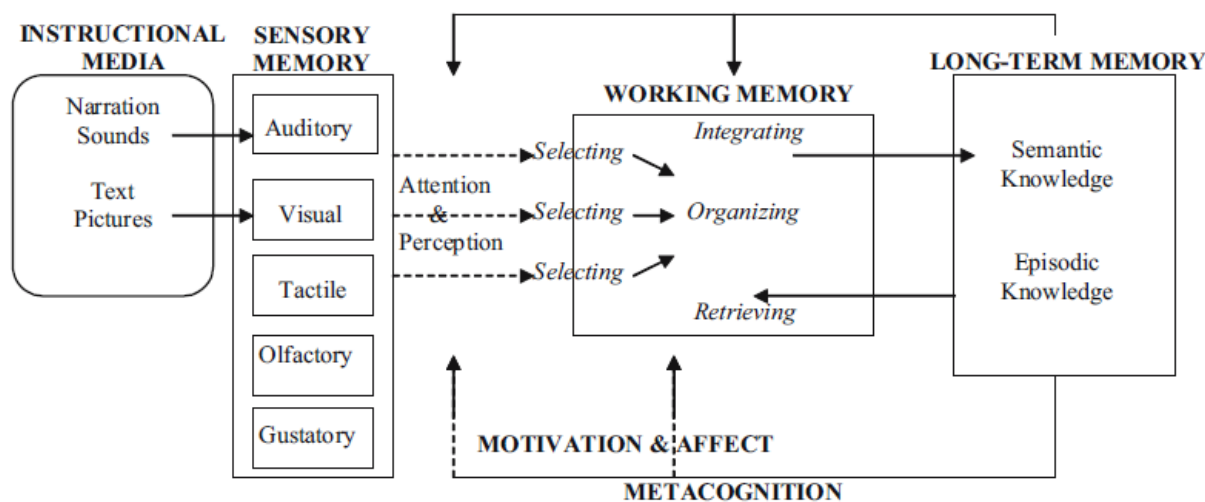


Figure 1. A cognitive-affective model of learning with media (Mayer & Moreno, 2007)

2.6. Dual coding theory

Dual coding theory (DCT) describes, in part, how the brain processes new information, i.e., ‘comprehensible input’, in Krashen’s terms (1981). Paivio argues that the human cognitive system has limited attention, and so it cannot process all sort of inputs, but rather selects key parts of auditory/visual messages and focuses on the most relevant information. Paivio’s dual coding posits that cognition consists of two distinct subsystems, one verbal (i.e., *logogens*) and one nonverbal (i.e., *imagens*) (Clark & Paivio, 1991). The verbal code is in charge of dealing with the language, and the nonverbal code is specialized in dealing with non-linguistic objects and events. Paivio claims that these information resources or ‘memory storages’ are functionally independent, but the two subsystems also cooperate and interact. There are two types of interactions: associative and referential. The former occurs if the verbal or the visual systems interact; the latter refers to an interaction between the verbal and nonverbal codes. Provided this, Paivio argues that “combining pictures, mental imagery, and verbal elaboration could be an effective method in promoting understanding and learning from text by students ranging from grade school to university level” (1991: 163). In other words, learning occurs more effectively through dual coding than by visual or verbal coding alone.

There is considerable evidence based on dual coding confirming that memory has a significant role in describing how learning occurs. DCT is particularly linked to memory in its educational implications for “memory remains crucial (...) [as] the basis of all knowledge and thought” (2006: 4). Paivio claims that imagery has powerful mnemonic effects: it

facilitates recall of verbal material because when a word evokes an associated image (either spontaneously, or through deliberate effort) two separate but linked memory traces are laid down, one in each of the memory stores. In this content, the chances that a memory will be retained and retrieved are much greater if it is stored in two distinct functional locations rather than in just one.

A DCT cognitive developmental hypothesis draws its theory from DCT and describes the different developmental stages of language learning. According to Paivio, the “Cognitive growth according to DCT is based on multiple learning processes of observation, classical conditioning, operant learning, and imitation” (1991: 7). Paivio states that nonverbal cognitive representations underpin the basis for later cognitive skills that include language acquisition (1991: 8). In this regard, humans’ first responses show recognition of people or objects before language appears. Language emerges as humans start making associations between words (*logogens*) and objects or events (*imagens*). At this stage, verbal responses appear in terms of ‘mimicry of speech sounds’, and, finally, at the end of the process, verbal associative skills related to comprehension prompt the production of sequences of two or more words (Paivio, 1991). Research has long supported these developmental stages (Mooser and Bregman, 1973; Strømnes, 2006).

3. The current study

3.1. Justification and objectives

Consistent with the studies referred to in section 2, the present research is intended to bridge the gaps in two dimensions of subtitles. First, and to the best of our knowledge, no previous studies have specifically explored whether there is variation in learning gains across different genres. Secondly, no previous studies have analysed the way EFL students learn MWEs through a combination of different videos with subtitles. Hence, the outcome of this study will contribute to existing knowledge of the acquisition of formulaic language using subtitles. In this context, this research might also help to improve multimedia task design in the EFL curriculum.

This project has a dual focus: on one hand, it will explore how L2 input enhancement might facilitate the acquisition of English MWEs. On the other hand, it will try to shed light on whether different genres might provide a better recall of learnt items, it will also compare the different gains that might be derived from different genres.

3.2. Research questions and hypotheses

The two main research questions and hypotheses are as follows:

RQ.1: Do different genres affect the acquisition of MWEs in English? If so, which genre is more appropriate to foster such type of formulaic language?

Hypothesis RQ.1: To our knowledge, this is the first empirical study to address this question. Thus, the null hypothesis is assumed.

RQ.2: What is the effect of input enhancement of target MWEs? Are there differences in the target gains between the groups (enhancement versus non-enhancement)?

Hypothesis RQ.2: Based on previous studies (Winke et al., 2010; Montero-Perez et al., 2013, 2015) it is predicted here that the input enhancement of target MWEs will have a limited but positive effect on learning gains.

4. Methodology

4.1. Participants

A total of 40 subjects, 18 women and 22 men, volunteered to participate in this study ($N=40$). Research participants were recruited using different social networks and also through professional and personal contacts. Participants' ages range from 15 to 52 years old, with a mean age of 27. Given the characteristics of this study, which required upper-intermediate to advanced learners, highly trained and highly advanced English language learners participated in the experiment: half of the participants reported to hold a Master's degree, 6 a Bachelor's diploma, and 4 subjects were PhD students. The remaining 10 were attending high school or doing professional training. Participants were all either native Spanish monolingual speakers or Spanish/Asturian, Basque, Catalan or Galician bilinguals. Participants were asked to self-report their English proficiency level in terms of the Common European Framework of Reference (CEFR). All participants self-rated their English language proficiency equal or higher than B1 (intermediate). Participants varied in English-language level from intermediate or upper-intermediate (B1, B2) to advanced (C1). All participants had studied English for a minimum of 5 year. 25 participants (a 60% of the sample) reported to have been studying English for 10 years or more.

Participants were told that they were going to take part in a research dealing with the use of TV series in learning English as a Foreign Language (EFL). Participants were not informed of the real goals of the experiment or the specific linguistic items that were going to be analysed. Research participants were divided into two groups: 20 subjects received the treatment, i.e., experimental group, whereas the remaining 20 participants served as a control group. Extreme care was taken to randomly include the same number of subjects in the experimental and control groups. In Table 1, relevant data collected from the participants is presented.

Table 1

Relevant information about participants

Feature	N° of participants
Gender:	
Male	22
Female	18
Total	40
Language	
Monolingual	14
Bilingual	26
Total	40
Level of education	
PhD	4
Masters	20
Bachelors	6
Professional training	6
High school	4
Total	40
Number of years studying English	
More than 5 years	15
More than 10 years	25
Total	40
Self-reported proficiency	
Intermediate	4
Upper intermediate	24
Advanced	12
Total	40

4.2. Research design

This is a within-subject, repeated measures design in which each participant had to watch four different subtitled clips in English and answer some questions based on the viewings. Prior to the tasks, participants read and signed a consent form informing about the aims of the

research and the procedures of the experiment (see Appendix B). Participants were given the opportunity to raise questions and withdraw from participation. Research participants were randomly assigned to either the experimental group or the control group. Both groups were tested before and after the one-week treatment period, and a delayed post-test was conducted three weeks later. At pre-test, learners took a working memory test, a vocabulary test, and they also completed a meaning recognition test that included the target expressions of the study. Testing included three distinct tests for each of the clips viewings: a video comprehension questionnaire, a word recognition test, and a meaning recognition test. The treatment involved two sessions of 30-45 minutes each (aprox.), which were carried out in successive two weeks.

Participants were exposed to a total of four short subtitled clips in English. The 4-minute videos were displayed in the original version with subtitles in English. After each of the viewings, participants were administered three different tests in order to analyse the effects of subtitles upon overall content comprehension as well as informal MWE(s) recognition and learning. The three tasks designed are as follows (in order of testing): i) overall comprehension, ii) form recognition, and iii) meaning recognition. More information on this will be provided in the Procedures section 4.4. Figure 2 illustrates the research design.

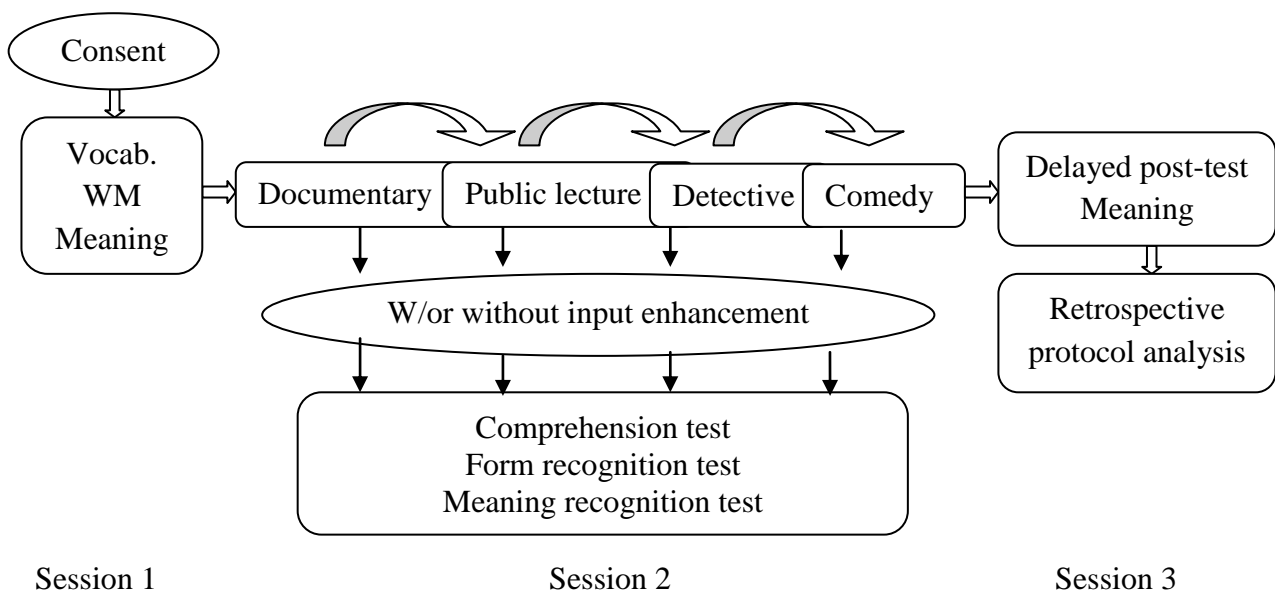


Figure 2. Research design

4.3. Instruments

4.3.1. Pilot study

The experiment was first piloted in order to evaluate feasibility, time, materials and methods, adverse events, and improve upon the study design prior to performance of the experiment. Five participants were recruited to conduct a small-scale rehearsal of the larger design. Subjects had an English proficiency level and language-learning background similar to the research participants. The researcher made minor changes to the initial design so as to refine it.²

4.3.2. Pre-test

As pre-test, three different tests were administered: a vocabulary size test, a working memory test, and the meaning recognition test (the latter is to be discussed in the Post-test section 4.3.3.).

Recent studies have shown a significant relationship between vocabulary size and overall proficiency (Meara & Milton, 2003; Meara, 2005; Milton, 2006, 2007). Hence, as an indicator of language proficiency, a receptive vocabulary size test was administered to estimate participants' overall proficiency in English. The Yes/No vocabulary test available at the Lognostics website was used. The Yes/No test measures receptive vocabulary knowledge by presenting a set of words, one at a time, in a context-free environment. This computer-based test covers a wide range of vocabulary ability (0-10K words). According to the authors, the test is more reliable with learners in the middle and upper levels of proficiency (Meara & Miralpeix, 2015). A manual to interpret the scoring of the Yes/No test is downloadable from the Lognostics website. The Yes/No scores are primarily used to assign a participant to a proficiency level. A potential limitation of this test is that there is no proven equivalence between its scores and the CEFR levels. Yet, there is a correspondence between vocabulary size and proficiency level, i.e., the higher the score, the higher the proficiency. Table 2 shows the scoring of the Yes/No test as the authors put it as well as an informal unproven correspondence to the CEFR levels for comparison's interest. Nevertheless, this is the only available reference.

² No major changes had to be made to the original design.

Table 2

Vocabulary Size and the CEFR Framework

Word range (maximum 10.000)	V_Yes/No guidelines	CEFR levels
6,000 to 10,000	Good for non-native speakers	C1, C2
3,500 to 6,000	Intermediate level	B1, B2
2,000 to 3,500	Competent beginner	A1, A2
Below 2,500 words	Unreliable	-

Note. Adapted from Meara & Miralpeix (2015)

In order to assess the working memory (WM) of the participants, a reading span task was administered. Studies in SLA typically use this type of task in order to assess both the executive control and storage component of memory (references). The reading span task used was developed by the GRAL research group (Gilabert & Muñoz, 2010).

The automated reading span task was used to test both reading span (RS) and letter span (LS). The LS task is a simple memory task that taps into the information storage aspect of working memory (WM) without requiring dynamic information processing. The reading span task (RS) is a complex memory task that looks into both the storage and the dynamic information manipulation aspects of WM. In the RS task participants are presented a series of to-be-remembered letters as well as a series of sentences that serves as a distractor task for it requires information processing (i.e., by deciding on the plausibility of a sentence). In other words, participants are asked to rate sentences and memorise letters simultaneously. Participants could take the task in Catalan or Spanish. For the purpose of this research, we will only be reporting on the results of the reading span as retrieving the results of the letter span is a tricky task that we had no time for.

4.3.3. *Post-test*

To collect data, a web-based questionnaire in English was designed using Google forms. In order to ensure randomization, eight different versions of the questionnaire were distributed. The assets and limitations of conducting web-based research have been widely discussed (Gosling et al., 2004; Dörnyei, 2007; Denissen et al., 2010; Dewaele & Wei, 2013). The aims

of this research showed more advantages than disadvantages to the use of online questionnaires –especially providing that the researcher monitored the pre-test and testing sessions.

The questionnaire consisted of seven different sections. The first section elicited information on the sociobiographic details of the respondents as well as relevant linguistic background information (e.g., number of years studying English). Specific instructions on the experiment were given in section two. Sections three, four, five, and six corresponded to the main testing, i.e., the viewings. Each genre video was attached to its specific section as well as the immediate tasks to be taken after each of the viewings. The last section enclosed a blank space for participants to leave their e-mail details (if interested in receiving the results) as well as a thank you note for their participation.

4.3.3.1. Genre videos

According to Talaván, “The sources of the authentic video clips to be used [for enhancing learning] are short sequences taken from English speaking sitcoms, soaps, series, films, short films, TV commercials, or documentaries” (2006: 45). Videos for this experiment varied in its format as well as in genre. The video clips to be viewed consisted of a selection of American and British TV series and programmes. These videos had been selected by the GRAL group, and they were chosen because they were representative of four distinct (and radically different) and easily recognisable generic characteristics. Two American TV series were selected: ‘The Mentalist’, labelled here as a sort of detective story, and ‘Friends’, a comedy often used in this area of research (Talaván, 2006; Frumuselu et al., 2015; Shabani & Pasha, 2015). As for the documentary, a BBC animal documentary by David Attenborough (*Life on Earth, Life*) was selected. The fourth video was a Ted Talk given by Sheena Iyengar (*The Art of Choosing, 2010*).

The four videos clips were transcribed and analysed in search for potential multi-words expressions. Different sources were checked to control for the level of proficiency of the MWEs: the Cambridge International Dictionary of Idioms, the English Vocabulary Profile (both British and American English versions), and the Oxford Idioms, and Oxford Phrasal Verbs Dictionary for Learners of English. Aegisub, a free advanced subtitle editing program, was used to create, edit the subtitles, and highlight the target MWE(s). In the case of the

experimental group, a total of 5 (out of 8) target MWEs were enhanced in order to harness the benefit of subtitles. The highlighting was removed from the target MWEs for the control group.

4.3.3.2. Comprehension test

Four comprehension tests were administered: one for each video genre. The overall comprehension tasks were taken immediately after each of the viewings. Participants were instructed to watch the videos just once. Each test consisted of two multiple choice questions. These tests were designed by the GRAL research group at the University of Barcelona. The GRAL group created these tests as part of a major project exploring the benefits of using subtitles to support English learning. An example of a comprehension test item is shown in Figure 3 (see Appendix A).

4.3.3.3. Word form recognition and meaning recognition tests

Both the word and meaning recognitions tests were modified versions of the tests designed by the GRAL group. These tests were designed and adapted from Rodgers and Webb (2017).

Word form recognition followed the comprehension test. The word form recognition test consisted of 10 Yes/No questions. Out of the ten items, five corresponded to target multi-word expressions (MWEs) and five were added as distracters. An example of a word form recognition item is shown in Figure 4 (see Appendix A).

Meaning recognition followed the word form recognition test. The meaning recognition test consisted of eight multiple choice questions. Learning was measured by checking whether the learners could recognise the meaning of the target MWEs among different given options. Participants were asked to provide the most accurate translation (in Spanish) for eight MWEs that had appeared in the videos. Three of the eight MWEs were distracters in the sense that these expressions corresponded to lower levels of proficiency (below B1). The five target MWEs corresponded to either B2 or C1, i.e., upper-intermediate and advanced levels of English. The logic of choosing the different options in the multiple choice test is explained in Rodgers and Webb (2017). An example of a meaning recognition item is shown in Figure 5 (see Appendix A).

4.3.4. Delayed post-test

An online delayed post-test was administered three weeks after the testing. In the delayed post-test participants were asked to re-take the meaning recognition test in order to measure their learning gains. The delayed post-test was delivered online due to time and space constraints – research participants reside in different areas of Spain.

4.3.5. Retrospective protocol analysis

The retrospective protocol analysis questionnaire was created to gather some qualitative information on the respondents' interests, engagement, and perceptions on the experiment. The questionnaire was a slightly modified version of a questionnaire designed by the GRAL group at the University of Barcelona. The questionnaire was completed after the participants once the participants finished the post-delayed test. It enclosed eight general and more specific questions in English on respondents' interests, perception of difficulty and watching practices (see Appendix D). Participants could answer in Spanish or English.

4.4. Procedure

The study was conducted over a 6-week period. Data collection (pre-test, post-test, and delayed post-delayed) took place in three sessions of 30-45 minutes each (approx.). Sessions took place in an isolated, quiet area to elicit natural behaviour and avoid setbacks. Participants sat facing the researcher and, after a few minutes of small talk, they signed the consent form and were introduced to the different tests designed to collect data in this experiment. Participants were encouraged to ask questions, and then they were asked to take the tasks. In order to convey the right messages and to control for possible language hitches, the researcher used Spanish and English to deliver the instructions. The pre-test was administered on the first week (Session 1). At pre-test participants took the vocabulary size test, the working memory test, and the meaning recognition test. Participants were tested on the following week (Session 2). The researcher monitored the pre-test and the post-test sessions. An online version of the delayed post-test was designed and distributed online three weeks after the post-test (Session 3). Participants were also asked to fill in a retrospective protocol analysis so as to gather relevant qualitative data.

5. Statistical analysis

Data obtained from the tests were analysed using R Software for Statistical Analysis. Two different types of tests were used in each of our research questions depending on the kind of distribution that was presented by the data, i.e., whether it was normally distributed or not by means of both Kolmogorov-Smirnov and Shapiro-Wilk tests of normality. The statistical analyses used in this study include descriptive statistics, correlations, Friedman tests, and Wilcoxon tests.

6. Results

The first research question asked whether genre had an effect on the learning gains of English multi-word expressions.

Prior to evaluating whether genre had an effect on the gains data was tested for normality of distribution. A Shapiro-Wilk normality test determined that the variable ‘gains’ was not normally distributed (.000) and it was decided to use non-parametric tests. The initial analysis examined the overall gains of the treatment. Overall learning gains were operationalized as positive raw changes from pre-test to delayed post-test. A non-parametric Friedman test of differences among repeated measures was conducted to assess the overall gains from pre-test to delayed post-test. Results rendered a Chi-square value of 10.567 which was significant ($p=0.014$). Results are shown in Table 3³.

Table 3

Gains from Pre-Test to Delayed Post-test

Test Statistics ^a	
N	40
Chi-square	10.567
Df	3
Asym. Sig.	0.014

a. Friedman Test

A more fine-grained analysis was conducted so as to provide a full report of the gains from pre-test to post-test, from post-test to delayed post-test, and from pre-test to delayed post-test

³ Summary tables of the Friedman tests performed from pre-test to post-test and from post-test to delayed post-test are available in Appendix A.

(i.e., overall gains). Descriptive statistics by genres are shown in Table 6. As indicated by the means, both the control and experimental groups showed learning gains in the four genres. In the case of the experimental group, the animal documentary ($M=1.60$, $SD=.82^4$) and comedy genres ($M=1.40$, $SD=1.00$) showed greater gains than the detective story ($M=1.00$, $SD=.79$). As for the control group, results revealed more gains in the detective story ($M=.80$, $SD=.83$) and the documentary ($M=.75$, $SD=.64$) than in the comedy ($M=.65$, $SD=.67$). Interestingly though, the genre with the least gains corresponded to the public lecture in the two groups ($M=.80$ and $M=.50$, respectively). A plot of means summarised the overall gains (Figure 6). Plots of means from pre-test to post-test and from post-test to delayed post-test are also available (see Appendix A).

Table 6

Descriptive Statistics for Gains

Genre	Group	Pre to post-test				Post to delayed post-test				Pre to delayed post-test			
		M	SD	Max	Min	M	SD	Max	Min	M	SD	Max	Min
G1	Control	0.25	0.44	1	0	0.50	0.51	1	0	0.75	0.64	2	0
	Exp.	0.55	0.51	1	0	1.05	0.60	2	0	1.60	0.82	3	0
G2	Control	0.10	0.31	1	0	0.40	0.50	1	0	0.50	0.69	2	0
	Exp.	0.25	0.44	1	0	0.55	0.60	2	0	0.80	0.89	3	0
G3	Control	0.20	0.41	1	0	0.60	0.60	2	0	0.80	0.83	2	0
	Exp.	0.35	0.49	1	0	0.65	0.59	2	0	1.00	0.79	2	0
G4	Control	0.30	0.47	1	0	0.35	0.49	1	0	0.65	0.67	2	0
	Exp.	0.35	0.59	2	0	1.05	0.76	2	0	1.40	1.00	4	0

Note. G1 is the animal documentary, G2 the public lecture, G3 the detective story, and G4 is the comedy

⁴ The high standard deviations (SD) values are due to the nature of the measurement.

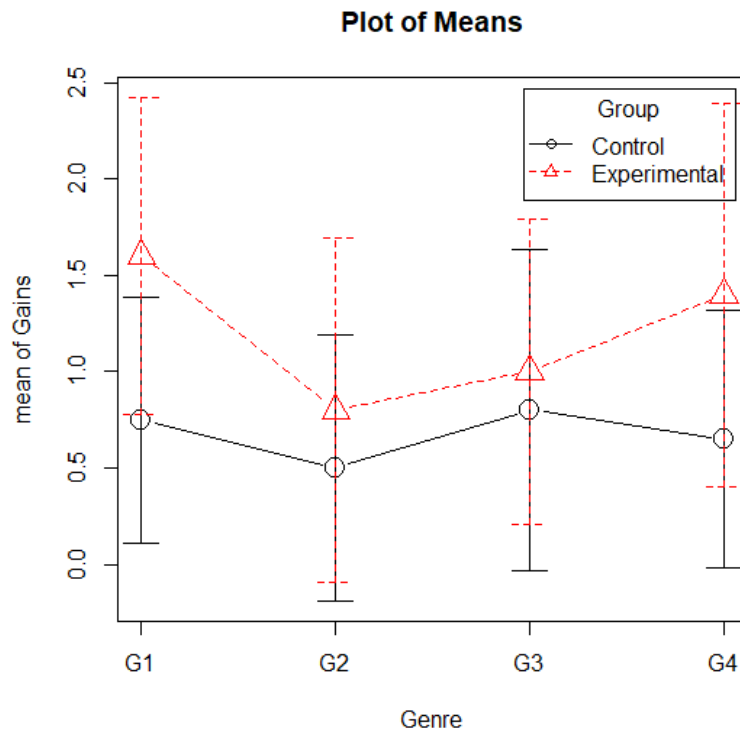


Figure 6. Plot of means for overall gains

Four one-sample sign-ranked Wilcoxon test were run so as to assess the gains by genres. Our null hypothesis was the median to be 0 whilst the alternative asked the median to be higher than 0⁵. Results showed that three of the four genres had an effect on the learning gains (see Table 7). The genre showing a greater impact was the documentary ($Z=104$, $p=.000$) followed by the comedy ($Z=731.5$, $p=.0009$), and the detective story ($Z=595$, $p=.004$). As for the public lecture, results showed that there was a non-significant effect ($Z=422.5$, $p=.2704$). A plot of means illustrates these findings (Figure 9).

⁵ Our null hypothesis is equivalent to say that at least half of the participants did not learn.

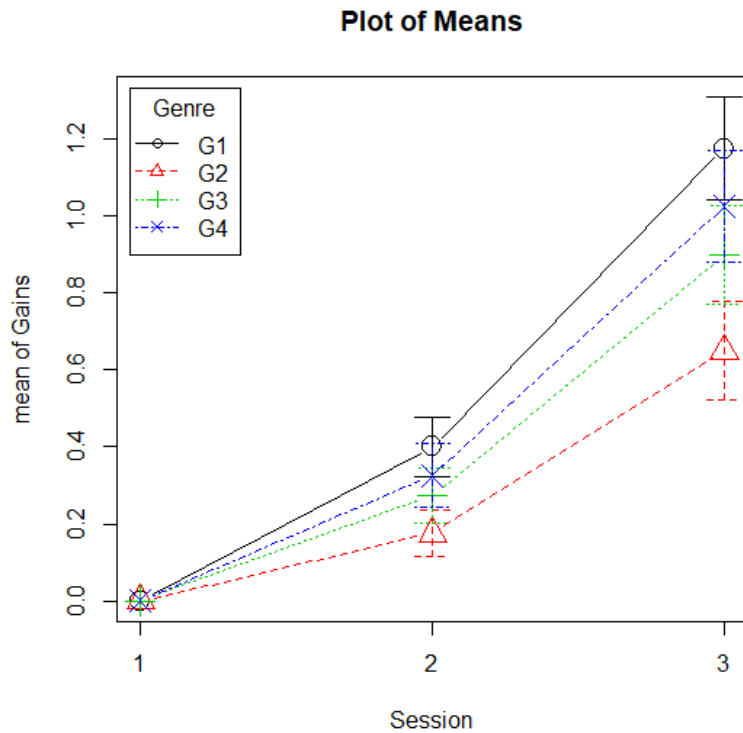


Figure 9. Plot of means showing the gains by genres

Note. X-axis: Session 1 refers to the pre-test, 2 to the post-test, and 3 to the delayed post-test
 Y-axis: G1 is the documentary, G2 the public lecture, G3 the detective story, and G4 the comedy

Table 7

<i>Tests Statistics^b</i>				
	Documentary	Public lecture	Detective story	Comedy
Z	704	422.5	595	731.5
Asym. Sig. (1-tailed)	.000	.270	.004	.000

a. Based on positive ranks

b. Wilcoxon Signed Ranks Tests

Further analyses, addressing the relationship between proficiency (as measured by the vocabulary size test) and learning gains from pre-test to post-test and from pre-test to delayed post-test, examined whether these two variables were correlated. Two Spearman correlations revealed non-significant correlations between proficiency and gains at post-test ($r= .13737$, $p=0,070$) nor between proficiency and overall gains ($r= .12687$, $p=0,239$).

In order to further explore research question 1, a set of correlations were run to explore the role of working memory (WM). Having provided that each pair of variables was normally distributed, a Pearson correlation revealed a non-significant correlation between working memory and proficiency (as measured by the Yes/No vocabulary size test) ($r= 0.98112$, $p=0,332$). Furthermore, two Spearman correlations were also run and determined that there were non-significant correlations between proficiency and gains at post-test ($r= 9166$, $p=0,388$) and overall gains (at the delayed post-test) ($r=10232$, $p=0.805$).

Regarding the second research question, which asked if there were differences in the target MWEs gains between the groups, we first checked if the measures were normally distributed. As shown in the box plot (Figure 10), the measures presented certain asymmetry and so a Shapiro-Wilks test by groups was run, confirming the lack of normality of distribution (values of $p= .012$ and $p= .021$, control and experimental, respectively). We can appreciate, however, a homogeneous dispersion of the data. Since the data deviate from a normal distribution, a Levene’s test was conducted so as to assess the homogeneity of the variances; a p-value of $.75$ was obtained.

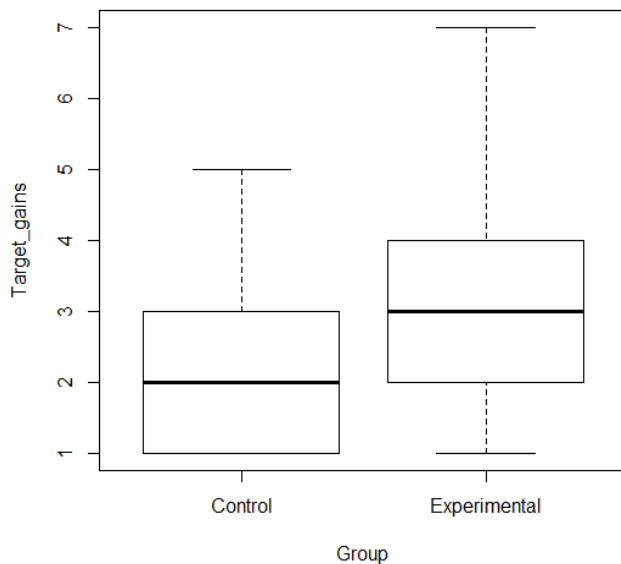


Figure 10. Box plots showing the dispersion of the target gains scores by groups

An independent samples t-test was run in order to compare the target gains in the control and experimental groups. A p-value of $.057$ indicates that, given the lack of normality of the distribution, non-parametric statistical tests had to be run. An unpaired two-sample Wilcoxon

test (also known as Mann-Whitney test) indicated that target gains were greater for the experimental group ($M^6=3.2$) than for the control group ($M=2.3$), ($U=134$, $p=.073$). Despite not having reached the level of significance ($p<.05$), there seems to be a strong trend towards significance ($p=.073$). The standard deviation (1.43) was obtained by taking the average of the deviations within the two groups, given the result of the Levene's test.

6.1. Results of the retrospective protocol analysis

A secondary data source for this study was the retrospective protocol analysis that participants were asked to complete immediately after they had finished the post-delayed test. The questionnaire was set up so the learners could provide more details to supplement the quantitative collected data. Many learners commented in general on whether they thought they learnt from the experiment, or which aspects of the specific video or genre they found most helpful. We will not report on general themes but on the issues that helped to shed some light on specific aspects regarding the results of the experiment.

Concerning the watching habits (*Do you watch with the soundtrack in English but the subtitles in Spanish? Do you sometimes have English-English?*), 50 % of participants claimed to watch English series with English subtitles, 25% English with subtitles in Spanish, and the remaining 25% watch English series without subtitles. A majority of participants (65%) rated the public lecture Ted Talk as the most difficult genre. More than a half of the participants (60 %) rated the comedy 'Friends' as the easiest to understand. Furthermore, learners showed an overwhelming positive evaluation of the experiment. To the question: *Do you think you learnt from watching those videos?* 100% of participants responded affirmatively. Relevant examples from 3 participants follow:

- (1) Yes I think I learnt a few things from those videos, like vocabulary and expressions.
- (2) Yes, I learnt the meaning of lightheaded. I think it was in Friends.
- (3) I think I learnt some slang and I also noticed some words.

In the following section findings from this research will be discussed as well as related to previous studies. Furthermore, some possible explanations and implications of the results will be provided in an attempt to further understand the findings.

⁶ Means instead of medians are reported here due to the nature of the measurement.

7. Discussion

Our main concern in this study was the issue of genre in relation to the acquisition of English multi-word expressions (MWEs).

To answer the first question (*Do different genres affect the acquisition of MWEs in English? If so, which genre is more appropriate to foster such type of formulaic language?*), the pre- and post- and post-delayed tests' scores were analysed. Results suggest that there is a significant effect of genre on the overall gains⁷ of MWEs. Thus, we reject the null hypothesis in favour of the alternative one. To our knowledge, no studies have been conducted on the effect of genre in learning MWEs from subtitled videos. Nevertheless, that gains varied from one genre to another was not unexpected. Our results suggest that under certain genres EFL learners must have detected and at least superficially processed more formulaic expressions than in others. For instance, our findings revealed that participants in the experimental showed more gains in the animal documentary. In this light, documentaries appear to create a relaxing learning environment in which the EFL learners are supported by a number of factors: the co-occurrence of action and text, the smooth speed of action or the clarity and speed of the delivery. This in turn creates better condition for noticing and for intake, which are the two stages in the early stages of acquisition (Schmidt, 1990). A more detailed analysis could look into these specific stages. Yet, we suggest that the distinct features of the language input in the documentary allowed for a more conscious information-gathering that may have aided chunking processing. In the case of experimental group, the documentary occupied the first position in terms of MWEs gains ($M=1.60$) whilst it occupied the second position for the control group ($M=.65$). Overall, our findings suggest that both the documentary and the comedy genres appear to be the most appropriate genres to foster learning gains ($p=.000$).

Results also showed that, regardless of the treatment, the genre showing fewer gains was the public lecture ($p=.270$). Interestingly though, the public lecture was rated as the most difficult genre by 65 % of the participants ($N=40$). In the absence of a complete narrative, as it was the case of the public lecture, the degree of abstractness might support the participants' higher perception of complexity. Furthermore, in the public lecture there was a lack of co-occurrence (image and text); this is something that might have also affected the learners' attention ability. In other words, in the case of the public lecture, attentional resources may

⁷ By overall gains we refer to target and non-target multi-word expressions (MWEs).

have been dispersed between and anxiety may have increased making learning gains too costly. In this regards, one learner expressed that “the Ted Talk was (...) the most difficult to focus on as there were no images but just the woman speaking” (Subject 23, Retrospective Protocol Analysis).

As for the comedy genre, it showed greater gains in the experimental group ($M=1.40$) than in the control group ($M=0.65$). Despite group differences, the comedy genre showed to be as appropriate as the comedy genre to enhance learning ($p=.000$). At this point it is worth noting that 65% of the learners in the experimental group (13 out of 20) rated the comedy as the most engaging of all genres. Similarly, almost half of the participants in the experimental group (8 subjects) also rated it as the most entertaining genre. In addition, some participants claimed to have watched the American TV series ‘Friends’ before. As for the engagement scores, there was, however, a greater deal in variation in the participants’ responses in the other three genres. This individual variation in the engagement ratings could explain some of the differences in the overall learning gains across genres. It is unfortunate that we had no time to look into the results of the form recognition tests so as to clarify whether participants noticed more forms under some genres than in others. Subsequent studies should include those results.

Ultimately, results from the detective story indicated that this genre also aided the learning of MWEs ($p=.004$) but in a lesser extent than the documentary and the comedy genres. We can argue, for instance, that the speed of action was quicker in this genre and so learners’ processing and intake might have suffered more than in the documentary or comedy genres. Furthermore, a greater number of characters appeared in the video which is something that might have also impacted on the learners’ attentional resources.

An interesting finding has to do with the role of proficiency (as measured by the self-ratings and the Yes/No vocabulary size test). Contrary to other studies (Danan, 2004; Mayer, 2014), proficiency level did not moderate the effect sizes of the overall learning gains. The higher-level proficiency learners in the present study had possibly impacted on the fact that proficiency did not have such an impact on the overall gains.

Broadly speaking, there are many factors that might have influenced the scores including learners’ interests in a specific genre, engagement, watching habits or the length of the

listening text. An additional possible explanation for our results may be explained by the impact of affective variables. For instance, if we examine the case of the comedy genre, we see that most participants reported it as being the easiest genre as well as the most engaging. In this regards, Rodgers and Webb (2017) also argue that there are a number of factors that are specific to viewing the videos: visual literacy of the participants, relationship of the images to the audio, and learners' interest in the text and video type. Additionally, learners' preferences towards a variety of English –either British or American accents or pronunciation in this experiment– might have also affected comprehension and learning gains. This was the case of a specific learner who argued so in the retrospective protocol analysis.

Concerning the second research question (*What is the effect of input enhancement of target MWEs? Are there differences in the target gains between the groups (enhancement versus non-enhancement?)*), we can see that, in line with findings from previous studies (Peters et al., 2016; Montero-Perez et al., 2013), our results also suggest that input enhancement may be playing a role in the process of learning. Despite significance was not attained ($p = .073$), there appears to be a strong trend pointing the fact that input enhancement of the target MWEs had a positive effect, which is evidenced by the mean number of MWEs learnt by the experimental group ($M=3.2$) in comparison to the control group ($M=2.3$). This suggests that the mapping of form to meaning was facilitated in the experimental group as input enhancement helped identify MWEs boundaries. In other words, input enhancement might have assisted these learners segment what otherwise be an incomprehensible string of words. As reported by other studies (Montero-Perez et al., 2013; Winke et al., 2013; Vanderplank, 1988), our findings also suggest that the use of input enhancement techniques result in a higher level of retention and recall of the target MWEs. This is evidenced by the feedback from the retrospective protocol analysis. In conclusion, we believe that our data show that input enhancement promoted noticing and so learners in the experimental group were able to unpack the speech in a more meaningful way. That is, input enhancement may have played a role in assisting these learners to remember and learn from the input presented.

9. Limitations

Notwithstanding the promising results, there are some limitations to this study that should be acknowledged. One of the main limitations in the present study, as in many others, is the sample size. In order to enhance the validity of this research a greater number of participants

is required. Another relevant aspect regarding the sample is the characteristics of the research participants. Future research will benefit from having a variety of participants in terms of their socioeconomic and linguistic background. As for the instruments, we must acknowledge that the Yes/No vocabulary size test poses certain limitations as it just measures receptive but not productive vocabulary knowledge. Additionally, the test does not have a certified equivalence to the CEFR levels. Another possible drawback is that participants self-rated their English competence. However, it is worth noting that most participants declared to possess a certified level of English. As for the videos used, the present research made use of four distinct genre videos of similar length. It is unclear whether findings are generalisable to other genres and videos of different lengths. Another flaw has to do with the fact that we did not have total control over each target MWEs, and we do not know if some were more difficult to learn than others. In this regard, we acknowledge that maybe the effects of genres was not such, but rather it was the effect of the difficulty of the target expressions contained in each genre. A more detailed follow-up study could look precisely at this issue.

Furthermore, a longer time of intervention and more than two exposures to subtitled video could be included in further research. It is unfortunate that this research was not able to include a more detailed data analysis due to the limited time-frame. As well as limitations pertaining to the methodology itself, a number of limitations arose due to the context of the research. In this experiment the researcher monitored the pre-test and post-test sessions but participants took the post-delayed test and the retrospective protocol analysis by themselves. As previously mentioned, web-research has obvious and inherent shortcomings. A specific flaw in our experiment has to do with the delayed post-test. Four participants claimed to have reviewed the videos and so this could entail certain problems regarding the interpretation of the results.

8. Conclusion and further research

The findings of this study make several contributions to the current research paradigm. First, this research has been an initial attempt to explore how different genres might affect the acquisition of English MWEs. Overall, the findings suggest that there is an effect of genre on the acquisition of MWEs through subtitles. More powerful results shall be expected provided a larger corpus of participants. Consequently, many more promising findings are to be found in future research.

In line with other studies this research has also found that input enhancement supports upper-level learners (Winke et al, 2013; Montero-Perez et al., 2013). All considered, the current research supports the idea that input enhancement in multimedia learning environments facilitates recognition of word forms and recall of MWEs meaning. At the same time, our results demonstrate that the presence of L2 subtitles can promote second language learning (Danan, 2004; Vanderplank, 2010; Talaván 2016). Additionally, the findings in this study seem to have certain pedagogical implications for EFL teachers. On one hand, EFL teachers may include appropriate audiovisual activities with target vocabulary forms in the curriculum. On the other hand, videos can be useful tools in the development of learner autonomy. One final observation emerged from the participants' perspective: they all showed positive attitudes toward learning from subtitled video.

Further research would benefit from examining both short and long-term effects of exposure and learning. Furthermore, a replication of this study using a variety of language measures and learners from different L1s may provide a more conclusive assessment of the effects of genre. A specific area that is yet to explore is the comparison between different languages. In conclusion, further analysis is required in order to uncover the impact of genre in the acquisition of English formulaic language.

In our concluding remark we would like to remind the reader that this study is a preliminary effort to empirically examine the effect of genre on the acquisition of formulaic language. This is certainly a challenging area of research that is yet to be explored.

Word count: 9250

Bibliography

Clark, J. & Paivio, A. (1991). Dual Coding Theory and Education. *Educational Psychology Review* 3, 149-210.

Danan, M. (1992). Reversed subtitling and dual coding theory: New directions for foreign language instruction. *Language Learning*, 42(4), 497-527.

Danan, M. (2004). Captioning and Subtitling: Undervalued Language Learning Strategies. *Meta: Translators' Journal*, 49(1), 67-77.

Díaz-Cintas, J. (2013). The technology turn in subtitling. In: Thelen, M and Lewandowska-Tomaszczyk, B, (eds.) *Translation and meaning: Part 9*, 119-132.

Gilabert R. & Muñoz C. (2010). Differences in Attainment and Performance in a Foreign Language: The Role of Working Memory Capacity. *International Journal of English Studies* 10(1), 19-42.

Kellner, D. (1980). Television Images, Codes and Messages. *Televisions* 7,(4).

Krashen, S.D. (1981). *Second Language Acquisition and Second Language Learning*. Oxford: Pergamon.

Mayer, R. (2014). *The Cambridge Handbook of Multimedia Learning*. Cambridge University Press.

Meara, P. & Miralpeix, I. (2015). V_YesNo v1.0. Retrieved from http://www.lognostics.co.uk/tools/V_YesNo/V_YesNo_Manual.pdf

Montero-Perez M., Peters E. & Desmet P. (2015). Enhancing Vocabulary Learning Through Captioned Video: An Eye-Tracking Study. *The Modern Language Journal*, 99(2).

Montero-Perez M, Van den Noortgate, W. & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System*, 41, 720–739.

Moreno, R., & Mayer, R. E. (1999a). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, 91, 358–368

Moreno, R., & Mayer, R. E. (2002). Verbal redundancy in multimedia learning: When reading helps listening. *Journal of Educational Psychology*, 94, 156–163.

Paivio, A. (1983). The empirical case for dual coding. In J. C. Yuille (Ed.), *Imagery, memory and cognition: Essays in honour of Allan Paivio*, 307–332. Hillsdale: Erlbaum.

Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University press.

Paivio, A. (1991). Mental representation in bilinguals. In A. G. Reynolds (Ed.), *Bilingualism, multiculturalism, and second language learning: The McGill conference in honour of Wallace E. Lambert*, 113–126. Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Peters E., Heynen E. & Puimège E. (2016). Learning vocabulary through audiovisual input: The differential effect of L1 subtitles and captions. *System*, 63, 134-148

Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.

Rodgers M. & Webb S. (2017). The Effects of Captions on EFL Learners' Comprehension of English-Language Television Programs. *Calico Journal*, (34)1.

Schmitt, N. (2004): Formulaic sequences: acquisition, processing and use. John Benjamins.

Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Sprenger, S. (2003). Fixed expressions and the production of idioms. *MPI Series in Psycholinguistics*, 21. Nijmegen, NL: Radboud University Nijmegen.

Sprenger, S. A., Levelt, W. J. M., & Kempen, G. (2006). Lexical access during the production of idiomatic phrases. *Journal of Memory and Language*, 54(2), 161-184.

Talaván, N. (2006). Using subtitles to enhance foreign language learning. *Porta Linguarum*, 6, 41-52.

Vanderplank, R. (1988). The value of teletext sub-titles in language learning. *English Language Teaching Journal*, 42(4), 272-281.

Vanderplank, R., (1990). Paying attention to the words: practical and theoretical problems in watching television programmes with unilingual subtitles. *System*, 18(2), 221-234.

Vanderplank R. (2016). Captioned media in foreign language learning: Subtitles for the deaf and hard-of-hearing as tools for language learning. London: Palgrave Macmillan UK.

Willmot, P., Bramhall, M., & Radley, K. (2012). Using digital video reporting to inspire and engage students. Retrieved from <http://www.raeng.org.uk>

Winke P., Gass S. & Sydorenko T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology*, 14, 65-86.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Wray, A. & Perkins, M. R. (2000). The functions of formulaic language: an integrated model, *Language & Communication*, 20, 1-28.

Appendix A: List of figures and tables

Why did the female bird leave in the end?

- a) She didn't like the male.
- b) She heard a person coming.
- c) Another male appeared.

Figure 3. Item #2 on the comprehension test for the BBC animal documentary

Did these expressions appear in the clip? Please choose 'yes' or 'no'. If you are not sure, choose 'no'.

- 5) To be a sweet tooth
- a) Yes
- b) No

Figure 4. Item #5 on the word recognition test for the sitcom 'Friends'

- 3) Have it your way
- a) *Hazlo a tu manera*
- b) *Haz tu camino*
- c) *Vete a tu ritmo*
- d) *Hacer senderismo*
- e) *No lo sé*

Figure 5. Item #3 on the meaning recognition test for the public lecture (a Ted Talk) by Sheena Iyengar

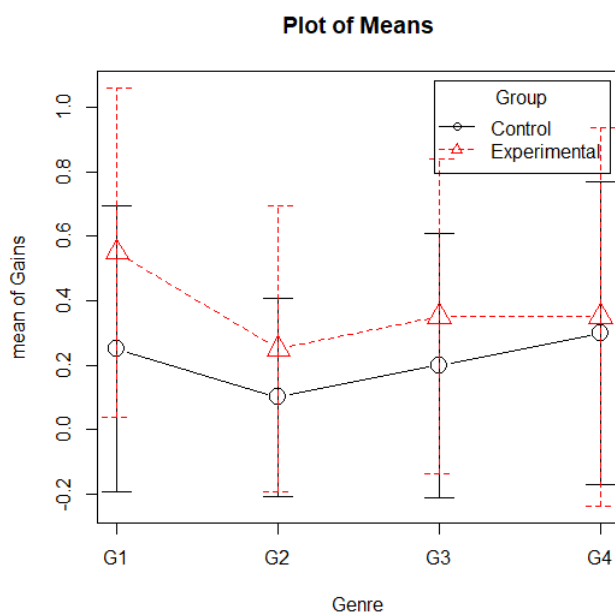


Figure 7. Plot of means from pre-test to post-test

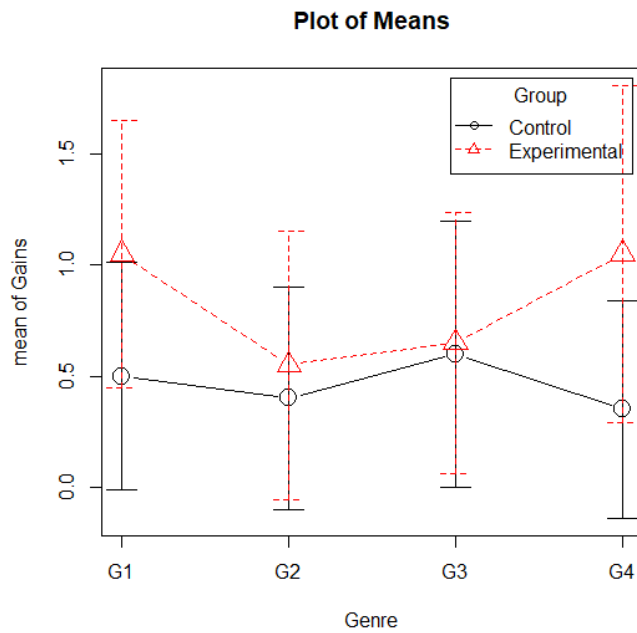


Figure 8. Plot of means from post-test to post-delayed test

Table 4
Gains from Pre-test to Post-test

Test Statistics ^a	
N	40
Chi-square	5.0204
Df	3
Asym. Sig.	0.1703

a. Friedman Test

Table 5
Gains from Post-test to Delayed Post-test

Test Statistics ^a	
N	40
Chi-square	4.0244
Df	3
Asym. Sig.	0.2588

a. Friedman Test

Appendix B: Consent form

TV series inside the EFL classroom

Description of the research

You are being invited to participate in a research study titled "TV series inside the EFL classroom".

This study is being conducted by Inés Sánchez de la Viña Rodríguez from the University of Barcelona. This research aims to investigate the potential benefits of watching audiovisuals in major aspects of learning English as a foreign language (EFL). The main objective of this research is to bridge the gap between teaching and learning English through multimedia.

This experiment consists of three sessions of 30-45 minutes (approx.).

The anonymity of information collected from research participants is guaranteed. Participation is completely voluntary.

Risks and discomforts

There are no known risks associated with this research.

Potential benefits

This research may help us to investigate the many benefits of implementing audiovisuals in the EFL classroom. Participants might also benefit from their participation on this experiment as it is aimed at practicing different English skills such as listening and overall content comprehension.

Protection of confidentiality

Your identity will not be revealed in any publication resulting from this study.

Voluntary participation

Participation in this research study is voluntary. You may refuse to participate or withdraw your participation from the study at any time.

Contact information

If you have any questions or concerns about this study or if any problems arise, please contact us: isanchro23@alumnes.ub.edu

Consent

I have read this consent form and have been given the opportunity to ask questions. I thereby agree to voluntarily participate in this research project.

Participant's signature _____ Date: _____