Effects of text, audio and learner control on text-sound association and cognitive load of EFL learners

by

Ana Maria Enciso Bernal

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2014 by the
Graduate Supervisory Committee:

Brian Nelson, Chair
Wilhelmina Savenye
Robert Atkinson

ARIZONA STATE UNIVERSITY

May 2014

ABSTRACT

This study investigated the effects of concurrent audio and equivalent onscreen text on the ability of learners of English as a foreign language (EFL) to form associations between textual and aural forms of target vocabulary words. The study also looked at the effects of learner control over an audio sequence on the association of textual and aural forms of target words. Attitudes towards experimental treatments and reported level of cognitive load were also examined in the context of a computer-based multimedia instructional program. A total of 200 college students took part in the study. Participants were randomly assigned to experimental conditions in a 2 x 3 factorial design with level of learner control (learner-controlled vs. not-learner-controlled) and format of presentation of information (audio + no text vs. audio + full text vs. audio + keyword text) as factors. The subjects completed a pretest, a posttest, cognitive load questions, and an attitude questionnaire. The results revealed the following findings: (a) groups in the audio + keyword text conditions outperformed those in the audio + no text and audio + full text conditions on text-sound association, (b) within the audio + keyword text conditions, the learner-controlled group outperformed the not-learner-controlled group on text-sound association, (c) within the learner-controlled conditions, the audio + keyword group outperformed the audio + no text and audio + full text groups on text-sound association, (d) a redundancy effect was not found for any treatment condition, and (e) overall, participants had positive attitudes towards the treatments. Implications, limitations, and future directions are discussed within the frameworks of cognitive load theory and cognitive theory of multimedia learning.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to Dr. Brian Nelson for the invaluable lessons, guidance, and support he provided during the time I spent at Arizona State University; his insight, patience, and critique of my work made this research project possible. I would also like to thank my other committee members, Dr. Robert Atkinson and Dr. Wilhelmina Savenye, for the priceless advice and encouragement they gave me throughout my studies at the program of Educational Technology. Special thanks to Dr. Samuel Green for being a willing teacher who was always there to give advice. In addition, thanks to the ACS-PUDI of the Autonomous University of Zacatecas for providing me with the conditions to conduct this study in the English language program. Also, thanks to Caroline Savio-Ramos for her important contribution to this work.

Completion of this project could not have been possible without the financial support of the Fulbright Program, the Autonomous University of Zacatecas, the Secretariat of Public Education in Mexico, and the Graduate and Professional Student Association at Arizona State University.

Finally, I wish to acknowledge my family and personal friends whose support helped me reach this important goal in my life.

TABLE OF CONTENTS

Page

LIST OF TABLES v

LIST OF FIGURES vi

CHAPTER

1 INTRODUCTION 1

Cognitive Load Theory 2

Cognitive Theory of Multimedia Learning 5

Verbal Redundancy 7

Learner Control 12

Overview of the Study 16

2 METHODS 18

Participants and Design 18

Materials 19

Measures 26

Procedure 29

3 RESULTS 31

Prior Knowledge 31

Text-Sound Association 31

Cognitive Load 34

Attitudes 42

4 DISCUSSION 57

Text-Sound Association 57

Cognitive Load 60
Attitudes 63
Limitations 65
Implications 66
Future Directions 67
REFERENCES 68
APPENDIX
A TEXT PASSAGE FOR THE LISTENING TASK 75
B WORD CLASSIFICATION FORM FOR INSTRUCTORS 77
C WORD CLASSIFICATION FORM FOR STUDENTS 79
D TARGET VOCABULARY WORDS 81
E TEXT-SOUND ASSOCIATION PRETEST 83
F TEXT-SOUND ASSOCIATION POSTTEST 85
G COGNITIVE LOAD MEASURES 87
H ATTITUDE QUESTIONNAIRE 89
I IRB APPROVAL 91

LIST OF TABLES

Table Page

1. Descriptive Statistics of Pretest and Posttest for Text-Sound Association.....32

2. Bivariate Correlations between Cognitive Load Measures.....34

3. Descriptive Statistics of Cognitive Load Measures.....35

4. Descriptive Statistics of Attitude Measures.....43

5. Summary of Factor Loadings of a One-Factor Solution for Student Attitudes.46

6. Descriptive Statistics of Factor "Attitudes towards the treatments".....47

7. Responses to Item 13 "What would help you better associate written words with their corresponding sounds?".....49

8. Responses to Item 14 "What did you like best about the activity?".....51

9. Responses to Item 15 "What could be done to improve the activity?" 54

LIST OF FIGURES

Figure Page

1. Screen of the vocabulary introduction section....21
2. Screen of the listening task for the not-learner-controlled + audio + no text condition 23
3. Screen of the listening task for the not-learner-controlled + audio + full text condition 23
4. Screen of the listening task for the not-learner-controlled + audio + keyword text condition 24
5. Screen of the listening task for the learner-controlled + audio + no text condition 24
6. Screen of the listening task for the learner-controlled + audio + full text condition 25
7. Screen of the listening task for the learner-controlled + audio + keyword text condition 25
8. Screen of a sample pretest item....27
9. Screen of a sample posttest item 28
10. Plot of adjusted means of frustration level showing a significant interaction between learner control and format of presentation of information 41

CHAPTER 1

INTRODUCTION

According to the British Council (n.d.), English is spoken as the official language in at least 75 countries whose combined population consists of over two billion human beings. Additionally, about 375 million people speak English as a second or foreign language (ESL/EFL). Therefore, the importance of learning ESL or EFL has increased in sufficient amount for it to become a subject matter on which many educators focus their efforts.

Multiple learning theories frame the design of ESL/EFL instruction: Cognitive load theory is not an exception. According to Sweller, Ayres and Kalyuga (2011), cognitive load theory is mainly concerned with the limitations of human working memory and how to overcome these limitations through the design of instructional materials. One of the principles of this theory indicates that when duplicated information is presented to learners simultaneously, this additional material unnecessarily increases the level of mental effort required for the learners to process the information (i.e. their "cognitive load"), thus hindering learning. This specific aspect of cognitive load theory is known as the redundancy effect (Kalyuga, Chandler, & Sweller, 1998, 1999; Sweller, 2005).

Research has shown that the negative impact of the redundancy effect may be less pronounced when it comes to learning of a second or foreign language (Plass & Jones, 2005). When a language being learned is non-phonetic in nature (e.g. English), the way words are pronounced frequently does not correspond to the way they are written and, as a consequence, inferring the written form of a word from its sound (and vice versa) may

pose a challenge for some learners. Consequently, presenting both written and spoken forms of the word simultaneously may not have a negative effect on learning. Moreover, some studies indicate that the presentation of duplicated information benefits learning of a second or foreign language (e.g. Baltova, 1999; Bird & Williams, 2002; Borrás & Lafayette, 1994; Garza, 1991; Hayati & Mohmedi, 2011; Koolstra & Beentjes, 1999; Markham, 1999; Neuman & Koskinen, 1992; Vanderplank, 1988). In order to contribute to the clarification of the aforementioned debate, this study investigated how simultaneously presented duplicated vocabulary words—in both visual and auditory formats—and control over the number of times an audio recording of a given word is listened to affects the ability of EFL learners to form text-sound association, their attitudes towards the treatments, and their reported level of cognitive load.

Cognitive Load Theory

Cognitive load theory (Clark, Nguyen, & Sweller, 2006; Plass, Moreno, & Brünken, 2009; Sweller, 2010; Sweller et al., 2011) consists of a series of constructs intended to guide effective instructional design based on the characteristics of the human cognitive architecture. Its goal is to direct "working memory resources to the intrinsic essentials of a curriculum area and away from extraneous aspects" (Sweller et al., 2011, p. 45). The act of processing instructional information during learning imposes a cognitive load on working memory. From this basic idea, three different categories of cognitive load emerge: intrinsic, extraneous, and germane. Intrinsic cognitive load originates from the demands of working memory resources imposed by the structure of the information to be learned; it is not only inherent to the instructional materials but also inalterable by instructional design without changing the nature of the task. Conversely,

extraneous cognitive load stems from the way instructional design presents the information and activities with which the learner must engage to accomplish learning goals; this type of cognitive load tends to be detrimental to the learning process when intrinsic and extraneous loads together exceed working memory capacity. Working memory resources devoted to the processing of information that is relevant to learning are known as germane cognitive load. A combination of intrinsic and extraneous cognitive load determine the total amount of cognitive load that is imposed on working memory at a specific moment; if working memory capacity is not replete, remaining cognitive resources are available to be occupied by germane cognitive load. Consequently, optimal instructional design should deal with a limited working memory capacity by promoting the availability of germane resources while minimizing extraneous cognitive load to foster learning.

As maintained by Sweller et al. (2011), the main factor from which levels of intrinsic and extraneous cognitive load derive is element interactivity, which is determined as a result of the amount of logically related components that must be processed simultaneously by working memory in order for the information contained in the instructional materials to be comprehended and ultimately learned. If the elements that constitute the information to be processed can be learned in isolation, the material is low in element interactivity. If, on the contrary, the constituents of the instructional material cannot be understood separately and thus interact, the level of element interactivity is considered high. Instructional materials with high element interactivity (i.e. high intrinsic cognitive load) demand for the minimization of extraneous cognitive

load, when working memory capacity may be exceeded, through the use of instructional design based on the principles of cognitive load theory.

Among other effects identified by cognitive load theory (e.g. modality, transience, split-attention, etc.), a redundancy effect (Kalyuga et al., 1998, 1999; Sweller et al., 2011) occurs when multiple sources of information that are intelligible in isolation are presented simultaneously, thus imposing an extraneous cognitive load on the working memory of a person in her attempt to integrate the redundant information. A redundancy effect is especially observed when complex content is presented to novice learners (Clark et al., 2006). For instance, when a subject who is learning to use a word processor is presented with identical instructions on screen and in a printed manual, working memory attempts to process this duplicated information simultaneously; duplicated information leads to the unavailability of working memory resources that could be allocated for germane cognitive load and are conversely occupied by redundant information, thus generating extraneous cognitive load. Conditions necessary for the redundancy effect to take place are that sources of information must be understandable independently (i.e. the information is duplicated), have a high level of element interactivity (i.e. the components of the instructional material are logically interrelated), and be long and complex enough to generate high cognitive load (Sweller et al., 2011). Degrees of expertise may also have an influence on this effect (Kalyuga, Chandler, & Sweller, 2000; Sweller et al., 2011). When experts are presented with information that may be essential to novices but already known to them, such information unnecessarily occupies working memory resources and may contribute to extraneous cognitive load, this is known as the expertise reversal effect (Kalyuga, Ayres, Chandler, & Sweller, 2003).

Cognitive Theory of Multimedia Learning

According to the cognitive theory of multimedia learning (Mayer, 2005, 2009, 2011; Mayer & Moreno, 2003), people learn better from a combination of pictures (visual information) and words (auditory information) than from words alone: This precept is known as the multimedia principle (Mayer, 2005). Three assumptions about human working memory are fundamental to multimedia learning: dual-channel, limited capacity, and active processing.

Human working memory relies on two complementary channels, both of them limited in capacity and duration (Clark et al., 2006; Sweller, 1998). One channel is devoted to the processing of visual/pictorial information—entering the system through the eyes—and receives the name of visuospatial sketchpad; the other channel is dedicated to process auditory/verbal information—acquired through the ears—and is known as the phonological loop (Baddeley, 1992; Mayer, 2005, Sweller et al. 2011). Nevertheless, a cross-channel representation may occur when information presented through one channel is represented in the other one (Mayer, 2005; Paivio, 1986). For example, in the case of text reading, information is initially acquired and represented through the visuospatial sketchpad and can ultimately be cross-represented in the phonological loop, given the verbal nature of the information. In addition, both the limited capacity and active processing assumptions relate respectively to the restrictions in cognitive capacity of the processing channels and the active cognitive processing in which learners must engage with the aim of constructing mental representations of experiences (Mayer, 2005). In order for active learning to occur, processes such as selection, organization, and integration must take place. In so doing, and given the constraints of cognitive channels,

learners pay attention to and select specific elements (words and images) from the presented information for further processing in working memory, where those elements are organized into models (verbal and pictorial) and integrated with prior knowledge (Mayer, 2009). For instance, a reader may select a group of letters that constitute a word or several words that constitute a sentence—depending on the reader's level of expertise—in order to organize them into a model that acquires meaning when it is integrated with prior knowledge.

Regarding second language acquisition (SLA), Plass & Jones (2005) proposed a model based on the cognitive theory of multimedia learning (Mayer, 2005, 2009, 2011; Mayer & Moreno, 2003) and the basic components of the interactionist SLA process proposed by Chapelle (1998). According to this model, the learner selects relevant and comprehensible input (i.e. target language) from the environment, for its subsequent representation in the corresponding cognitive channels. The process of focusing attention and selecting information is denominated as apperception and is facilitated through the use of multimedia (Plass & Jones, 2005). The organization of information into semantically and syntactically interrelated verbal and pictorial models constitutes the foundation for language comprehension, also referred to as intake, which incorporates potential to develop the linguistic system of the learner (Chapelle, 1998). The ensuing integration in working memory of intake and prior knowledge directly influences the production of comprehensible output in the target language, which in turn will elicit new input (e.g. feedback) from the environment in a continuing process of language development, defined by Long (1996) as negotiation of meaning.

The multimedia principle has been found to occur in multimedia learning of a second language (see Meskill, 1996; Plass & Jones, 2005). Other principles that also have been found to apply in this context are the individual differences and the advance organizer principles, which indicate that language acquisition is facilitated when learners can choose between verbal or visual annotations and when advance organizers (i.e. material intended to activate prior knowledge) are presented prior to listening and reading comprehension tasks (Plass & Jones, 2005; Yeh & Lehman, 2001). Learners of a second language may also benefit from the use of learner control, especially those learners with low ability (Yeh & Lehman, 2001). It is worth mentioning that when practice and testing modes differ, output of tested knowledge is hindered, thus practice and testing modes should be equivalent (Plass & Jones, 2005).

Verbal Redundancy

The simultaneous presentation of identical text and narration is defined by Moreno and Mayer (2002) as verbal redundancy. During the past years, multiple studies have been conducted regarding verbal redundancy in multimedia presentations finding negative effects on learning (Gerjets, Scheiter, Opfermann, Hess, & Eysink, 2009; Kalyuga et al., 1999, 2000, 2004; Leahy, Chandler, & Sweller, Experiment 2, 2003; Mayer, Heiser, & Lonn, 2001). In a lesson about soldering theory, Kalyuga et al. (1999, Experiment 1) found on transfer tests that the use of simultaneous duplicated information was disadvantageous to learners, due to the generation of additional cognitive load, while information presented in auditory format rendered performance effective. In another study, Mayer et al. (2001, Experiment 2) compared the use of one of three conditions—no text, summary text, and full text—in a multimedia presentation, not under control of

the learner, about lightning formation with concurrent animation and audio; the results indicated that the learners in the no text condition outperformed the other two groups in both retention and transfer tests. Moreover, in a study involving hypermedia, Gerjets et al. (2009) obtained a redundancy effect when simultaneous spoken and written explanations, in comparison to written instructions only, rendered instruction less effective; in contrast, a redundancy effect was not observed when comparing the spoken and written condition to the spoken text condition.

Conversely, some studies have found verbal redundancy to be beneficial in some cases (Lewandowski & Kobus, 1993; Mayer, 2011; Mayer & Johnson, 2008; Montali & Lewandowski, 1996; Moreno & Mayer, 2002). For instance, Moreno and Mayer (2002) found that learners presented with concurrent text and narration, in a multimedia presentation about lightning formation, performed significantly better on retention, transfer, and matching tests when animation was not presented simultaneously. Based on these findings, the authors concluded that students learn better when the visual channel is not overloaded by processing concurrent on-screen text and animation (Mayer, 2011; Mayer & Moreno, 2003; Moreno & Mayer, 2002). In another study, Mayer and Johnson (2008) used short redundant on-screen text displayed next to the visual they described and observed a significant improvement in performance on retention tests attributed to the guidance that redundant labels offered; performance on transfer tests did not differ between groups.

In regards to SLA, some studies indicate that concurrent text and audio render negative effects on learning of a second or foreign language (Diao, Chandler, & Sweller, 2007; Diao & Sweller, 2007; Moussa, 2008; Sweller et al., 2011). Diao and Sweller

(2007) found that the presentation of concurrent full text and audio rendered text comprehension and recall of EFL less effective. With regards to listening comprehension, Diao, Chandler, and Sweller (2007) concluded that the simultaneous presentation of audio and scripts or subtitles benefitted comprehension of EFL passages but interfered with learning to listen. Moreover, as said by Moussa (2008), reading led to better performance in EFL listening comprehension than listening only or simultaneous listening and reading of full texts.

On the other hand, when learners are nonnative speakers of a language, the use of on-screen text may be valid in pedagogical terms (Mayer, 2010). Furthermore, some studies have found the use of concurrent text and audio to be beneficial in learning a second or foreign language (Baltova, 1999; Bird & Williams, 2002; Borrás & Lafayette, 1994; Garza, 1991; Hayati & Mohmedi, 2011; Koolstra & Beentjes, 1999; Markham, 1999; Neuman & Koskinen, 1992; Vanderplank, 1988). Borrás & Lafayette (1998) studied the effects of subtitled video, varying the complexity level of tasks, on speaking skills of intermediate/advanced learners of French as a foreign language. They concluded that such a condition improved both comprehension and speaking performance, especially in more complex tasks. Bird & Williams (2002) found that the presentation of words using a combination of visual (i.e. text) and auditory formats improved word recognition and facilitated learning of new words in explicit and implicit memory tests, particularly for EFL learners. In another study, Hayati and Mohmedi (2011) examined the effects of video presented with or without subtitles—in English or Persian language—on listening comprehension of Iranian intermediate EFL learners; the findings indicated that

subtitles in English facilitated listening comprehension whilst subtitles in the learners' first language presumably imposed an extraneous cognitive load.

Overall, the current body of literature is mixed with regards to whether the use of simultaneous text and audio is effective in learning a second or foreign language. Multiple studies, whose results concerned the use of simultaneously presented text and audio, were found to be somewhat beneficial in learning a second or foreign language (Baltova, 1999; Borrás & Lafayette, 1994; Garza, 1991; Hayati & Mohmedi, 2011; Koolstra & Beentjes, 1999; Markham, 1999; Neuman & Koskinen, 1992; Vanderplank, 1988) and were conducted using instructional materials that included dynamic visuals (e.g. video or animation); such a combination contradicts the idea that students learn better when the visual channel is not overloaded by processing concurrent on-screen text and animation (Mayer, 2011; Mayer & Moreno, 2003; Moreno & Mayer, 2002).

Additionally, looking at an investigation of the attitudes of learners towards simultaneous presentation of identical on-screen text and narration may enrich the results of studies regarding this matter, even though the evaluation of attitudes does not necessarily reflect learning but offers a perspective as to the reactions of learners towards the treatments. Along these lines, it is worth mentioning that some studies (Ayres, 2002; Heller, 2005; Stepp-Greany, 2002; Wiebe & Kabata, 2010) have found that learners have positive attitudes towards the integration of computers in second or foreign language learning. For instance, in a study related to the use of computer assisted language learning (CALL) conducted mostly with undergraduate students of EFL, Ayres (2002) found that the majority of learners had positive attitudes towards the use of CALL in their foreign language courses, given that they considered it relevant to their needs and indicated their

preference for a higher frequency of its use, as long as it was linked to the course curriculum.

There is ample evidence in the literature that the use of concurrent written and spoken information improves comprehension of a second or foreign language. As said by Sweller et al. (2011), considering that comprehension differs from learning, transfer of knowledge must be tested in studies about the use of concurrent text and verbatim audio by means of using contents different from those presented during instruction in order to ensure that it is learning, and not comprehension of a specific material, that is being measured.

When referring to learning of a second or foreign language, it can be argued that, for some learners, the simultaneous presentation of identical on-screen text and narration will not generate a redundancy effect because the two sources of information may not be completely intelligible in isolation. Moreover, given that learners must take comprehensible input from the environment to acquire linguistic competencies (Chapelle, 1998), integration of both sources may foster learning of a second or foreign language instead of hindering it, not only in the form of full phrases but also in the form of keywords or short phrases. This was seen in a study conducted by Mayer and Johnson (2008), where subjects presented with short, redundant on-screen text significantly outperformed those in the non-redundant conditions on retention tests given that, according to the authors, redundant short phrases foster information processing by guiding the selection of relevant information without increasing extraneous cognitive load. EFL learners may benefit from being presented with both sources in order to appropriately associate the visual and aural forms of target language of a non-phonetic

nature. This idea will be referred to, in the context of the current study, as text-sound association.

Some studies have explored the concept of association of words (or images) and sounds from a neurological perspective. For instance, in a study about memory retrieval, Nyberg, Habib, McIntosh, and Tulving (2000) used positron-emission tomography (PET) to observe the activation of brain regions engaged during encoding of word-sound pairs, to ultimately find that activation of some regions of the auditory responsive cortex took place during information retrieval when the visual word of the encoded pair was used as stimulus. Similarly, in a study about retrieval of sensory-specific information, subjects learned pairs of pictures and sounds; functional magnetic resonance imaging (fMRI) was used during a recall test to identify activation of the brain's visual and auditory cortex regions involved in this process. The authors concluded that retrieval of visual and auditory information leads to activation of sensory regions also active during perception of this same information (Wheeler, Petersen, & Buckner, 2000).

In the area of language acquisition, in a study about instructional encoding methods with beginning learners of Chinese as a foreign language, Shen (2010) compared verbal encoding only versus verbal plus imagery encoding to examine learning of concrete and abstract words. The results indicated that the verbal plus imagery encoding method significantly improved learners' retention of shape and meaning of abstract words but not their sounds.

Learner Control

The degree to which a learner can determine the pace of instruction and/or its content is known as learner control (Clark et al., 2006); the term locus of control

describes if the learner, an instructor, or a computer program decides the course of action of instruction (Alessi & Trollip, 2001; Merrill, 1984). According to Clark et al. (2006), a learner should be given control over the pace of instruction. Along these lines, Merrill (1974) argues that each learner possesses specific aptitudes at a particular moment and thus she should be given the opportunity to select the tactics that best suit those aptitudes in order to accomplish a learning goal. Such tactics in combination should constitute a scenario with all possible learning strategies for the learner to select the most appropriate one. In a learner-controlled environment, the student must learn to "recognize her own learning needs" (Merril, 1980, p. 89).

Moreover, learners with high prior knowledge should be provided with higher levels of learner control than students with low prior knowledge or lower ability (Kopcha & Sullivan, 2008; Merrill, 2002). When it comes to less experienced learners, making decisions regarding instruction can impose an overwhelming cognitive load due to both making such decisions and dealing with their consequences; accordingly, learners should make use of learner control over instruction only when their levels of expertise allow them to understand the consequences of their choices (Sweller et al., 2011). Nevertheless, as said by Mayer and Chandler (2001), when novices are presented with information that possesses a high level of element interactivity, introducing isolated fundamental elements separately, before the presentation of information with high levels of element interactivity, renders learning more effective and reduces cognitive load, especially in the case of learners who are given control over the pace of the information presented first. This is known in cognitive load theory as the isolated elements effect (Sweller et al., 2011).

In addition, restricting learner control from more experienced subjects who favor it may lead to lower performance levels than those of peers whose preference for learner control is matched (see Kopcha & Sullivan, 2008), showing an expertise reversal effect. These learners benefit from learner control when they are able to make the decision of investing cognitive resources on information that is new or more relevant.

Instructionally paced learning environments (e.g. instructor-led classes or uninterrupted multimedia animations) impose a high cognitive load on learners; appropriate instructional design should be used to manage cognitive load in such cases. The transient nature of verbal and visual information (e.g. spoken presentations or instructional videos) may impose additional demands to working memory since such elements must be held in working memory and processed before they are forgotten. A transient information effect is observed when increased cognitive load affects learning due to the demands imposed on working memory that stem from processing transient information (Leahy & Sweller, 2011; Sweller et al., 2011). Learner control offers a way to reduce loss of learning due to transience by means of allowing the learner to control the speed of presented transient information in ways such as stopping and reducing the speed of audio or animation. Some studies have found this instructional strategy to be beneficial for learners. In a study about the use of interactive features of videos, Schwan and Riempp (2004) found that providing learners with control over the speed and direction of instructional videos improved learning outcomes and time in comparison to those who were not given control over the video sequence. Similarly, Hasler, Kersten, and Sweller (2007) found that groups with learner control over the pace of a narrated animation reported lower cognitive load and better learning performance than groups

with no learner control, especially for information with a high level of element interactivity. Additionally, in a meta-analysis of related studies, Kraiger and Jerden (2007) suggested that the presence of learner control has a positive impact on learning, that learner control over navigation and pace renders better results than control over content, and that the amount of learner control given to a student should be proportional to the amount of knowledge possessed by that learner in a particular domain.

In the area of SLA, some authors have found the use of learner control to be beneficial. Meskill (1996) argues that providing learners with control over the pace of aural streams contributes to a better identification of discourse chunks which, in turn, favors the development of listening skills by means of facilitating retention, sound discrimination, inference from context, and aural processing. Moreover, in a study about the effects of learner control of speech rate on listening comprehension, conducted with non-native speakers enrolled in an intensive English program, Zhao (1997) concluded that listening comprehension of participants in the learner-controlled conditions was higher than that of participants who were not given control over speech rate while listening to a series of sentences; positive attitudes towards the use of learner control were found among participants of the corresponding conditions. Furthermore, in a study about the effects of learner control over the use of subtitles on listening comprehension, Gibbs (2009) concluded that, albeit not statistically significant, scores for the learner control group were higher than those for the non-controlled condition but, when grouped by listening skills, participants with lower ability significantly improved their listening skills in comparison to higher ability subjects who ultimately learned more words. In another study, Heift (2002) analyzed the effects of learner control on error correction of

students of an introductory German course and found that lower performers made more use of learner-controlled traits of the instructional environment to correct their errors than did high performers.

Overview of the Study

The main purpose of this study was to investigate the effects of concurrent onscreen text and equivalent audio as well as their relationship with levels of learner control on text-sound association, attitudes, and cognitive load experienced by EFL learners, in the context of a computer-based multimedia instructional program. The study investigated the following research questions:

1. What is the impact of different types of text (i.e. full and keyword), presented in combination with equivalent audio in a multimedia instructional program, on learner text-sound association, cognitive load, and attitudes towards the treatments?
2. What is the relationship between learner control and different types of text (i.e. full and keyword) in combination with equivalent audio regarding text-sound association and cognitive load?
3. What is the effect of learner control on learner text-sound association, cognitive load, and attitudes towards the treatments?

The two independent variables manipulated in this study were format of presentation of information, with three levels (audio + no text vs. audio + full text vs. audio + keyword text) and level of learner control with two levels (learner-controlled vs.

not-learner-controlled). Performance on text-sound association, experienced level of overall cognitive load, and attitudes of the learners towards the treatments were measured as dependent variables. Time on vocabulary introduction was examined as a moderator variable for all the treatment conditions. Time on the listening task and number of times a participant replayed the audio (i.e. clicked on the play button) were analyzed as moderator variables for the learner-controlled conditions. Other factors were controlled to be constant among groups.

CHAPTER 2

METHODS

Participants and Design

The sample for this study was comprised of 200 students enrolled in a basic-level course of EFL at a large Mexican university. A total of 137 (68.5 %) participants were females and 63 (31.5 %) were males. They were undergraduate students majoring in one of the following areas of the Health Sciences: Medical Surgery (13.5%), Dentistry (20%), Nutrition (24%), or Biopharmaceutical Chemistry (42.5%). They were 18 years or older and their average age was 18.56 (SD = 1.83). Spanish was their first language. The level of knowledge of the English language possessed by the participants was established by the institution through the administration of a diagnostic test designed for this purpose. Subjects participated in the study for course credit and on a voluntary basis, as part of the laboratory activities of the course. Monetary incentives were given to subjects in order to encourage them to take part in the study.

A pretest-posttest, 2 (learner-controlled vs. not-learner-controlled) x 3 (audio + no text vs. audio + full text vs. audio + keyword text) factorial design was used in this study. The subjects were randomly assigned to one of the following treatment conditions:

1. Not-learner-controlled + audio + no text
2. Not-learner-controlled + audio + full text
3. Not-learner-controlled + audio + keyword text
4. Learner-controlled + audio + no text
5. Learner-controlled + audio + full text
6. Learner-controlled + audio + keyword text

Materials

A computer-based multimedia instructional program, developed using Adobe Flash, was used to deliver the lesson and measures to collect data. Six different versions of the program were developed to match the treatment conditions. The instructional program included the following sections in the order listed: a tutorial to familiarize the participants with the use of the program, a text-sound association pretest, a vocabulary introduction, a listening task, a text-sound association posttest, cognitive load measures, and an attitude questionnaire.

The listening task, consisting of a short passage about a fable, was adapted from the original text to meet the needs of the study (see Appendix A). The target words included in the lesson as well as the pre and posttest were selected by two instructors of basic-level courses and 26 basic-level students of the EFL program. The word selection process comprised two phases:

1. The instructors rated the level of difficulty of all the adjectives, nouns, verbs, and prepositions in the passage using a five-point Likert scale. The scale ranged from 1 = easiest to 5 = most difficult. They also classified the words, to the best of their knowledge, as either known or unknown by basic-level EFL students (see Appendix B).
2. The words were also classified by the students as either known or unknown to them (see Appendix C).

Those words rated by the instructors as being unknown to their students and having a level of difficulty of 4 or 5, as well as classified as unknown by the students, comprised the target language presented on the lesson (see appendix D). On the basis of

time availability of the students, the EFL program determined which subjects would take part in this stage of the study.

As for the computer-based multimedia program, the first part consisted of a brief tutorial to show the participants how to interact with the software. The second part was devoted to the lesson and the elements comprising it were presented in the following order: text-sound association pretest, vocabulary introduction, listening task, text-sound association posttest, cognitive load measures, and attitude questionnaire. The instructions, word definitions, cognitive load measures, and attitudes questionnaire were presented in the participants' first language. Audio and text corresponding to the target words and the listening task, along with the items of the pre and posttest, were presented in English.

Each target word was presented to the subjects on a screen displaying the word, its definition, and an example of its use in a sentence along with a translation of the sentence into the participants' first language (see Figure 1). The listening task consisted of a narrated version of the passage accompanied by text when required by the treatment condition. A play button started the audio reproduction. In the learner-controlled conditions, a pause button intermitted the audio reproduction and a control bar allowed the user to move backwards in the audio sequence to listen a specific segment again. Pressing the arrow up and down keys on the computer keyboard allowed the users in any treatment condition to adjust the volume of the audio as needed, even though a volume adjustment section was presented in the tutorial with the purpose of setting the volume of the audio prior to the listening task. Abandoning the narration screen was not possible until the entire audio sequence had been played. All the audio recordings were made by a young adult, female, native English speaker.

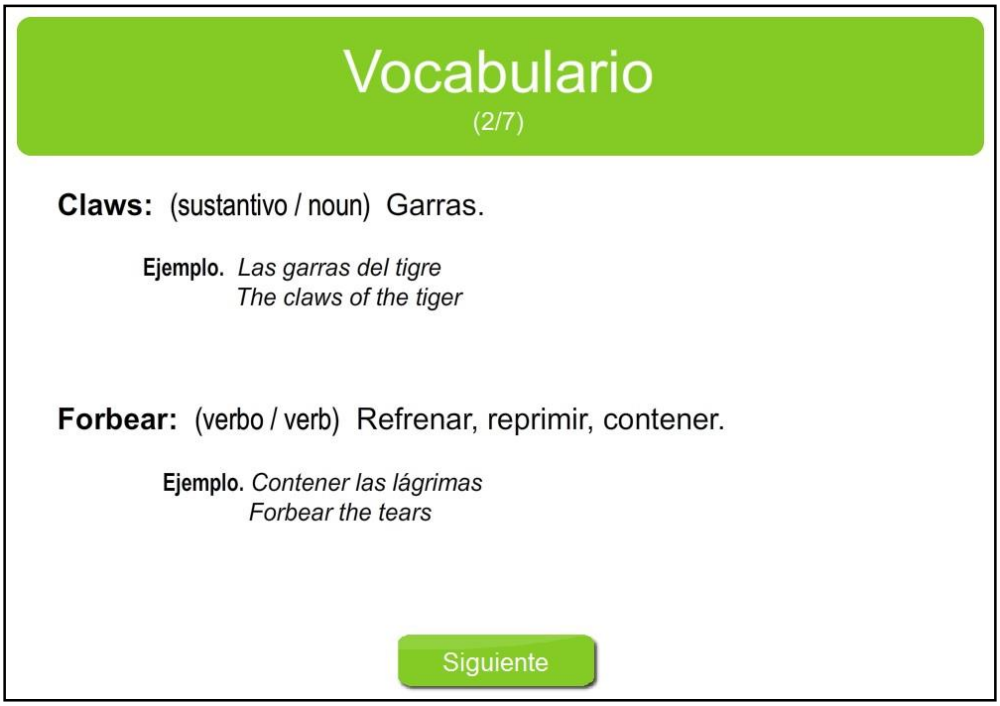

Figure 1. Screen of the vocabulary introduction section.

Six different versions of the instructional software were used to implement the treatment conditions; text and audio did not vary between treatments sharing specific conditions. The characteristics of the treatments were as follows:

- *Not-learner-controlled* + *audio* + *no text*: The play button was disabled once pressed to start the audio reproduction. Repetition of audio segments was not allowed. A timer showed the progress of the audio reproduction in seconds (see Figure 2).
- *Not-learner-controlled* + *audio* + *full text*: In addition to the characteristics of the not-learner-controlled + audio + no text condition, this version of the software displayed onscreen synchronous verbatim text for the full passage (see Figure 3).

- *Not-learner-controlled + audio + keyword text*: It differed from the not-learner-controlled + audio + full text condition in that only keywords showing the target words were displayed on the screen (see Figure 4).
- *Learner-controlled + audio + no text*: The play button was not disabled after being pressed to start the audio reproduction. Instead, its appearance changed to that of a pause button that, once pressed, stopped the audio reproduction and went back to its original appearance of a play button. A control bar allowed the user to navigate back in the narration and repeat specific audio segments; a short legend was displayed under the control bar to inform the user that interaction with the control bar was possible in order to repeat audio segments. The audio stopped playing each time the user interacted with the control bar or the pause button so she had to click on the play button to continue listening to the audio. A log entry was created each time the user clicked on the play button. A timer showed the progress of the audio reproduction in seconds (see Figure 5).
- *Learner-controlled + audio + full text*: Besides the characteristics of the learner-controlled + audio + no text condition, synchronous verbatim text for the full passage appeared on the screen (see Figure 6).
- *Learner-controlled + audio + keyword text*: In addition to the features of the learner-controlled + audio + no text condition, keywords showing the target words were displayed on the screen (see Figure 7).

Figure 2. Screen of the listening task for the not-learner-controlled + audio + no text condition.

Figure 3. Screen of the listening task for the not-learner-controlled + audio + full text condition.

Figure 4. Screen of the listening task for the not-learner-controlled + audio + keyword text condition.

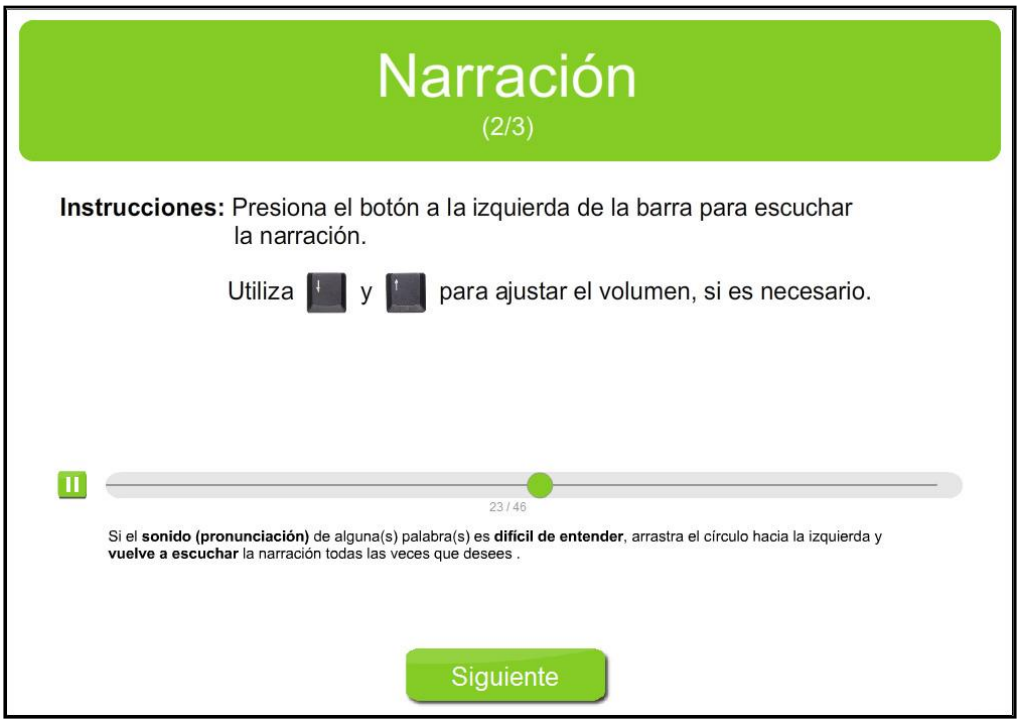

Figure 5. Screen of the listening task for the learner-controlled + audio + no text condition.

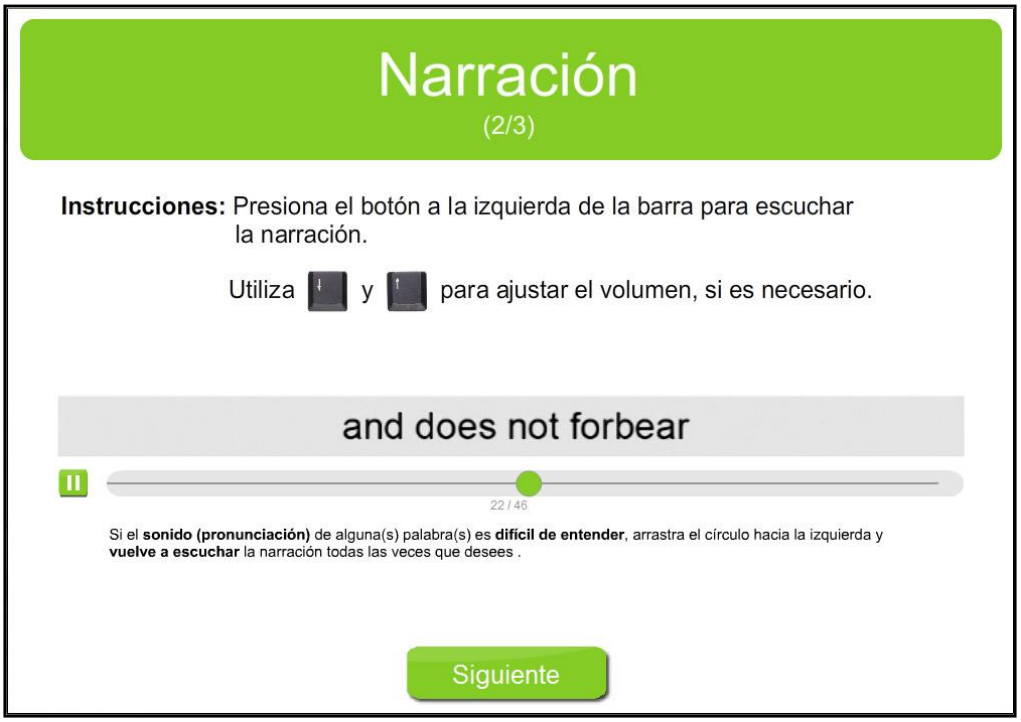

Figure 6. Screen of the listening task for the learner-controlled + audio + full text condition.

Figure 7. Screen of the listening task for the learner-controlled + audio + keyword text condition.

Measures

Text-sound association refers to relating the written and aural forms of a word. Learners of a non-phonetic foreign language, as is the case of English, frequently find it difficult to carry out text-sound association of particular words given that, in non-phonetic languages, it is common that the way some words are pronounced does not match the way they are written and, consequently, inference of such written word forms from their sound (and contrariwise) may impose an additional amount of cognitive load for some learners. In order to measure text-sound association in this study, the following pre and posttest were employed.

A pretest, consisting of 10 multiple choice items (i.e. one item per target word), measured text-sound association of the participants with regards to the target words prior to taking the lesson (see appendix E). The participants read the word and clicked on a play button to reproduce each of the four sounds presented. The audio for each word could be played up to three times, after which the button would be disabled and showing a diagonal red line. Continuing to the next item was not possible until all audio sequences had been played at least once. Participants then selected the option that they believed corresponded to the sound of the target word by dragging and dropping the target word onto the rectangle associated to the selected sound (see Figure 8). Sounds used as distractors for each item were words or phrases whose sound is similar to that of the target word (e.g. distractors for the item *forbear* were "fur-bear", "forbore", and "for-beer"). Results of the pretest served as a baseline to be compared against posttest results.

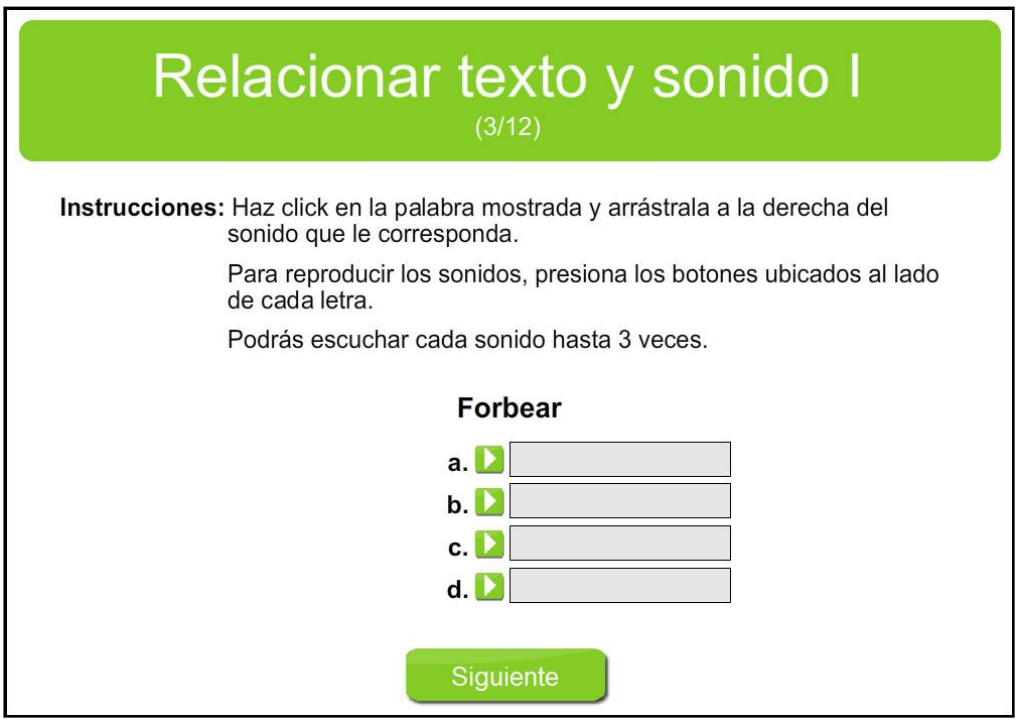

Figure 8. Screen of a sample pretest item.

A posttest, containing 10 multiple-choice items (i.e. one per target word), measured text-sound association of the participants with regards to the target words after taking the lesson (see appendix F). The items were constituted by the target words within the context of a phrase unrelated to the topic of the passage so that measurement of text-sound association of the target vocabulary would take place in a context different to that used during target vocabulary introduction (See Sweller et al., 2011). Each phrase was comprised of one target word accompanied by five to six additional words that contributed to increase element interactivity and, consequently, intrinsic cognitive load. Using the same method they applied for answering the pretest items, the participants selected the option that they believed corresponded to the aural form of the phrase (see Figure 9). Sounds used as distractors for each item were phrases whose sound was similar

to that of the original phrase. Correct textual and aural forms of each target word were presented at the end of the posttest. Results obtained were compared against the pretest results to determine performance with regards to text-sound association after having received a specific treatment.

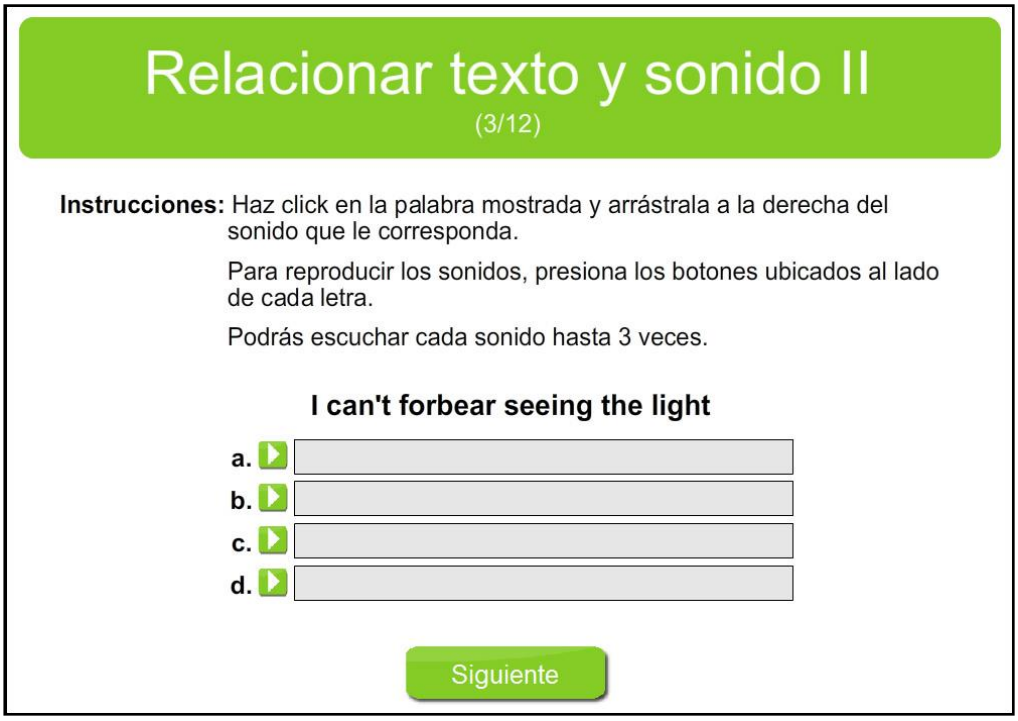

Figure 9. Screen of a sample posttest item.

In order to subjectively measure cognitive load, an adaptation of the NASA Task Load Index (NASA-TLX, Hart & Staveland, 1998) was employed. This instrument included five different measures presented in the form of a nine-point Likert scale and each of them corresponding to one of the following categories (see Appendix G): task demand (i.e. mental and physical activity required to complete the learning task), effort (i.e. work necessary to understand the contents of the learning environment), feeling of

success (i.e. how successful the subject felt in understanding the contents of the learning environment), navigation (i.e. effort invested to navigate the learning environment), and frustration level (i.e. insecurity, discouragement, irritation, stress, and annoyance felt during the learning task). The use of NASA-TLX to measure cognitive load has proven successful in other studies (see Gerjets, Scheiter, & Catrambone, 2004, 2006).

A questionnaire was used to measure the attitudes of the participants towards the treatments. This instrument consisted of 10 to 12 (depending on the treatment condition) five-point Likert type items and three open-ended questions (see Appendix H). Cronbach's alpha for the Likert type items was .81. The time each participant took to complete the listening task was constant for the not-learner-controlled conditions (i.e. 46 seconds) and varied for the learner-controlled conditions. To calculate the latter, the software recorded the number of seconds each subject spent listening to the audio sequence; the timer was not visible to the user. Similarly, the number of times an audio segment was played by the participants was constant for the not-learner-controlled conditions (i.e. one time) and varied for the learner-controlled conditions. To compute the latter, the software kept track of the number of times a participant clicked on the play button. As for the time users spent on the vocabulary section, the software stored the number of seconds it took to each participant to complete this stage.

Procedure

The study was conducted in a computer laboratory setting equipped with 18 personal computers each provided with a set of headphones that facilitated individual listening for each subject. The participants were randomly assigned to the treatment conditions. To accommodate the totality of participants, the study was completed in 22

sessions of one hour approximately. At the beginning of the session the participants were informed that they would be taking part in a research study and received instructions about the activities that would take place during the session. Instructions on how to access and use the program were provided in the participants' first language. Once the subjects clicked on the start button of the multimedia computer-based program, the software randomly assigned them to one of the treatment conditions. The participants then went through the tutorial and individually completed the stages of the study in the following sequence: text-sound association pretest, vocabulary introduction, listening task, text-sound association posttest, cognitive load measures, and attitude questionnaire. At the end of the session, participants were thanked and received the incentives.

CHAPTER 3

RESULTS

Family-wise type I error was set at the .05 level. With regards to effect size, partial eta squared (η_p^2) was used to report these results. According to Hatcher (2013), Cohen's criteria for eta squared (η^2) is pertinent to interpret effect sizes indicated by η_p^2. Consequently, .01, .06, and .14 are considered as small, medium, and large effects, respectively.

Prior Knowledge

A two-way ANOVA was conducted on the pretest scores to determine whether the subjects' prior knowledge differed across treatment conditions. The results indicated that there was no significant difference between the learner-controlled and not-learner-controlled conditions, F (1,194) = 3.49, MSE = 7.77, p = .06, η_p^2 = .02. Additionally, no significant differences were found between the three conditions involving format of presentation of information (i.e. audio + no text, audio + full text, and audio + keyword text), F (2,194) = 1.48, MSE = 3.29, p = .23, η_p^2 = .02; nor was there any significant interaction, F (2,194) = .72, MSE = 1.61, p = .49, η_p^2 < .01. Means and standard deviations are presented in Table 1.

Text-Sound Association

A two-way analysis of covariance (ANCOVA) was conducted to evaluate the potential effects of learner control and format of presentation of information on the posttest scores. Descriptive statistics are presented in Table 1.

Table 1

Descriptive Statistics of Pretest and Posttest for Text-Sound Association

			Pretest		Posttest	
		n[a]	*M*	*SD*	*M*	*SD*
Not learner controlled	AN	27	3.63	1.39	3.41	1.45
	AF	29	4.21	1.05	3.59	1.72
	AK	35	3.74	1.56	3.63	2.03
Learner controlled	AN	37	4.00	1.49	3.24	1.82
	AF	42	4.31	1.68	3.52	1.71
	AK	30	4.47	1.59	4.57	1.79

Note. *M* = Mean. *SD* = Standard Deviation. AN = Audio + No Text. AF = Audio + Full Text. AK = Audio + Keyword Text.
[a]Sample size within conditions.

The pretest scores were used as the covariate to control for the potential effects of prior knowledge on text-sound association. There were no main effects of learner control, F (1, 193) = .20, MSE = .58, p = .66, η_p^2 = .001, or format of presentation of information, F (2, 193) = 3.00, MSE = 8.81, p = .05, η_p^2 = .03; nor was there any interaction, F (2, 193) =1.51, MSE = 4.43, p = .22, η_p^2 = .02. However, a two-way ANCOVA was conducted after doing a median split on the pretest scores that were used as the covariate, where subjects were grouped by low (between 0 and 4) and high (between 5 and 9) prior knowledge; a significant main effect of format of presentation of information was found, F (2, 193) = 3.33, MSE = 10.11, p = .04, η_p^2 = .03, power = .63. On the other hand, there was no significant effect of learner control, F (1, 193) = .57, MSE = 1.72, p = .45, η_p^2 < .01, and there was a non-significant interaction, F (2, 193) = 2.43, MSE = 7.37, p = .09,

η_p^2 = .03, power = .49. Pairwise comparisons were conducted to assess differences between groups with regards to format of presentation of information. To control for type I error, the modified sequentially rejective Bonferroni (MSRM) procedure (Schaffer, 1986) was used. Pairwise comparisons were significant between the keyword text and no text conditions (p = .02) and between the keyword text and full text conditions (p = .04); the pairwise comparison between the no text and full text condition was not significant (p = .71). Additionally, within the keyword text condition, there was a significant difference between the not-learner-controlled and learner-controlled groups (p = .03). Although there was no significant main effect for learner control, differences were seen between the keyword text vs. no text (p < .01) groups and between the keyword text vs. full text (p = .01) groups, both within the learner-controlled condition.

A two-way ANCOVA was conducted to evaluate whether the independent variables varied as a function of the time subjects spent on the vocabulary introduction for all the treatment conditions; time spent on the vocabulary introduction was used as the covariate. All interactions between the independent variables and the time spent on the vocabulary introduction were non-significant (Fs < 2.67 and ps > .07). Additionally, a one-way ANCOVA was conducted to evaluate whether the time subjects spent on the listening task had a moderating effect on format of presentation of information with respect to posttest scores for the learner controlled conditions; time spent on the listening task was used as the covariate. There was no significant interaction between format of presentation of information and time spent on the listening task, F (3, 103) = 2.43, p = .07. Similarly, a one-way ANCOVA was conducted to evaluate whether the format of presentation of information or posttest scores varied as a function of the number of times

participants replayed the audio (i.e. clicked on the play button) for the learner controlled conditions; the number of times participants replayed the audio was used as the covariate. The interaction between the independent variables and the time spent on the listening task was non-significant, F (3, 103) = 1.86, p = .14.

Cognitive Load

Separate two-way ANOVAs were conducted on the scores of each measure of cognitive load included in NASA-TLX and overall cognitive load. Bivariate correlations between the cognitive load measures are presented in table 2. In addition, descriptive statistics are presented in table 3.

Table 2

Bivariate Correlations between Cognitive Load Measures

	CL1	CL2	CL3	CL4
CL2	.64**			
CL3	-.18*	-.26**		
CL4	.20**	.27**	-.08	
CL5	.40**	.40**	-.16*	.37**

Note. CL1 = Task Demand. CL2 = Effort. CL3 = Feeling of Success. CL4 = Navigation. CL5= Frustration Level.
* $p < .05$.
** $p < .01$.

Table 3

Descriptive Statistics of Cognitive Load Measures

			Task demand		Effort		Feeling of success		Navigation		Frustration level		Overall cognitive load	
		n[a]	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Not learner controlled	AN	27	4.15	2.27	4.07	2.29	5.52	1.65	1.63	1.31	1.93	1.47	3.46	1.01
	AF	29	4.59	1.59	4.69	1.83	5.76	1.68	2.28	2.22	2.97	1.96	4.06	1.25
	AK	35	4.54	2.01	4.60	1.77	5.46	1.36	2.09	1.95	2.11	1.66	3.76	1.05
Learner controlled	AN	37	4.22	1.70	4.46	1.99	5.68	1.77	1.92	1.61	2.41	1.85	3.74	1.11
	AF	42	4.17	1.62	4.71	1.76	5.36	1.81	1.67	1.16	2.17	1.31	3.61	0.73
	AK	30	4.00	1.74	4.00	1.64	5.70	1.88	2.00	1.44	2.47	1.61	3.63	0.73

Note. *M* = Mean. *SD* = Standard Deviation. AN = Audio + No Text. AF = Audio + Full Text. AK = Audio + Keyword Text.

[a]Sample size within conditions.

Overall cognitive load. A two-way ANOVA was conducted to evaluate the potential effects of learner control and format of presentation of information on the subjects' overall level of cognitive load. The average of the five NASA-TLX measures was used as the dependent variable. Results indicated that neither of the two main effects was significant; for learner control, F (1, 194) = .47, MSE = .46, p = .49, η_p^2 < .01; for format of presentation of information, F (2, 194) = .96, MSE = .94, p = .39, η_p^2 = .01. The interaction was non-significant, F (2, 194) = 2.15, MSE = 2.10, p = .12, η_p^2 = .02. A two-way ANCOVA was conducted to evaluate whether the time subjects spent on the vocabulary section had a moderating effect on the independent variables; time spent on the vocabulary section was used as the covariate. No significant moderating effect was found for learner control, F (1, 188) = .32, MSE = .31, p = .57, η_p^2 < .01; format of presentation of information, F (2, 188) = .96, MSE = .92, p = .39, η_p^2 = .01; or their interaction F (2, 188) = 2.87, MSE = 2.78, p = .06, η_p^2 = .03, power = .56. In addition, a one-way ANCOVA was conducted to evaluate whether the time subjects spent on the listening task had a moderating effect on format of presentation of information for the learner-controlled conditions; time spent on the listening task was used as the covariate. No moderating effect was found, F (3, 103) = .77, MSE = .60, p = .51, η_p^2 = .02. Similarly, a one-way ANCOVA was conducted to evaluate whether the number of times subjects clicked on the play button had a moderating effect on format of presentation of information for the learner-controlled conditions; number of times subjects clicked on the play button was used as the covariate. No moderating effect was found, F (3, 103) = .24, MSE = .19, p = .87, η_p^2 < .01.

Task demand. A two-way ANOVA was conducted to evaluate the potential effects of learner control and format of presentation of information on the subjects' ratings of task demand (i.e. mental and physical activity required to complete the learning task). Results indicated that neither of the two main effects was significant; for learner control, F (1, 194) = 1.31, MSE = 4.34, p = .25, $\eta_p^2 < .01$; for format of presentation of information, F (2, 194) = .18, MSE = .62, p = .82, $\eta_p^2 < .01$. Additionally, there was a non-significant interaction, F (2, 194) = .50, MSE = 1.66, p = .61, $\eta_p^2 < .01$. A separate two-way ANCOVA was conducted to evaluate whether the time subjects spent on the vocabulary section had a moderating effect on the independent variables; time spent on the vocabulary section was used as the covariate. No significant moderating effect of time spent on the vocabulary section was found for learner control, F (1, 188) = .84, MSE = 2.73, p = .36, $\eta_p^2 < .01$; format of presentation of information, F (2, 188) = 1.53, MSE = 4.99, p = .22, η_p^2 = .02; or their interaction F (2, 188) = 2.74, MSE = 8.91, p = .07, η_p^2 = .03, power = .53. Along these lines, a one-way ANCOVA was conducted to evaluate whether the time subjects spent on the listening task had any moderating effect on format of presentation of information for the learner-controlled conditions; time spent on the listening task was used as the covariate. No moderating effect was found, F (3, 103) = 1.95, MSE = 5.38, p = .13, η_p^2 = .06. Additionally, a one-way ANCOVA was conducted to evaluate whether the number of times subjects clicked on the play button had any moderating effect on format of presentation of information for the learner-controlled conditions; number of times subjects clicked on the play button was used as the covariate. No moderating effect was found, F (3, 103) = 1.00, MSE = 2.71, p = .42, η_p^2 = .03.

Effort. A two-way ANOVA was conducted to evaluate the potential effects of learner control and format of presentation of information on the subjects' ratings of effort (i.e. work necessary to understand the contents of the learning environment). Neither of the two main effects was found to be significant; for learner control, F (1, 194) = .06, MSE = .20, p = .81, η_p^2 < .001; for format of presentation of information, F (2, 194) = 1.11, MSE = 3.92, p = .33, η_p^2 = .01. Moreover, there was a non-significant interaction, F (2, 194) = 1.12, MSE = 3.96, p = .33, η_p^2 = .01. A separate two-way ANCOVA was conducted to evaluate whether the time subjects spent on the vocabulary section had a moderating effect on the independent variables; time spent on the vocabulary section was used as the covariate. No significant moderating effect of time spent on the vocabulary section was found for learner control, F (1, 188) = .19, MSE = .66, p = .67, η_p^2 < .01; format of presentation of information, F (2, 188) = 1.63, MSE = 5.76, p = .20, η_p^2 = .02; or their interaction F (2, 188) = .58, MSE = 2.06, p = .56, η_p^2 < .01. Along these lines, a one-way ANCOVA was conducted to evaluate whether the time subjects spent on the listening task had a moderating effect on format of presentation of information for the learner-controlled conditions; time spent on the listening task was used as the covariate. No moderating effect was found, F (3, 103) = .70, MSE = 2.33, p = .55, η_p^2 = .02. Furthermore, a one-way ANCOVA was conducted to evaluate whether the number of times subjects clicked on the play button had a moderating effect on format of presentation of information for the learner-controlled conditions; number of times subjects clicked on the play button was used as the covariate. No moderating effect was found, F (3, 103) = .43, MSE = 1.44, p = .73, η_p^2 = .01.

Feeling of success. A two-way ANOVA was conducted to evaluate the potential effects of learner control and format of presentation of information on the subjects' ratings of feeling of success (i.e. how successful the subject felt in understanding the contents of the learning environment). Results showed that neither of the two main effects was significant; for learner control, F (1, 194) < .001, MSE < .001, p = .99, η_p^2 < .001; for format of presentation of information, F (2, 194) = .01, MSE = .03, p = .99, η_p^2 < .001. In addition, a non-significant interaction was found, F (2, 194) = .71, MSE = 2.05, p = .49, η_p^2 = .01. A separate two-way ANCOVA was conducted to evaluate whether the time subjects spent on the vocabulary section had a moderating effect on the independent variables; time spent on the vocabulary section was used as the covariate. No significant moderating effect of time spent on the vocabulary section was found for learner control, F (1, 188) = 2.73, MSE = 7.85, p = .10, η_p^2 = .01; format of presentation of information, F (2, 188) = .95, MSE = 2.72, p = .39, η_p^2 = .01; or their interaction F (2, 188) = .77, MSE = 2.22, p = .46, η_p^2 < .01. Subsequently, a one-way ANCOVA was conducted to evaluate whether the time subjects spent on the listening task had any moderating effect on format of presentation of information for the learner-controlled conditions; time spent on the listening task was used as the covariate. No moderating effect was found, F (3, 103) = .47, MSE = 1.58, p = .70, η_p^2 = .01. Additionally, a one-way ANCOVA was conducted to evaluate whether the number of times subjects clicked on the play button had any moderating effect on format of presentation of information for the learner-controlled conditions; number of times subjects clicked on the play button was used as the covariate. No moderating effect was found, F (3, 103) = 1.75, MSE = 5.62, p = .16, η_p^2 = .05.

Navigation. A two-way ANOVA was conducted to evaluate the potential effects of learner control and format of presentation of information on the subjects' ratings of navigation (i.e. effort invested to navigate the learning environment). Neither of the two main effects was found to be significant; for learner control, F (1, 194) =.33, MSE = .89, p = .56, η_p^2 < .01; for format of presentation of information, F (2, 194) = .46, MSE = 1.23, p = .63, η_p^2 < .01. There was a non-significant interaction, F (2, 194) = 1.25, MSE = 3.35, p = .29, η_p^2 = .01. A separate two-way ANCOVA was conducted to evaluate whether the time subjects spent on the vocabulary section had a moderating effect on the independent variables; time spent on the vocabulary section was used as the covariate. No significant moderating effect of time spent on the vocabulary section was found for learner control, F (1, 188) = .13, MSE = .34, p = .73, η_p^2 = .001; format of presentation of information, F (2, 188) = .46, MSE = 1.24, p = .63, η_p^2 < .01; or their interaction F (2, 188) = 1.29, MSE = 3.49, p = .28, η_p^2 = .01. Furthermore, a one-way ANCOVA was conducted to evaluate whether the time subjects spent on the listening task had a moderating effect on format of presentation of information for the learner-controlled conditions; time spent on the listening task was used as the covariate. No moderating effect was found, F (3, 103) = .13, MSE = .26, p = .94, η_p^2 < .01. Additionally, a one-way ANCOVA was conducted to evaluate whether the number of times subjects clicked on the play button had a moderating effect on format of presentation of information for the learner-controlled conditions; the number of times subjects clicked on the play button was used as the covariate. No moderating effect was found, F (3, 103) = .34, MSE = .67, p = .80, η_p^2 = .01.

Frustration level. A two-way ANOVA was conducted to evaluate the potential effects of learner control and format of presentation of information on the subjects' ratings of frustration (i.e. insecurity, discouragement, irritation, stress, and annoyance felt during the learning task). Results indicated that neither of the two main effects was significant; for learner control, F (1, 194) < .01, MSE = .01, p = .96, η_p^2 < .001; for format of presentation of information, F (2, 194) = 1.03, MSE = 2.78, p = .36, η_p^2 = .01. However, the interaction was significant, F (2, 194) = 3.06, MSE = 8.27, p < .05, η_p^2 = .03, power = .59 (see Figure 10).

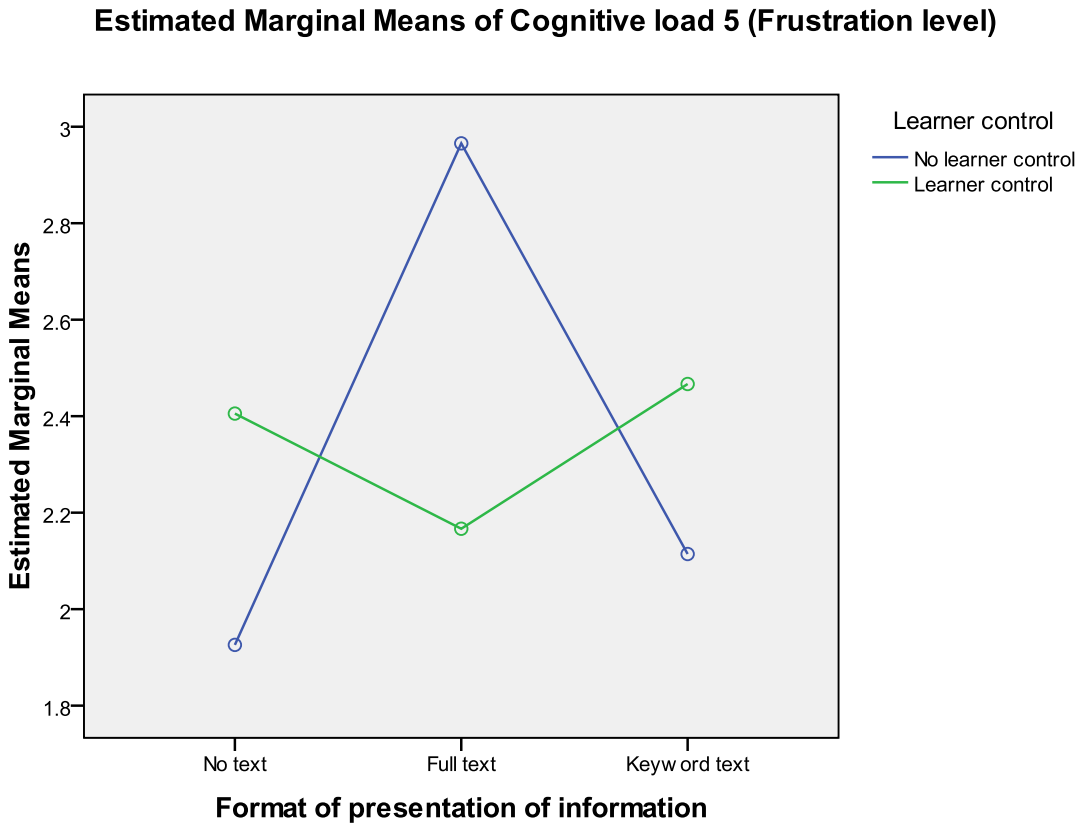

Figure 10. Plot of adjusted means of frustration level showing a significant interaction between learner control and format of presentation of information.

A two-way ANCOVA was conducted to evaluate whether the time subjects spent on the vocabulary section had a moderating effect on the independent variables; time spent on the vocabulary section was used as the covariate. No significant moderating effect of time spent on the vocabulary section was found for learner control, F (1, 188) = 1.35, MSE = 3.51, p = .25, $\eta_p^2 < .01$. A significant moderating effect of time spent on the vocabulary section was found for format of presentation of information, F (2, 188) = 3.14, MSE = 8.18, $p < .05$, $\eta_p^2 = .03$. The interaction was not significant, F (2, 188) = 1.29, MSE = 3.49, p = .28, $\eta_p^2 = .01$. In addition, a one-way ANCOVA was conducted to evaluate whether the time subjects spent on the listening task had a moderating effect on format of presentation of information for the learner-controlled conditions; time spent on the listening task was used as the covariate. No moderating effect was found, F (3, 103) = .63, MSE = 1.62, p = .60, $\eta_p^2 = .02$. Moreover, a one-way ANCOVA was conducted to evaluate whether the number of times subjects clicked on the play button had any moderating effect on format of presentation of information for the learner-controlled conditions; number of times subjects clicked on the play button was used as the covariate. No moderating effect was found, F (3, 103) = .03, MSE = .07, p = .99, $\eta_p^2 = .001$.

Attitudes

Depending on the treatment condition, a questionnaire comprised by 10 to 12 items was used to measure the attitudes of the subjects towards the treatments. Each item consisted of a question and a five-point Likert scale ranging from *strongly agree* (5) to *strongly disagree* (1); values in parentheses were assigned for data analysis purposes. Questionnaire items along with means and standard deviations are presented in Table 4.

Table 4

Descriptive Statistics of Attitude Measures

		NLC			LC		
		n[a]	*M*	*SD*	*n*[a]	*M*	*SD*
1. I liked this activity.	AN	27	4.63	.49	37	4.59	.50
	AF	29	4.72	.46	42	4.69	.47
	AK	35	4.51	.56	30	4.43	.68
2. The activity was well organized.	AN	27	4.70	.54	37	4.68	.58
	AF	29	4.90	.31	42	4.76	.43
	AK	35	4.77	.49	30	4.53	.68
3. The instructions in the activity were easy to follow.	AN	27	4.96	.19	37	4.73	.56
	AF	29	4.83	.47	42	4.71	.55
	AK	35	4.89	.32	30	4.77	.57
4. The content was relevant for me to learn English.	AN	27	4.19	.88	37	4.22	.75
	AF	29	4.59	.68	42	4.19	.77
	AK	35	4.20	.72	30	4.07	.87
5. The story was interesting to me.	AN	27	3.67	.78	37	3.89	.77
	AF	29	4.14	.74	42	3.98	.75
	AK	35	3.71	.83	30	3.73	.83
6. The activity helped me learn words I did not know.	AN	27	4.33	.83	37	4.59	.64
	AF	29	4.69	.54	42	4.64	.62
	AK	35	4.49	.61	30	4.60	.50
7. The activity helped me learn the sounds of new words.	AN	27	4.78	.42	37	4.68	.53
	AF	29	4.66	.55	42	4.76	.43
	AK	35	4.69	.47	30	4.73	.52
8. Seeing text of the words I heard in the listening task helped me learn the sounds of words[b].	AN	27	4.44	.64	37	4.38	.72
	AF	29	4.55	.57	42	4.43	.63
	AK	35	4.40	.55	30	4.60	.56
9. I enjoyed seeing text of the words I heard in the listening task[c].	AN	0			0		
	AF	29	4.48	.63	42	4.33	.72
	AK	35	4.49	.56	30	4.30	.79
10. Repeat and pause controls in the listening task helped me learn the sounds of words[d].	AN	27	4.37	.69	37	4.62	.59
	AF	29	4.38	.68	42	4.45	.59
	AK	35	4.29	.83	30	4.40	.77

		NLC			LC		
		n^a	*M*	*SD*	n^a	*M*	*SD*
11. I enjoyed using the repeat and pause controls in the listening task[e].	AN	0			37	4.51	.65
	AF	0			42	4.21	.75
	AK	0			30	4.13	.90
12. I would like to use more activities like this one to learn the sounds of words.	AN	27	4.96	.19	37	4.86	.35
	AF	29	4.90	.31	42	4.83	.44
	AK	35	4.83	.38	30	4.77	.50

Note. *M* = Mean. *SD* = Standard Deviation. 1=strongly disagree, 5=strongly agree. NLC = Not-Learner Control. LC = Learner Controlled. AN = Audio + No Text. AF = Audio + Full Text. AK = Audio + Keyword Text

[a]Sample size within conditions.

[b]For audio + no text conditions, this item read "*Seeing text of the words I heard in the listening task would have helped me learn the sounds of words*".

[c]Item displayed to participants in the full text and keyword text conditions.

[d]For not-learner-controlled conditions, this item read "*Repeat and pause controls in the listening task would have helped me learn the sounds of words*".

[e]Item displayed to participants in the learner-controlled conditions.

An exploratory factor analysis (EFA) was performed to determine the underlying structure of latent variables for further analysis. As mentioned earlier, depending on the treatment conditions, participants responded between 10 and 12 attitude items. Only those items that were responded by the totality of participants in the study were included in the factor analysis (i.e. items 9 and 11 were analyzed separately).

Given the variations in wording for items 8 and 10 in accordance to specific treatment conditions, separate one-way ANOVAs were conducted on the results of these items to define if wording influenced the participants' responses. Treatment condition (1 - 6) served as the independent variable in both cases. For item 8, the comparison between audio + no text conditions vs. audio + text conditions was non-significant, F (1, 194) = .78, MSE = .30, p = .38, $\eta_p^2 < .01$. As for item 10, the comparison between not-learner controlled conditions vs. learner controlled conditions was non-significant, F (1, 194) =

2.19, *MSE* = 1.04, p = .14, η_p^2 = .01. Given that item wording did not influence participant responses to items 8 and 10, such items were then included in the analysis.

The EFA was conducted using maximum likelihood as the extraction method and a varimax (orthogonal) rotation of 10 of the 12 Likert type items from the attitude questionnaire, using data from all the participants. The data were suitable for structure detection, as indicated by Kaiser-Meyer Olkin measure of sampling adequacy (KMO = .79) and Bartlett's test of sphericity < .001. Following the recommendations made by Costello and Osborne (2005), a value of .32 was considered as the minimum loading for a given item, and a minimum of three items loading onto a given factor set the criterion to keep it or discard it. Moreover, a value of .50 was used to indicate strong loading of an item onto a factor, with at least five of these indicating factor solidity. Item 3 was the only one that loaded onto a second factor and was then removed from the analysis. Results of the final orthogonal rotation are presented in Table 5.

The analysis yielded a one-factor solution with a simple structure. Five of these items' loads were greater than or equal to .50, indicating a solid factor (Costello & Osborne, 2005). This factor was labeled "Attitudes towards the treatments".

A two-way ANOVA was conducted to evaluate the potential effects of learner control and format of presentation of information on the factor "Attitudes towards the treatments". Means of the items comprising this factor were used as the dependent variable. No significant main effect of learner control, *F* (1, 194) = .07, *MSE* = .01, p = .79, η_p^2 = .03, or format of presentation of information, *F* (2, 194) = 2.82, *MSE* = .34, p = .06, η_p^2 = .03, power = .55 were found. There was a non-significant interaction, *F* (2, 194) = .65, *MSE* = .08, p = .52, η_p^2 = .01. However, a significant difference was found between

the keyword text vs. full text groups, (p = .02). Means and standard deviations are presented in Table 6.

Table 5

Summary of Factor Loadings of a One-Factor Solution for Student Attitudes

Item	Factor 1
1. I liked this activity.	.50
2. The activity was well organized.	.44
3. The instructions in the activity were easy to follow[a].	
4. The content was relevant for me to learn English.	.55
5. The story was interesting to me.	.53
6. The activity helped me learn words I did not know.	.61
7. The activity helped me learn the sounds of new words.	.56
8. Seeing text of the words I heard in the listening task helped me learn the sounds of words.	.43
9. I enjoyed seeing text of the words I heard in the listening task[a].	
10. Repeat and pause controls in the listening task helped me learn the sounds of words.	.40
11. I enjoyed using the repeat and pause controls in the listening task[a].	
12. I would like to use more activities like this one to learn the sounds of words.	.37

[a]Excluded from the exploratory factor analysis.

Table 6

Descriptive Statistics of Factor "Attitudes towards the treatments"

		n^a	*M*	*SD*
Not learner controlled	AN	27	4.45	.34
	AF	29	4.61	.29
	AK	35	4.43	.35
Learner controlled	AN	37	4.50	.35
	AF	42	4.53	.32
	AK	30	4.43	.41

Note. *M* = Mean. *SD* = Standard Deviation. AN = Audio + No Text. AF = Audio + Full Text. AK = Audio + Keyword Text.
[a]Sample size within conditions.

Regarding the items that were not a part of the EFA, a two-way ANOVA was conducted to evaluate the potential effects of learner control and format of presentation of information on the subjects' attitudes about item 3 (i.e. *The instructions in the activity were easy to follow*). Results indicated that there was a significant main effect of learner control, F (1, 194) = 5.20, MSE = 1.18, p = .02, η_p^2 = .03, power = .62. In addition, the main effect of format of presentation of information was not significant, F (2, 194) = .45, MSE = .10, p = .64, η_p^2 < .01, and a non-significant interaction was found, F (2, 194) = .32, MSE = .07, p = .72, η_p^2 < .01. As for item 9, a two-way ANOVA was conducted to evaluate the potential effects of learner control and format of presentation of information on the subjects' attitudes regarding this item (i.e. *I enjoyed seeing text of the words I heard in the listening task*). No significant main effect of learner control, F (1, 132) = 2.00, MSE = .93, p = .16, η_p^2 = .02, or format of presentation of information, F (1, 132) = .02, MSE = .01, p = .90, η_p^2 < .001, were found. There was a non-significant interaction, F

(1, 132) = .02, *MSE* = .01, p = .88, η_p^2 < .001. In regards to item 11, a one-way ANOVA was conducted to evaluate the potential effects of format of presentation of information on the subjects' attitudes about this item (i.e. *I enjoyed using the repeat and pause controls in the listening task*), for subjects in the corresponding treatment conditions. No significant main effect was found, F (2, 106) = 2.42, *MSE* = 1.41, p = .09, η_p^2 = .04. Yet there was a significant difference between the keyword text vs. no text groups, (p < .05).

Open-ended questions. The participants responded to three open-ended questions asking about their opinion regarding (1) what would help them better associate written words with their corresponding sounds, (2) what they liked best about the activity, and (3) what could be done to improve it. Multiple participants described more than one aspect of the questions. Qualitative content analysis (Babbie, 2013; Berg, 2001; Russ-Eft & Preskill, 2009) was used to derive emerging categories of manifest content from the open-ended responses, using themes as the unit of analysis. The procedure for content analysis consisted in the inductive, systematic, and objective identification of emerging thematic categories, of a mutually exclusive nature, that were generated through recognition of meaningful patterns and then labeled according to the substance of the data they contained. This, in turn, established the selection criteria through which responses of the participants were classified, according to the characteristics of their content, and included into the corresponding thematic category by tallying the frequencies of relevant themes. Additionally, representative pieces of text were selected during the analysis to further explain and exemplify each category. Frequencies of responses per category are shown in Tables 7, 8, and 9.

Table 7

Responses to Item 13 "What would help you better associate written words with their corresponding sounds?" (n = 200*)*

	Not-learner controlled			Learner controlled			
Category	AN	AF	AK	AN	AF	AK	Total
Concurrent audio and text	23	18	19	22	31	18	131
More practice	1	8	9	5	2	2	27
Visuals	2	5	6	6	3	3	25
Repeated listening	2	3	1	6	2	4	18
Knowing the meaning of words	0	0	4	3	2	2	11
None	0	1	1	0	3	1	6

Note. AN = Audio + No Text. AF = Audio + Full Text. AK = Audio + Keyword Text.

Five categories emerged from the analysis of this item that asked students what would help them better associate written words with their sounds. The category with the most frequent mentions by the participants was the concurrent use of audio and text. Responses provided by the participants include comments indicating that they consider the combination of concurrent text and audio to be advantageous in the association of aural and written forms of words. Participants made statements such as "What would help to facilitate learning are activities like this one to associate in a better way the written form and the sound. Also, there is a great advantage: we would be practicing two aspects at a time", "To see the word at the same time as I listen to it…", and "Practicing with programs like this one where, at the moment of reading the word, I'm listening to it, too". The second most prevalent category referred to an increase in the amount of time allocated for language practice. Participants produced comments like "Undoubtedly,

practicing the English language more. I consider that it is by means of practicing and familiarizing with words that the association with the respective sound is facilitated" and "Using the aural form of the words repeated times in a dialogue". The next arisen category was the use of visuals (e.g. images, pictures, drawings, animations, etc.) related to the target word being studied. Comments in this category include "Maybe images although the written word and the sound being listened to helped me a lot to learn to pronounce and write" and "To include images with each sound". The following category identified is related to the use of repeated listening of the words under study. An example of a statement made by a participant regarding this theme was "To listen to the words several times until I have them clear". Additionally, some participants indicated that knowing the meaning of the words would help them associate their written and aural forms. A comment produced by a participant was "Knowing their meaning or relating them with phrases". Finally, it must be noted that a few participants provided responses that were not applicable to the question.

In addition, a chi-square test of independence was performed to examine the relationship between learner control and format of presentation of information for the most recurrent category (i.e. concurrent audio and text). The relationship between the independent variables was not significant, X^2 (2, N=131) = 2.59, p = .27. A chi-square test of independence was not conducted on the remaining categories due to the sample size for each of them being insufficiently large to conduct the test without violating the sample size assumption.

Table 8

Responses to Item 14 "What did you like best about the activity?" (n = 200)

	Not-learner controlled			Learner controlled			
Category	AN	AF	AK	AN	AF	AK	Total
Associate written words with sounds	20	16	18	27	18	19	118
Instructional design	4	8	10	7	16	8	53
Concurrent audio and text	0	6	7	0	11	5	29
Activity was fun / interesting / dynamic	4	3	1	4	5	6	23
Learning new words	3	2	0	7	5	2	19
Learner control	0	0	0	6	4	3	13
Audio replay capability (pre and posttest)	4	1	1	2	1	1	10
Everything	2	2	0	2	0	0	6

Note. AN = Audio + No Text. AF = Audio + Full Text. AK = Audio + Keyword Text.

Eight categories emerged from the analysis of item 14, which was used to ask participants what they liked best about the activity. The most frequent category was the association of written forms of words with their sound. Examples of responses provided by the participants are "Being able to associate the written words with their sound because it allowed me to better know which word corresponded to the audio that I was listening to" and "To associate and differentiate the sounds of written words". The second category identified was the instructional design. Students indicated that they liked the didactics of the activity as well as the navigation features and content. Comments identified in this category include "It's very didactic and easy to understand", "The way the activity was designed", "The examples and controls", "The instructions are in Spanish, which makes it easier to understand what you're being asked", and "Everything

was perfectly explained and it's a very interactive way to learn". The third most prevalent category was the use of concurrent text and audio. Participants made statements such as "What I liked the most was seeing and of course listening to the words and their sounds. It is interesting this activity because one learns more with the audio and reading", "…seeing the word and at the same time listening how it is pronounced and, above all, using those same words within a defined text…" and "When I listened to the audio. I think it was really good to be able to see the words of the story on the screen". The fourth most liked category was that the activity was fun / interesting / dynamic. Comprising this category are statements such as "Everything is very interesting and it is an easy and fun way to learn English", "The dynamics. You learn and it is very interactive. You have fun while learning. I loved it" and "I liked that it is fun. It is not the typical, boring way of learning, on the contrary, it is interesting and creative". For the next emerged category, the participants indicated that they liked best having learnt new words. Some examples of phrases used by the participants are "To learn new words and to know how to pronounce them correctly" and "That I learnt the meanings of words I did not know". The subsequent category was the use of learner control for those participants in the corresponding treatment conditions. The following are comments made by the participants regarding this theme: "Being able to reproduce the phrases or words as many times as I wanted so I could understand better", "I could repeat the audio as many times as I wanted and that helped me know how to differentiate the pronunciation" and "They gave you the sound and the phrase and you could listen to it as many times as you wanted!". The penultimate category corresponded to the capability of replaying the audio up to three times on the pre and posttest sections. Some representative phrases provided

by the participants are “We could repeat the pronunciation of the words more than once, in case I was unsure about their pronunciation” and “I could replay several times the sound of a word”. Finally, there were a few students who stated that they liked everything about the activity. Examples in this category include “In general, I liked everything. It was a dynamic and fun activity”, “Everything was very cool because I learnt the pronunciation of words I didn’t know…” and “Everything in general. It was very interesting”.

In addition, a chi-square test of independence was performed to examine the relationship between learner control and format of presentation of information for the two most recurrent categories. The relationship between the independent variables was not significant for either the association of written words with sounds, X^2 (2, N=118) = .34, p = .84, or instructional design, X^2 (2, N=53) = 2.24, p = .33. A chi-square test of independence was not conducted on the remaining categories due to the sample size for each of them being insufficiently large to conduct the test without violating the sample size assumption.

Table 9

Responses to Item 15 "What could be done to improve the activity?" (n = 200*)*

	Not-learner controlled			Learner controlled			
Category	AN	AF	AK	AN	AF	AK	Total
No changes	9	11	12	14	13	15	74
Extended content	6	6	11	9	12	8	52
Visuals	3	8	4	10	8	4	37
Learner control	7	4	12	0	0	0	23
Slower narration	0	0	0	4	4	1	9
Simultaneous text	3	0	0	4	0	1	8
Immediate feedback	1	1	0	0	3	2	7
Use activities like this one more often	1	1	1	0	1	0	4

Note. AN = Audio + No Text. AF = Audio + Full Text. AK = Audio + Keyword Text.

Regarding the item about what could be done to improve the activity, eight categories emerged from the analysis. The most prevalent theme was that no changes were needed for the activity. Some representative comments made by the participants about this category were "I think it is very well designed, I don't think it needs anything" and "In my opinion, the activity is very complete and well structured". The second most frequent reaction was related to the addition of content; this category included the increase in the number of examples and practice exercises. Instances of statements related to this category are "Increase the number of exercises", "Increase the length of the activities", and "That the person who is doing this activity could also write the words being listened to and, in doing so, she would improve her writing". The third most recurrent category was the inclusion of visuals in the form of pictures, animation, or

video. The following are some of the comments that constituted this category: “I would recommend the utilization of images to provide us with an idea of what it is being talked about. Like the words, maybe we knew them but I would say that it would help better identify the sound”, “To add video”, and “To include some images related to the content of the text”. The fourth category has to do with the inclusion of learner control for those treatment conditions that lacked this feature. Statements made by the participants regarding this theme are “I think it is ok as it is. I would only add some controls to pause the narration because unknown words are difficult for us to understand” and “Maybe the implementation of a narration stopper to better listen to the audio”. The next emerged theme was related to the reduction of the narration speed. Some examples of comments in this category include “That the narration was slower to better understand what it is about and the words it says” and “That the pronunciation in the narration was slower”. It is of note that this category appeared from the comments made by the participants in the learner controlled conditions. The next thematic category identified was the simultaneous use of text and audio for the treatment conditions where no text was displayed during the narration; a representative statement of this was “I think that, when we listen to the narration, the text of the narration appeared on the screen”. A participant in the learner controlled + keyword condition indicated a preference for the entire narration to be presented on screen. The second to last arisen category was the implementation of immediate feedback after each item of the tests. An example of a statement identified in this category was “…that when we select an answer, the program marks it as either correct or incorrect”. The final category identified during the analysis is the desire to use activities like this one more often. An example of a comment made by a participant in

this category was “This activity was good. I would like to do more activities like this one because I believe it motivates the students more, especially those of us who find it difficult to understand this language”.

Additionally, a chi-square test of independence was performed to examine the relationship between learner control and format of presentation of information for the two most recurrent categories. The relationship between the independent variables was not significant for either the no changes category, X^2 (2, N=74) = .24, p = .89, or extended content, X^2 (2, N=52) = 2.41, p = .30. A chi-square test of independence was not conducted on the remaining categories due to the sample size for each of them being insufficiently large to conduct the test without violating the sample size assumption.

CHAPTER 4

DISCUSSION

The purpose of this study was to investigate the effects of learner control as well as the use of concurrent audio and different types of equivalent onscreen text on the ability of EFL learners to associate the textual and aural forms of target vocabulary words. The study also looked at reported levels of cognitive load and attitudes towards experimental treatments in the context of a computer-based multimedia instructional program. The results revealed the following findings: (a) groups in the audio + keyword text conditions outperformed those in the audio + no text and audio + full text conditions on text-sound association, (b) within the audio + keyword text conditions, the learner-controlled group outperformed the not-learner-controlled group on text-sound association, (c) within the learner-controlled conditions, the audio + keyword group outperformed the audio + no text and audio + full text groups on text-sound association, and (d) a redundancy effect was not found for any treatment condition. Implications, limitations, and future directions are discussed next within the frameworks of cognitive load theory and cognitive theory of multimedia learning.

Text-Sound Association

In learning a second or foreign language of a non-phonetic nature, the learner commonly encounters that written forms of words do not correspond to their aural forms. As a consequence, inference of word pronunciation from its written form and vice versa may be challenging for some learners. Along these lines, this study was concerned with looking at the effects on text-sound association of the use of concurrent audio and equivalent on-screen text in combination with learner control.

According to the initial results, no significant main effect of either learner control or format of presentation of information on text-sound association was found when holding pretest scores constant. However, when pretest scores were grouped by low and high prior knowledge, the results revealed that format of presentation of information had a significant effect on posttest scores, with a small to medium effect size and relatively low power. There was no significant effect of learner control. Although the interaction between the factors was not significant, a higher power design may lead to significant results regarding this matter. In keeping with the findings, participants in the keyword text conditions significantly outperformed those in the no text and full text conditions. Similarly, the learner-controlled group significantly outperformed the not-learner-controlled group within the keyword text condition. Furthermore, within the learner-controlled conditions, the keyword text group significantly outperformed both the no text and full text groups. There were no moderating effects of the variables time spent on the vocabulary introduction, time on the listening task and number of times a participant replayed the audio (i.e. clicked on the play button).

Outcomes are consistent with those of the study conducted by Mayer and Johnson (2008) in which groups presented with short redundant on-screen text significantly outperformed those in the no text condition on retention but not on transfer tests. Conversely, the results of the current study were obtained from a transfer test to ensure that learning, (i.e. text-sound association) was measured through the use of contents different from those presented during instruction, as suggested by Sweller et al. (2011) regarding studies about the use of concurrent text and verbatim audio. The use of single keywords displayed during the narration may have had a positive effect on the subjects'

performance regarding text-sound association given that no other visual was competing for their attention. Visuals (e.g. images, animations, etc.) were not included as part of the materials used in this study with the purpose of not overloading the participants' visual channel. As it occurred in other studies (Mayer & Moreno, 2003; Moreno & Mayer, 2002), results support the statement that students learn better when redundant text is presented, provided that the visual channel is not overloaded by processing simultaneous on-screen text and animation, confirming what Moreno and Mayer (2002) denominate as a verbal redundancy effect. Moreover, the single keywords may have facilitated the input selection process described by Chapelle (1998) in the interactionist model of second language acquisition, by signaling the learner to focus her attention on a specific feature of the narration (i.e. identifying the keyword) and relating the word on the screen with the audio and, as said by Vanderplank (1988), allowing the learner to consciously monitor the narration and identify target words within the stream of the speech. In turn, this also corroborates the argument that second language learning is facilitated by the use of multimedia (Meskill, 1996; Plass & Jones, 2005). Likewise, findings endorse those of other SLA studies (Bird and Williams, 2002; Hayati and Mohmedi, 2011), in which conclusions indicate that the use of simultaneous audio and equivalent keyword text may bolster text-sound association of target vocabulary in a foreign language, considering that one modality may compensate for deficiencies in the other, in cases where the written form of the input cannot be inferred reliably from sound (Bird and Williams, 2002) or in the other direction, as in the case of this study.

As previously mentioned, the learner-controlled group significantly outperformed the not-learner-controlled group within the keyword text condition and the keyword text

group significantly outperformed both the no text and full text groups within the learner-controlled conditions in posttest scores. Results suggest that participants in the learner-controlled + audio + keyword condition benefitted from the presence of learner control presumably due to having a higher level of prior knowledge than subjects in the other treatment conditions, endorsing the view that recommends learners with low prior knowledge or lower ability to be provided with higher levels of learner control (Kopcha & Sullivan, 2008; Merrill, 2002). Furthermore, the combination of learner control and keywords in the audio sequence may have allowed the subjects to discriminate between target words to focus on and the rest of the narration, comparable to the argument proposed by Meskill (1996) referring to the benefits of language identification through the concurrence of learner control and language presented in more than one modality.

Overall, findings suggest that performance on text-sound association of basic-level EFL learners improved due to a presumable signaling effect provided by the presentation of equivalent keyword text and concurrent audio during a narration in English. Additionally, the combination of the aforementioned factors along with the use of learner control ostensibly contributed to improved performance on text-sound association of those learners with a higher level of prior knowledge. The latter finding to some extent contradicts the view of Vanderplank (1988), which indicated that the concurrent use of verbatim (i.e. full) text may be of limited value for learners with a lower level of prior knowledge and may be of benefit for more advanced learners.

Cognitive Load

Results indicate that no significant effects of learner control and format of presentation of information were found for overall level of cognitive load, task demand,

effort, feeling of success, or navigation; neither were there moderating effects of the variables time spent on the vocabulary introduction, time on the listening task and number of times a participant replayed the audio. Nevertheless, a significant interaction between learner control and format of presentation of information was present for frustration level, with a small to medium effect size and relatively low power, indicating that the effect of learner control on frustration level varied as a function of format of presentation of information and vice versa. Similarly, time spent on the vocabulary section varied as a function of format of presentation of information and contrariwise, indicating a moderating effect on frustration level. Additionally, a higher power design may lead to significant results regarding the moderating effects of time spent on the vocabulary section on the interaction between the two independent variables for overall level of cognitive load and task demand. No moderating effects of time on the listening task and number of times a participant replayed the audio were found.

Albeit not statistically significant in its majority, results of the analysis of cognitive load measures yielded useful information. Overall, the group in the not-learner-controlled + audio + full text condition reported the highest level of cognitive load, presumably due to the subjects having to listen to the entire narration accompanied by verbatim text. Comparable results were obtained by Diao, Chandler, and Sweller (2007) and Diao and Sweller (2007) although, unlike the cited studies, a redundancy effect was not present in the current study. On the other hand, the lowest level of cognitive load was reported by the not-learner-controlled + audio + no text condition. This may have occurred because the subjects dealt only with aural information processed through the auditory channel without an overload of the visual channel, as suggested by Mayer and

Moreno (2003). In addition, the absence of learner control ostensibly contributed to the minimization of cognitive load.

A redundancy effect occurs when two simultaneously presented sources of information are intelligible to the learner in isolation (Sweller et al, 2011). As stated by Plass and Jones (2005), the negative impact of the redundancy effect may be less pronounced when it comes to learning of a second or foreign language. In the case of this study, assuming there was a learning gap to overcome by means of administering the instructional treatments, the two sources in isolation (i.e. audio and text) presumably were not completely intelligible for the learners, therefore producing a high level of element interactivity. According to the results, no redundancy effect was found for any treatment condition. The use of concurrent text and audio may not have interfered with text-sound association since significant amounts of extraneous cognitive load were not present or sufficient to produce such an effect because the combination of intrinsic and extraneous cognitive load did not exceed working memory capacity. This was ostensibly due to the instructional design of the treatments that kept cognitive load within the limits of working memory by not overloading the visual channel with simultaneous on-screen text and competing visuals (Mayer & Moreno, 2003; Moreno & Mayer, 2002). Additionally, presentation of target vocabulary in an isolated fashion prior to the listening task may have contributed to the reduction of cognitive load, as suggested by Mayer and Chandler (2001), particularly because the subjects, regardless of treatment condition, were given control over the pace of the vocabulary section.

Findings contradict those of the study conducted by Diao and Sweller (2007) in which, after having found a redundancy effect, they concluded that presenting novice

EFL/ESL learners with both written and spoken material was not optimal, especially for those who had difficulties with associating written and aural forms of words. Similarly, Diao, Chandler, and Sweller (2007) reported a redundancy effect for EFL learners that were presented with concurrent audio and verbatim text. These contradictory results may be due to the effects of the use of verbatim (i.e. full) text, in the case of the two studies that reported a redundancy effect, whereas results of the current study showed evidence that the use of redundant keyword text facilitated text-sound association of novice EFL learners.

Attitudes

Overall, the subjects tended to have positive attitudes towards the treatments, as indicated by the answers they provided in the attitude questionnaire. About the instructional activity, the majority of participants in all treatment conditions specified that they liked best to associate the written and aural forms of words and acknowledged the use of concurrent audio and text as the preferred instructional approach to help them better associate the written and aural forms of words in EFL. These findings are comparable to those of the study conducted by Borrás & Lafayette (1994) in which participants that were presented with video and equivalent text had significantly more positive attitudes towards the multimedia instructional program than participants that were not presented with on-screen text. Additionally, Vanderplank (1988) observed that EFL learners had more positive attitudes towards the use of video and equivalent text, along with reduced or non-existing levels of anxiety which, in turn, appeared to foster confidence.

Concerning the analysis of the one-factor model, although no significant main effects of the independent variables were found between groups, it is worth mentioning that the non-significance of the effect of format of presentation of information is likely due to relatively low power; a significant effect of this factor might be found with a higher power design. Nevertheless, there was a difference between the keyword text and full text groups, the latter being significantly higher. This finding suggests that subjects in the audio + full text conditions felt more positive towards the use of concurrent audio and verbatim text than those in the audio + keyword text conditions, showing that subjects may have felt more comfortable by having the support of verbatim text on screen during the entire narration. Additionally, as mentioned before, attitudes do not necessarily reflect performance and this finding somewhat contradicts the results of the analysis of text-sound association in which subjects in the audio + keyword text conditions outperformed, although not significantly, to those in the audio + full text conditions.

Similarly, regarding the analysis of item 11 (i.e. *I enjoyed using the repeat and pause controls in the listening task)* for subjects in the corresponding conditions, no significant main effect of format of presentation of information was found. Nonetheless, significant differences were found between the audio + keyword text and audio + no text groups, indicating that subjects in the latter had significantly more positive attitudes towards the use of learner control as presented in the instructional program. This finding presumably indicates that subjects in the audio + no text condition felt more comfortable—and perhaps confident—by having the possibility of controlling the number of times they could listen to the narration, or specific segments of it, given that no concurrent text was presented on-screen and, as mentioned before, subjects tended to

favor the use of concurrent audio and text to facilitate text-sound association. This finding somewhat supports the study conducted by Vanderplank (1988), in which EFL learners appeared to show lower confidence when presented with video in the absence of text.

Moreover, regarding item 3 (i.e. *The instructions in the activity were easy to follow)*, significant differences were found between the not-learner-controlled and the learner-controlled conditions, showing that participants in the former groups had more positive perceptions of the instructions presented in the materials than those in the latter groups. Even though attitudes towards this item across groups were close to the highest possible level of satisfaction, results indicate that the implementation of learner control in the instructional program was presumably perceived as more elaborated by the subjects in the corresponding conditions due to them having more controls at their disposal to interact with.

Limitations

Results obtained are limited to populations whose characteristics are similar to those of the sample in the current study. Potential limitations regarding text-sound association results include a small to medium effect size of format of presentation of information, accompanied by relatively low power. Additionally, reliability of the measures used to assess this variable may pose another restriction since Chronbach's alphas for the pre and posttest were .06 and .37, respectively; this was likely due to a high level of difficulty of the tests. Typically accepted values of Chronbach's alpha in the social sciences are those greater than .70 (Cortina, 1993).

Moreover, potential limitations of the results corresponding to cognitive load include the small to medium effect size and relatively low power of the interaction found between learner control and format of presentation of information on frustration level. In addition, results of cognitive load must be interpreted with caution given the number of statistical tests run on the overall level of cognitive load and each separate measure of NASA-TLX, taking into account that this may have increased the probability of committing a type I error.

Additionally, potential limitations of the significant attitude results found include a small to medium effect size of learner control and relatively low power in what refers to item 3 of the attitude questionnaire (i.e. *The instructions in the activity were easy to follow).*

Implications

According to findings, the groups presented with text in the form of keywords significantly outperformed those in the no text and full text conditions on text-sound association. For the field of instructional design, these findings implicate that the use of audio along with concurrent equivalent text in the form of keywords does not necessarily yield a redundancy effect, considering that the visual channel is not overloaded by processing visual stimuli other than keyword text as part of the instructional activity. In the arena of foreign language learning, findings suggest that basic-level learners benefit from the use of the simultaneous presentation of narration and onscreen text in the form of keywords, when referring to learning of a foreign language of a non-phonetic nature (e.g. English), since a redundancy effect is not present when learners are nonnative speakers of the language only if, for such learners, the two identical sources of

information are not completely intelligible in isolation. Moreover, the integration of both sources of information fosters the association of visual and aural forms of words.

Future Directions

More research is still needed to investigate the effects of the use of concurrent audio and text (full and keyword) for learners of EFL of different levels of expertise (e.g. basic, intermediate, advanced, etc.), along with the effects of the use of learner control on text-sound association and the analysis of their relationship with the redundancy effect.

Future research also includes the study of lengthier narrations as another factor to consider while designing future experiments regarding text-sound association, given that the current study focused on the use of a short narration only. Examination of the effects of narration length as a moderator on text-sound association may enrich the findings here presented.

In addition, future studies comprise the use of physiological measures, such as electroencephalography (see Antonenko & Niederhauser, 2010), to complement the measurement of cognitive load as an instantaneous reflection of the mental effort invested during completion of learning tasks. Along these lines, physiological measures may also supplement the analysis of anxiety levels that, according to Vanderplank (1988), EFL learners tend to experience in the absence of on-screen text wile dealing with narration or video. The study of the effects of learner control on anxiety levels of EFL learners is also considered as a part of future research work.

REFERENCES

Aesop (1834). The fox in the well. In *Aesop's fables for the instruction and improvement of youth* (pp. 20-21). London: J. Clements.

Alessi, S. M. & Trollip, S. R. (2001). *Multimedia for learning*. Needham Heights, MA: Allyn and Bacon.

Antonenko, P. D., & Niederhauser, D. S. (2010). The influence of leads on cognitive load and learning in a hypertext environment. *Computers in Human Behavior, 26*(2), 140-150.

Ayres, R. (2002). Learner attitudes towards the use of CALL. *Computer Assisted Language Learning, 15*(3), pp. 241-249.

Babbie, E. (2013). *The basics of social research*. Belmont, CA: Cengage Learning.

Baddeley, A. (1992). Working memory. *Science, 255*(5044), 556-559.

Baltova, I. (1999). Multisensory language teaching in a multidimensional curriculum: The use of authentic bimodal video in core French. *Canadian Modern Language Review, 56*(1), 31-48.

Berg, B. L. (2001). *Qualitative research methods for the social sciences*. Needham Heights, MA: Allyn & Bacon.

Bird, S. A., & Williams, J. N. (2002). The effect of bimodal input on implicit and explicit memory: An investigation into the benefits of within-language subtitling. *Applied Psycholinguistics*, *23*, 509–533.

Borrás, I., & Lafayette, R. C. (1994). Effects of multimedia courseware subtitling on the speaking performance of college students of French. *Modern Language Journal, 78*(1), 61-75.

British Council (n.d.). *Frequently asked questions: The English language*. Retrieved from http://www.britishcouncil.org/learning-faq-the-english-language.htm

Chapelle, C. A. (1998). Multimedia CALL: Lessons to be learned from research on instructed SLA. *Language Learning and Technology, 2*(1), 22-34.

Clark, R.C., Nguyen, R, & Sweller, J. (2006). *Efficiency in learning*. San Francisco, CA: Jossey-Bass Pfeiffer.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of applied psychology*, *78*(1), 98.

Costello, A. B., & Osborne, J. W. (2005). Exploratory Factor Analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation*, *10*(7), 1-9.

Cottam, M. E. (2010). *The effects of visual and textual annotations on Spanish listening comprehension, vocabulary acquisition and cognitive load*. (Doctoral thesis, Arizona State University, Arizona, United States). Retrieved from http://129.219.247.59/attachments/56050/content/Cottam_asu_0010E_10102.pdf

Diao, Y., Chandler, P., & Sweller, J. (2007). The effects of written text on comprehension of spoken English as a foreign language. *American Journal of Psychology*, *120*(2), 237-261.

Diao, Y., & Sweller, J. (2007). Redundancy in foreign language reading comprehension instruction: Concurrent written and spoken presentations. *Learning and Instruction*, *17*, 77-88.

Garza, T. J. (1991). Evaluating the use of captioned video material in advanced foreign language learning. *Foreign Language Annals*, *24*(3), 239-258.

Gerjets, P., Scheiter, K., & Catrambone, R. (2004). Designing instructional examples to reduce intrinsic cognitive load: Molar versus modular presentation of solution procedures. *Instructional Science*, *32*, 33–58.

Gerjets, P., Scheiter, K., & Catrambone, R. (2006). Can learning from molar and modular worked examples be enhanced by providing instructional explanations? *Learning and Instruction*, *16*, 104-121.

Gerjets, P., Scheiter, K., Opfermann, M., Hesse, F. W., & Eysink, T. H. S. (2009). Learning with hypermedia: The influence of representational formats and different levels of learner control on performance and learning behavior. *Computers in Human Behavior*, *25*, 360-370.

Gibbs, C. (2009). *Learner-controlled captioning: A new frontier? Exploring the impact of learner control on the development of listening skills in a multimedia environment* (Doctoral thesis, Concordia University, Quebec, Canada). Retrieved from http://search.proquest.com.ezproxy1.lib.asu.edu/pagepdf/577055293?accountid=4485

Hart, S. G., & Staveland, L. E. (1998). Development of NASA-TLX (task load index): Results of empirical and theoretical research. In P. Hancock (Ed.), *Human mental workload* (pp. 139-183). Amsterdam: North Holland.

Hasler, B. S., Kersten, B., & Sweller, J. (2007). Learner control, cognitive load and instructional animation. *Applied Cognitive Psychology, 21*(6), 713-729.

Hatcher, L. (2013). *Advanced statistics in research: Reading, understanding, and writing up data analysis results*. ShadowFinch Media, LLC.

Hayati, A, & Mohmedi, F. (2011). The effect of films with and without subtitles on listening comprehension of EFL learners. *British Journal of Educational Technology*, *42*(1), 181-192.

Heift, T. (2002). Learner control and error correction in ICALL: browsers, peekers, and adamants. *CALICO Journal*, *19*(2), 295-313.

Heller, I. (2005). Learner experiences and CALL-tool usability—Evaluating the Chemnitz InternetGrammar. *Computer Assisted Language Learning*, *18(*1), pp. 119-142.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38*(1), 23-31.

Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors, 40*(1), 1-17.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology, 13*, 351-371.

Kalyuga, S., Chandler, P., & Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology, 92*(1), 126-136.

Kalyuga, S., Chandler, P., & Sweller, J. (2004). When redundant on-screen text in multimedia technical instruction can interfere with learning. *Human Factors, 46*(3), 567-581.

Koolstra, C. M., & Beentjes, J. W. J. (1999). Children's vocabulary acquisition in a foreign language through watching subtitled television programs at home. *Educational Technology Research and Development*, *47*(1), 51-60.

Kopcha, T. J. & Sullivan, H. (2008). Learner preferences and prior knowledge in learner-controlled computer-based instruction. *Educational Technology Research and Development*, *56*(3), 265-286.

Kraiger, K. & Jerden, E. (2007). A meta-analytic investigation of learner control: Old findings and new directions. In S. M. Fiore & E. Salas (Eds.), *Toward a science of distributed learning* (pp. 65-90). Washington, D.C.: American Psychological Association.

Leahy, W., Chandler, P., & Sweller, J. (2003). When auditory presentations should and should not be a component of multimedia instruction. *Applied Cognitive Psychology*, *17*, 401–418.

Leahy, W., & Sweller, J. (2011). Cognitive load theory, modality of presentation and the transient information effect. *Applied Cognitive Psychology*, *25*(6), 943-951.

Lewandowski, L. J., & Kobus, D. A. (1993). The effects of redundancy in bimodal word processing. *Human Performance*, *6*(3), 229-239.

Long, M. H. (1996). The role of linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia, (Eds.), *Handbook of second language acquisition* (pp. 413-468). San Diego, CA: Academic Press.

Markham, P. (1999). Captioned videotapes and second-language listening word recognition. *Foreign Language Annals*, *32*(3), 321-328.

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31-48). New York, NY: Cambridge University Press.

Mayer, R. E. (2009). *Multimedia learning*. New York, NY: Cambridge University Press.

Mayer, R. E. (2010). Research-based solutions to three problems in web-based training. In W. J. Kozlowski & E. Salas (Eds.), *Learning, training, and development in organizations* (pp. 203-227). New York, NY: Taylor and Francis Group.

Mayer, R. E. (2011). Applying the science of learning to multimedia instruction. In J. P. Mestre & B. H. Ross (Eds.), *The psychology of learning and motivation* (pp. 77-108). San Diego, CA: Elsevier.

Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages?. *Journal of educational psychology*, *93*(2), 390.

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology*, *93*(1), 187-198.

Mayer, R. E., & Johnson, C, I. (2008). Revising the redundancy principle in multimedia learning. *Journal of Educational Psychology*, *100*(2), 380-386.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, *38*(1), 43-52.

Merrill, M. D. (1974, March). *Learner control: Beyond aptitude treatment interactions*. Paper presented at the Annual Meeting of the Association for Educational Communications and Technology, Atlantic City, NJ. Retrieved from http://eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_ &ERICExtSearch_SearchValue_0=ED095822&ERICExtSearch_SearchType_0= no&accno=ED095822

Merrill, M. D. (1980). Learner control in computer-based learning. *Computers and Education, 4*, 77-95.

Merrill, M. D. (1984). What is learner control? In R. K. Bass & C. D. Dills (Eds.), *Instructional development: The state of the art II* (pp. 221–242) Dubuque, IA: Kendall Hunt Pub Co.

Merrill, M. D. (2002). Instructional strategies and learning styles: Which takes precedence? In R. A. Reiser & J. V. Dempsey (Eds.), *Trends and issues in instructional technology*. Upper Saddle River, NJ: Merrill Prentice Hall.

Meskill, C. (1996). Listening skills development through multimedia. *Journal of Educational Multimedia and Hypermedia*, *5*(2), 179-201.

Montali, J., & Lewandowski, L. (1996). Bimodal reading: Benefits of a talking computer for average and less skilled readers. *Journal of Learning Disabilities*, *29*(3), 271-279.

Moreno, R., & Mayer, R. E. (2002). Verbal redundancy in multimedia learning: When reading helps listening. *Journal of Educational Psychology*, *94*(1), 156-163.

Moussa, J. (2008). *The impact of spoken English on learning English as a foreign language: A cognitive load perspective* (Doctoral thesis, University of New South Wales, Sidney, Australia). Retrieved from http://www.unsworks.unsw.edu.au/primo_library/libweb/action/dlDisplay.do?vid=UNSWORKS&docId=unsworks_3879

Neuman, S. B., & Koskinen, P. (1992). Captioned television as comprehensible input: Effects of incidental word learning from context for language minority students. *Reading Research Quarterly, 27*(1), 94-106.

Nyberg, L., Habib, R., McIntosh, A. R., & Tulving, E. (2000). Reactivation of encoding-related brain activity during memory retrieval. *Proceedings of the National Academy of Sciences, 97*(20), 11120-11124.

Paivio, A. (1986). *Mental representations*. New York, NY: Oxford University Press.

Plass, J., & Jones, L. (2005). Multimedia learning in second language acquisition. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 467-488). New York, NY, US: Cambridge University Press.

Plass, J., Moreno, R., & Brünken R. (2009). Introduction. In J. Plass, R. Moreno, & R. Brünken (Eds.), *Cognitive load theory* (pp. 1-6). New York, NY: Cambridge University Press.

Russ-Eft, D. F. & Preskill, H. (2009). *Evaluation in organizations: A systematic approach to enhancing learning, performance, and change*. New York, NY: Basic Books.

Shaffer, J. P. (1986). Modified sequentially rejective multiple test procedures. *Journal of the American Statistical Association, 81*(395), 826-831.

Schwan, S., & Riempp, R. (2004). The cognitive benefits of interactive videos: Learning to tie nautical knots. *Learning and Instruction, 14*(3), 293-305.

Shen, H. H. (2010). Imagery and verbal coding approaches in Chinese vocabulary instruction. *Language Teaching Research, 14*(4), 485-499.

Stepp-Greany, J. (2002). Student perceptions on language learning in a technological environment: Implications for the new millennium. *Language Learning & Technology, 6*(1), pp. 165-180.

Sweller, J. (1998). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257-285.

Sweller, J. (2005). The Redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 159-168). New York, NY: Cambridge University Press.

Sweller, J. (2010). Cognitive load theory: Recent theoretical advances. In J. Plass, R. Moreno, & R. Brünken (Eds.), *Cognitive load theory* (pp. 29-47). New York, NY: Cambridge University Press.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. doi:10.1007/978-1-4419-8126-4

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

Wheeler, M. E., Petersen, S. E., & Buckner, R. L. (2000). Memory's echo: Vivid remembering reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences*, *97*(20), 11125-11129.

Wiebe, G., & Kabata, K. Students' and instructors' attitudes toward the use of CALL in foreign language teaching and learning. *Computer Assisted Language Learning 23*(3), pp. 221–234.

Yeh, S., & Lehman, J. D. (2001). Effects of learner control and learning strategies on English as a foreign language (EFL) learning from interactive hypermedia lessons. *Journal of Educational Multimedia and Hypermedia*, *10*(2), 141-159.

Zhao, Y. (1997). The effects of listeners' control of speech rate on second language comprehension. *Applied linguistics*, *18*(1), 49-68.

APPENDIX A

TEXT PASSAGE FOR THE LISTENING TASK

A fox falls in a well and sticks his claws into the sides of the well to keep his head above the water level. Soon after, a wolf comes and peeps over the well's brink. The fox asks the wolf for help, entreating that the wolf throws him a rope or something of that kind to favor his escape. The wolf is moved with compassion at the misfortune of the fox and does not forbear expressing his concern: "Poor fox! I am sorry for you with all my heart for being in this somber condition". The fox replies: "No, my dear friend! If you wish me well, do not pity me. Instead, lend me some succor as fast as you can. Pity is just useless comfort when I am really close to starvation or drowning".

Note. Passage adapted from the fable *The fox in the well* (Aesop, 1834).

APPENDIX B

WORD CLASSIFICATION FORM FOR INSTRUCTORS

Word classification form for instructors

Directions: To the best of your knowledge, please classify the words in the list by circling the response that best matches your opinion according to the following criteria:

1. Level of difficulty (1 = easiest to 5 = most difficult) it takes to the students in your course to learn each word.
2. Word knowledge: known (K) or unknown (U) by EFL students in your course?

Word	**Difficulty** *(1 = easiest to 5 = most difficult)*	**Knowledge** *Known (K) / Unknown (U)*
Adjectives		
Close	1 2 3 4 5	K U
Dear	1 2 3 4 5	K U
Fast	1 2 3 4 5	K U
Poor	1 2 3 4 5	K U
Somber	1 2 3 4 5	K U
Sorry	1 2 3 4 5	K U
Useless	1 2 3 4 5	K U
Nouns		
Brink	1 2 3 4 5	K U
Claws	1 2 3 4 5	K U
Comfort	1 2 3 4 5	K U
Compassion	1 2 3 4 5	K U
Concern	1 2 3 4 5	K U
Condition	1 2 3 4 5	K U
Drowning	1 2 3 4 5	K U
Escape	1 2 3 4 5	K U
Fox	1 2 3 4 5	K U
Friend	1 2 3 4 5	K U
Head	1 2 3 4 5	K U
Heart	1 2 3 4 5	K U
Help	1 2 3 4 5	K U
Kind	1 2 3 4 5	K U
Level	1 2 3 4 5	K U
Misfortune	1 2 3 4 5	K U
Pity	1 2 3 4 5	K U
Rope	1 2 3 4 5	K U
Sides	1 2 3 4 5	K U
Starvation	1 2 3 4 5	K U
Succor	1 2 3 4 5	K U
Water	1 2 3 4 5	K U
Well	1 2 3 4 5	K U
Wolf	1 2 3 4 5	K U

Word	**Difficulty** *(1 = easiest to 5 = most difficult)*	**Knowledge** *Known (K) / Unknown (U)*
Prepositions		
Above	1 2 3 4 5	K U
After	1 2 3 4 5	K U
At	1 2 3 4 5	K U
For	1 2 3 4 5	K U
In	1 2 3 4 5	K U
Into	1 2 3 4 5	K U
Of	1 2 3 4 5	K U
Over	1 2 3 4 5	K U
To	1 2 3 4 5	K U
With	1 2 3 4 5	K U
Verbs		
Am	1 2 3 4 5	K U
Asks	1 2 3 4 5	K U
Being	1 2 3 4 5	K U
Can	1 2 3 4 5	K U
Comes	1 2 3 4 5	K U
Do	1 2 3 4 5	K U
Does	1 2 3 4 5	K U
Entreating	1 2 3 4 5	K U
Expressing	1 2 3 4 5	K U
Falls	1 2 3 4 5	K U
Favor	1 2 3 4 5	K U
Forbear	1 2 3 4 5	K U
Is	1 2 3 4 5	K U
Keep	1 2 3 4 5	K U
Lend	1 2 3 4 5	K U
Moved	1 2 3 4 5	K U
Peeps	1 2 3 4 5	K U
Pity	1 2 3 4 5	K U
Replies	1 2 3 4 5	K U
Sticks	1 2 3 4 5	K U
Throws	1 2 3 4 5	K U
Wish	1 2 3 4 5	K U

APPENDIX C

WORD CLASSIFICATION FORM FOR STUDENTS

Word classification form for students

Directions: To the best of your knowledge, please classify the words in the list as either known or unknown to you by circling the response that best matches your opinion.

Word	**Knowledge** *Known (K) / Unknown (U)*		**Word**	**Knowledge** *Known (K) / Unknown (U)*	
Adjectives			***Prepositions***		
Close	K	U	Above	K	U
Dear	K	U	After	K	U
Fast	K	U	At	K	U
Poor	K	U	For	K	U
Somber	K	U	In	K	U
Sorry	K	U	Into	K	U
Useless	K	U	Of	K	U
Nouns			Over	K	U
Brink	K	U	To	K	U
Claws	K	U	With	K	U
Comfort	K	U	***Verbs***		
Compassion	K	U	Am	K	U
Concern	K	U	Asks	K	U
Condition	K	U	Being	K	U
Drowning	K	U	Can	K	U
Escape	K	U	Comes	K	U
Fox	K	U	Do	K	U
Friend	K	U	Does	K	U
Head	K	U	Entreating	K	U
Heart	K	U	Expressing	K	U
Help	K	U	Falls	K	U
Kind	K	U	Favor	K	U
Level	K	U	Forbear	K	U
Misfortune	K	U	Is	K	U
Pity	K	U	Keep	K	U
Rope	K	U	Lend	K	U
Sides	K	U	Moved	K	U
Starvation	K	U	Peeps	K	U
Succor	K	U	Pity	K	U
Water	K	U	Replies	K	U
Well	K	U	Sticks	K	U
Wolf	K	U	Throws	K	U
			Wish	K	U

APPENDIX D

TARGET VOCABULARY WORDS

1. Claws
2. Forbear
3. Peeps
4. Pity
5. Poor
6. Rope
7. Starvation
8. Succor
9. Throw
10. Useless

APPENDIX E

TEXT-SOUND ASSOCIATION PRETEST

1. Claws
   Class
   Clues
   Close

2. For beer
   Fur bear
   Forbore
   Forbear

3. Petes
   Peeps
   Peps
   Pets

4. Pretty
   Pie-tee
   Pity
   Potty

5. Půr
   Pure
   Poor
   Power

6. Rho-pee
   Rope
   Rupee
   Rap

7. Starvation
   Stur-vation
   Steer-vation
   Stare-vation

8. Succor
   Sacker
   Sook-ore
   Sock-ore

9. Truck
   Throb
   Trow
   Throw

10. Us-less
    Useless
    Oos-less
    Juice-less

APPENDIX F

TEXT-SOUND ASSOCIATION POSTTEST

1. The old cat has big class
   The old cat has big close
   The old cat has big clues
   The old cat has big claws

2. I can't forbore seeing the light
   I can't forbear seeing the light
   I can't fur-bear seeing the light
   I can't for-beer seeing the light

3. She pets at her mom's dog
   She petes at her mom's dog
   She peeps at her mom's dog
   She peps at her mom's dog

4. It is a potty to replace them
   It is a pity to replace them
   It is a pie-tee to replace them
   It is a pretty replace them

5. The poor people in that house
   The power people in that house
   The pure people in that house
   The půr people in that house

6. We use a rap to tie them
   We use a rho-pee to tie them
   We use a rope to tie them
   We use a rupee to tie them

7. Stare-vation is a big problem here
   Stur-vation is a big problem here
   Steer-vation is a big problem here
   Starvation is a big problem here

8. You give sook-ore to your friends
   You give succor to your friends
   You give sock-ore to your friends
   You give sack-er to your friends

9. They throb a pen to Rose
   They throw a pen to Rose
   They trow a pen to Rose
   They truck a pen to Rose

10. It is useless to keep trying it
    It is juice-less to keep trying it
    It is oos-less to keep trying it
    It is us-less to keep trying it

APPENDIX G

COGNITIVE LOAD MEASURES

	Item	Measure
1.	How much mental and physical activity (thinking, deciding, remembering, looking, searching, etc.) was required to complete the learning task? Was the task easy or demanding? *Easy* ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ *Demanding*	Task demand
2.	How hard did you have to work in order to understand the contents of the learning environment? *Not hard at all* ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ *Very hard*	Effort
3.	How successful did you feel in understanding the contents of the learning environment? *Not successful at all* ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ *Very successful*	Feeling of success
4.	How much effort did you have to invest to navigate the learning environment (deciding which link/button takes you to a different section)? *Very little effort* ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ *Very much effort*	Navigation
5.	How insecure, discouraged, irritated, stressed, and annoyed did you feel during the learning task? *Not at all* ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ *Very much*	Frustration level

Note. Measures adapted from the NASA Task Load Index (NASA-TLX, Hart & Staveland, 1998).

APPENDIX H

ATTITUDE QUESTIONNAIRE

Attitude questionnaire

Part 1

Directions: Select the response that best matches your opinion according to the following key

SA=Strongly agree A=Agree N=Neutral D=Disagree SD=Strongly disagree

1.	I liked this activity.	*SA*	*A*	*N*	*D*	*SD*
2.	The activity was well organized.	*SA*	*A*	*N*	*D*	*SD*
3.	The instructions in the activity were easy to follow.	*SA*	*A*	*N*	*D*	*SD*
4.	The content was relevant for me to learn English.	*SA*	*A*	*N*	*D*	*SD*
5.	The story was interesting to me.	*SA*	*A*	*N*	*D*	*SD*
6.	The activity helped me learn words I did not know.	*SA*	*A*	*N*	*D*	*SD*
7.	The activity helped me learn the sounds of new words.	*SA*	*A*	*N*	*D*	*SD*
8.	Seeing text of the words I heard in the listening task helped me learn the sounds of words[a].	*SA*	*A*	*N*	*D*	*SD*
9.	I enjoyed seeing text of the words I heard in the listening task[b].	*SA*	*A*	*N*	*D*	*SD*
10.	Repeat and pause controls in the listening task helped me learn the sounds of words[c].	*SA*	*A*	*N*	*D*	*SD*
11.	I enjoyed using the repeat and pause controls in the listening task[d].	*SA*	*A*	*N*	*D*	*SD*
12.	I would like to use more activities like this one to learn the sounds of words.	*SA*	*A*	*N*	*D*	*SD*

Part 2

Directions: Answer the following questions.

13. What would help you better associate written words with their corresponding sounds?
14. What did you like best about the activity?
15. What could be done to improve the activity?

Note. Some items were adapted from the attitude survey used by Cottam (2010).

[a]For audio + no text conditions, this item read "*Seeing text of the words I heard in the listening task would have helped me learn the sounds of words*".

[b]Item displayed to participants in the full text and keyword text conditions.

[c]For not-learner-controlled conditions, this item read "*Repeat and pause controls in the listening task would have helped me learn the sounds of words*".

[d]Item displayed to participants in the learner-controlled conditions.

APPENDIX I

IRB APPROVAL

ASU Knowledge Enterprise Development

Office of Research Integrity and Assurance

To:	Brian Nelson BYENG
From:	Mark Roosa, Chair Soc Beh IRB
Date:	09/21/2012
Committee Action:	**Exemption Granted**
IRB Action Date:	09/21/2012
IRB Protocol #:	1209008246
Study Title:	Effects of text, audio, and learner control on text-sound association and cognitive load of EFL learners

The above-referenced protocol is considered exempt after review by the Institutional Review Board pursuant to Federal regulations, 45 CFR Part 46.101(b)(1) .

This part of the federal regulations requires that the information be recorded by investigators in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects. It is necessary that the information obtained not be such that if disclosed outside the research, it could reasonably place the subjects at risk of criminal or civil liability, or be damaging to the subjects' financial standing, employability, or reputation.

You should retain a copy of this letter for your records.