The Effects of Visual and Textual Annotations on Spanish Listening Comprehension, Vocabulary Acquisition and Cognitive Load

By

Michael Evan Cottam

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

Approved November, 2010 by the
Graduate Supervisory Committee:

Wilhelmina Savenye, Chair
James D. Klein
Robert Atkinson

ARIZONA STATE UNIVERSITY

December 2010

ABSTRACT

The purpose of this experimental study was to investigate the effects of textual and visual annotations on Spanish listening comprehension and vocabulary acquisition in the context of an online multimedia listening activity. 95 students who were enrolled in different sections of first year Spanish classes at a community college and a large southwestern university were randomly assigned to one of four versions of an online multimedia listening activity that contained textual and visual annotations of several key words. Students then took a comprehension and vocabulary posttest and a survey to measure cognitive load and general attitudes towards the program.

Results indicated that textual annotations had a significant positive effect on listening comprehension and that visual annotations had a significant positive effect on how successful students felt. No statistically significant differences were found for other variables. Participants also reported positive attitudes towards vocabulary annotations and expressed a desire to see more annotations during multimedia listening activities of this type. These findings provide further evidence of the impact that multimedia may have on language acquisition.

These findings have implications for multimedia design and for future research. Language listening activities should include a variety of vocabulary annotations that may help students to understand what they hear and to help them learn new vocabulary. Further research is needed outside of the laboratory, in the online and increasingly-mobile language learning environment in order to align the research with the environment in which many students currently study. The

incorporation of motivation into multimedia learning theory and cognitive load should be explored, as well as new measures of cognitive load.

DEDICATION

To my wife, Corena, and my children, JuliAn, Loren, Samuel, Emily, Caitlyn, and Andrew, and to my father and my now-departed mother: thank you for being patient with me, listening to me, and giving me your love and support. We have been through births, deaths and everything in between over the past few years. Without you, I never would have made it this far.

ACKNOWLEDGEMENTS

To my committee, thank you. Dr. Savenye, your years of support in the labor of completing this research have strengthened me and made me a better scholar. Your friendship and all of our long, shared conversations over stacks of paper are invaluable to me. Dr. Klein and Dr. Atkinson, I appreciate your teachings, your experience and your insights. Throughout my years in the Ed Tech program, you have helped kindle a fire for learning in me that will burn throughout my life. You probably do not know how much you have taught me or how many of your insights I have applied in my work as a result of your teaching. Thank you.

To my Rio Salado family, Vernon, Jennifer, Jarred, and a couple of dozen more wonderful people in the Faculty, IDS and CP, thank you for your creativity, your dedication, your loyalty, your friendship and your support. From the beginnings of my research project to the end, you have each contributed in numerous ways. You have listened to my ramblings, brainstormed ideas, critiqued, cajoled, commiserated, laughed and cried with me. You are family and there are no words for the depth of my gratitude. Thank you!

TABLE OF CONTENTS

LIST OF TABLES

TABLE Page

TABLE Page

TABLE Page

LIST OF FIGURES

FIGURE Page

CHAPTER 1

INTRODUCTION

Online learning is becoming more and more pervasive in higher education institutions. According to a recent Sloan-C report (Allen & Seaman, 2009) more than 4.6 million students were enrolled in online courses in the United States in 2008. That number represents a 17% increase over the previous year, which far exceeds the 1.2% growth rate for higher education enrollments as a whole over the same time period. The Sloan-C report also indicates that most institutions in the U.S. expect student demand for online education to grow and plan to increase their online course offerings (Allen & Seaman, 2009).

Concurrent with the overall increase in online education since the 1990's, foreign language course offerings have increased dramatically in distance learning catalogues across the country (White, 2004). White (2004) also contends that foreign language courses vary in the technology they use and the teaching and learning activities they employ. Some rely heavily on the latest technology, while others use a mix of well-established and emerging technologies. Some are offered in an asynchronous format, while others have at least some synchronous virtual meetings. However, irrespective of delivery format or technology used, the learning goals are the same.

The goals of modern foreign language instruction include development of multiple language competencies. The American Council on the Teaching of Foreign Languages (ACTFL) establishes national standards for language learning.In their standards document, the council states, "Communication is at the

heart of second language study, whether the communication takes place face-to-face, in writing, or across centuries through the reading of literature" (American Council on the Teaching of Foreign Languages, 2001, p. 3). Communication takes several forms and requires skills in reading and listening (receptive skills) as well as writing and speaking (productive skills), which are reflected in the ACTFL proficiency guidelines.

All of the language skills are equally important; however, in their research summary, Plass and Jones (2005) report that between the two receptive language skills, more multimedia research has been published on reading comprehension skills than on listening comprehension skills. Due to this apparent disparity, the current study is focused on the receptive communication skill of Spanish listening comprehension among beginning-level students. More specifically, this study includes an examination of aspects of one type of multimedia learning activity intended to improve students' Spanish listening skills.

Acquiring language skills requires extensive second-language input at an appropriate level (Krashen, 1985; Gass & Selinker, 1994; Lafford & Salaberry, 2003). Krashen states that "We acquire by understanding language that contains structure a bit beyond our current level of competence (i+1). This is done with the help of context or extralinguistic information" (1982, p. 21). Participants in this study, due to their progress in college Spanish courses, were near the novice-high level in the ACTFL proficiency scale. The proficiency guidelines at this level state, "At times, but not on a consistent basis, the Novice-High level reader may be able to derive meaning from material at a slightly higher level where

context and/or extralinguistic background knowledge are supportive" (American Council on the Teaching of Foreign Languages, 2001)

In the classroom, listening comprehension instruction commonly includes listening to the teacher, other students and recorded language samples, while at the same time observing extralinguistic cues such as body language, tone of voice, facial expressions and prepared visual aids. In a fully asynchronous online course, the main form of aural input is audio or video recordings from the instructor and from other publisher-prepared materials. Such materials sometimes lack the extralinguistic information that makes the aural input comprehensible to students. Adding multimedia components, such as videos or pictures, to accompany the words, may help to improve the comprehensibility of the language input students receive. Two key theories may help to explain why this may be so: the theory of multimedia learning and the cognitive load theory.

Multimedia Learning

Mayer, in his generative theory of multimedia learning, contends that students learn more deeply when information is presented in both verbal (written or spoken) and pictorial (illustrations, photos, animations or videos) forms (Mayer, 2001). In his description of the multimedia principle he states, "When words and pictures are both presented, students have an opportunity to construct verbal and pictorial mental models and to build connections between them. When words alone are presented, students have an opportunity to build a verbal mental model but are less likely to build a pictorial mental model and make connections between the verbal and pictorial mental models" (Mayer, 2001, p. 63).

Mayer examined the presentation and learning of mechanical systems, such as how a pump works and how lightning is formed (Mayer, 2001; Moreno & Mayer, 2000; Moreno & Mayer, 2002; Leahy, 2003). Subsequently, other researchers extended the application of multimedia learning theory to other contexts, including foreign language instruction.

Multimedia Theory and Second Language Acquisition: Reading Comprehension

Typical instructional reading tasks in language classes may not generally follow Mayer's multimedia principle and may not include illustrations or images to depict what the text describes. Chun and Plass (1996a, 1996b) were among the first researchers to consider multimedia learning theory and add multimedia elements to instructional reading materials in order to investigate their effects on second-language reading comprehension. The multimedia elements that they and subsequent researchers added to reading materials usually consisted of vocabulary annotations that could be accessed by students while reading on-screen text. Vocabulary annotations, as defined by these researchers, are in-line hypermedia glossaries that may include textual definitions or translations and pictorial illustrations.

Chun and Plass (1996a, 1996b) and Plass, Chun, Mayer, and Leutner (1998) found that annotations with both textual and visual information aided students' second-language reading comprehension and vocabulary learning more than did textual information alone. Dubois and Vial (2000) also noted that students are able to memorize words better when both textual and visual information is provided. In three separate studies, Yeh and Wang (2003), Yoshii

(2006) and Yoshii and Flaitz (2002) also found that students in a text-plus-picture annotation treatment outperformed those in text-only and picture-only treatments on vocabulary recall assessments.

Many of the preceding studies included annotations in the learners' first language and results seem to have indicated that annotations may be helpful to students during reading tasks. Yoshii (2006) expanded the research by examining different types of annotations. He investigated the effects of annotations supplied in the learners' first language compared to annotations in the students' second language. He found a significant effect for his text+picture treatment over a text-only treatment, but also found that text annotations in either language were effective for vocabulary learning.

However, some researchers have found evidence that annotations in reading comprehension activities may not always have a positive effect on comprehension. Sakar and Ercetin (2004) and Ariew and Ercetin (2004) found that students had positive attitudes towards visual annotations, but they also observed a negative effect of such annotations on reading comprehension. There is a need for more research into multimedia annotations to help clarify their effects on comprehension.

Multimedia Theory and Second Language Acquisition: Listening Comprehension

Other researchers have focused not on reading comprehension, but on students' listening comprehension skills and vocabulary learning. In his multimedia research on listening activities Brett (1997) indicated that students performed better on comprehension and vocabulary assessments if they were

presented with a multimedia listening activity rather than with audio or video alone. In a study on the effects of illustrations on TOEFL test takers' listening comprehension, Ginther (2003) found a positive effect for the presence of images as well. Jones and Plass (2002) and Plass and Jones (2005) indicated that the effects seen in reading comprehension and vocabulary acquisition studies are also present on listening comprehension tasks. Participants who accessed both verbal and visual annotations performed better on vocabulary recall as well as on listening comprehension tasks than did participants who did not access these annotations.

Cognitive Load Theory

Cognitive load theory is sometimes cited to explain effects found in multimedia research studies. This theory is concerned with working memory limitations and strategies to overcome those limitations (Sweller, 1999, 2005). It is based upon some fundamental assumptions about human cognitive architecture, long-term memory and working memory.

Long-term memory capacity is very large and plays a central role in learning (Sweller, 2005). Long-term memories are organized into schema, which are described as "cognitive constructs that allow multiple elements of information to be categorised as a single element" (Sweller 2005, p. 21). Thus, learning involves schema acquisition and practice subsequent to initial exposure can allow schema to be processed automatically rather than consciously (Sweller, 2003, 2005).

In contrast to long-term memory, working memory is very limited in capacity (about 7 plus or minus 2 items can be held in working memory at a time) and duration (items remain for only a few seconds) (Miller 1956; Sweller, 2005). Baddeley (1986, 1992, 1999) describes working memory as being made up of an executive function and two subsystems: a visio/spatial system and an auditory loop. This division of labor within memory has led other researchers, such as Penny (1989), to find that using both subsystems can increase the capacity of working memory, taking advantage of the modality effect and the split-attention effect identified by Sweller (2003).

Cognitive load theory describes three different types of load on our memory systems: intrinsic cognitive load, extrinsic cognitive load, and germane cognitive load (Sweller, 1999, 2005). Intrinsic cognitive load is created by the natural complexity of the material to be learned, while extrinsic cognitive load is characterized as that load caused by inefficient instructional design that requires energy to be spent in things other than schema acquisition (Sweller, 2005). Germane cognitive load is that load created by the effort used to create and to make schema automatic (Sweller, 2005). The goal of instruction, therefore, should be to reduce extraneous cognitive load and increase germane cognitive load.

In spite of years of study, research into cognitive load has been focused principally on the areas of math, science and technology education "for reasons of convenience" (Sweller, 1999, p. 2). However, Sweller (1999) has asserted that cognitive load theory could be generalized to non-technical, language-based

subjects as well. Indeed, within the literature in multimedia learning theory several researchers apply cognitive load theory to their findings. Plass and Jones (2005) confirm the need for more research in the area of cognitive load theory, multimedia learning theory and language acquisition.

Cottam and Savenye (2008) conducted a research study designed to be similar to that of Jones and Plass (2002) but in a different environment, with participants spread across multiple classes and locations. They examined the effects of textual and pictorial annotations on listening comprehension, vocabulary acquisition and cognitive load for online college-level Spanish students. Unlike previous studies, the 35 participants in this pilot study were not in a lab or classroom environment, but participated in a completely online environment. The inclusion of a cognitive-load measure also set this study apart from previous multimedia research into second-language acquisition or other language-based material. These studies typically included discussion on cognitive load, but did not attempt to measure it directly or indirectly (Sweller, 1999; Jones, 2004; Jones & Plass, 2002; Plass, Chun, Mayer & Leutner, 2002, Plass & Jones, 2005).

Cottam and Savenye (2008) examined two variables: text and pictorial vocabulary annotations. Participants were randomly assigned to four conditions: no-annotations, text-only-annotations, picture-only-annotations, and both-annotations. They were presented with a listening activity that prompted them to read an introduction to the topic in English and then listen to a cultural-event description in Spanish. The description was split into five segments and was

presented on five different screens with accompanying vocabulary annotations appropriate to their assigned treatment. There were seven vocabulary terms available on each screen, for a total of 35 words available in the entire presentation. Immediately following the presentation, participants completed a comprehension and vocabulary posttest and a survey to measure cognitive load.

Analyses revealed that visual annotations yielded a significant positive effect on listening comprehension and both types of annotations had a significant positive effect on students' perceptions of success, one of the cognitive-load measures. The researchers did not find a significant effect for annotations on vocabulary acquisition or other aspects of cognitive load. Nonetheless, with the limited number of participants, the significant results indicate a need for further investigation.

Research Design

The current study was designed to extend the research of Cottam and Savenye (2008) and Jones and Plass' (2002) previous research. It was designed to investigate the effects of visual and textual annotations in a multimedia listening activity on aural comprehension, vocabulary acquisition, and cognitive load using a larger sample of participants in online college-level Spanish classes. Perhaps due to the small number of participants in the Cottam and Savenye (2008) study, significant differences were only found among a few of the treatment variables. The current study will help to clarify potential effects observed in the pilot study.

Most research on multimedia theory and cognitive load theory has been conducted in laboratory settings with content such as math, statistics and well-

defined mechanical systems (Sweller, 1999). Similar to the Cottam and Savenye (2008) pilot study, this study was conducted in an actual online course environment with students enrolled in elementary-level Spanish courses at an online community college and a large public university in the southwestern United States.

The independent variables in the study were visual and textual annotations. Visual annotation consisted of two levels: visuals included or excluded. The visual annotations, when included, were pictorial representations of the vocabulary term. For instance, the key word *cuerno* (horn) was illustrated with a photograph of a bull's horn. Textual annotation also consisted of two levels: included or excluded. Textual annotations, when included, were English translations of the key words. For the word *cuerno* participants saw the word "horn" beside the keyword in Spanish. Participants in the combination treatment saw both the picture and the translation.

The dependent variables in the study were vocabulary acquisition, listening comprehension, and cognitive load. Student attitudes and time-in-program were also examined. The research questions were:

1. What are the effects of textual and visual annotations on listening comprehension and vocabulary acquisition?
2. What are the effects of textual and visual annotations on cognitive load?
3. What are the effects of textual and visual annotations on student attitudes?

CHAPTER 2

METHOD

Participants

Participants (n=95) in the study were recruited from two institutions, one community college and one large university in the southwestern United States.

Students enrolled in first-year college-level online Spanish classes at the selected community college were invited to participate in this study via email and online course announcements. All sections of Spanish at the college are completely online with no in-person or synchronous meeting requirements. These courses were offered in a 14-week, "flex schedule" format at the college, which is similar to an open entry/open exit format. This allows individual students to enroll almost any Monday of the year and their assignment due-dates are calculated according to their individual start date. Students work through the instructional Spanish content individually, not necessarily with a cohort of students on the same schedule.

Individual students enrolled in first semester Spanish (SPA101) were offered the opportunity to participate in the study as they entered the 13th week of the course. Students in second-semester Spanish (SPA102) were invited to participate in the first two weeks of class. With this language study background, students were at approximately the correct level of proficiency to benefit from the listening activity that was used in this study.

Due to limited responses from the online community college students (n=54), participants were also recruited from a large public southwestern

university (n=41). These participants were recruited by the researcher, who visited in-person classes in the final weeks of SPA101 and the first weeks of SPA102. The researcher verbally invited students to participate and handed out flyers with the URL of the online research materials. Like the community college students, the university students also completed the study outside of class time over the internet either at computer labs on campus or from home. A few students (n=7) enrolled in summer session classes went directly to the computer lab with the researcher after a classroom invitation. In these cases the researcher remained in a different section of the lab and did not offer any assistance or extra instructions to any of the participants. Again, in all cases, participants completed the study online and independently.

Although a total of 95 participants completed the listening activity, posttests, and surveys, there was a programming error that invalidated many responses to the cognitive load measures in the survey. Valid data from only 36 of the 95 participants were analyzed to measure cognitive load, while data from all 95 participants were analyzed for all other measures.

Procedures

During the 13th week of the online community college SPA101 course, all students received an invitation via email and an online course announcement to participate in the study and complete the listening activity, posttest and survey. In the university setting, the invitations were offered in-person rather than via email. Students at both institutions were informed that upon completion of the research study they would receive a $10 incentive.

Participants were randomly assigned to one of the four treatments: no-annotations, textual-definitions-only, visual-illustrations-only, and a combination of both types of annotations. Random assignment was accomplished by a computerized random number generator; each student was randomly assigned to one of the four versions of the activity upon accessing the hyperlinks in the study invitations.

Materials

Materials were similar to those used in the Cottam and Savenye (2008) study, with some adaptations to accommodate a different online delivery platform and some improvements made to the graphics within the program. The content of the listening activity was an original, researcher-written description of the Festival of San Fermín and the Running of the Bulls in Pamplona, Spain. The lead researcher was a Spanish faculty member and instructional designer. He collaborated with another Spanish faculty member at the online community college to select the topic and create a basic outline before developing the materials. The topic was chosen for its general appeal to language learners and those interested in foreign travel. The topic also prompts the use of new, unfamiliar vocabulary, vivid descriptions and memorable images. Furthermore, cultural festivals are a common topic for beginning and intermediate foreign-language courses. A complete copy of the both-annotations version of the activity is included in Appendix A.

A total of 35 key words in the listening passage were identified to receive annotation support in the activity. A complete table of the key words used in the

study is included in Appendix B. Words were selected based on the lead researcher's experience with beginning-level Spanish students and knowledge of the course content. Key words were those deemed to be more unfamiliar to students at this level of instruction and thus were more likely to require instructional support within the activity. A subset of the most unfamiliar and more concrete 25 key terms was used to assess vocabulary learning in the posttest. More abstract terms that were more difficult to effectively illustrate were not included in the vocabulary posttest. Most obvious English-Spanish cognates (words that are similar to one another in the two languages) were also excluded from the posttest.

The activity began with an introductory screen (Figure 1) containing instructions on how to navigate through the web-based program. The help option, accessible by clicking a button at the top right of the screen, was available throughout the program. Following the instructions were two screens of information about the Running of the Bulls in English. This design was similar to the design of the program used in the Jones and Plass (2002) study, which also included introductory screens in English before presenting the listening passage. These pages served as an advance organizer and were intended to activate students' existing knowledge of the topic since they may have seen or heard of this festival previously. Each screen of introductory text was accompanied by a photograph of the festival to aid in prior-knowledge activation.

Figure 1: Introduction screen

Following the introduction, students were presented with five screens of Spanish listening content. The current screen number and the total number of screens in the program appeared at the bottom of the screen so that participants always knew where they were within the program. Participants could navigate forwards and backwards through the activity or access individual pages freely, with no time or sequence restrictions on individual screens. Although participants had complete control over their movement throughout the program, a log of their elapsed time in the program was recorded for later analysis.

One difference between the instructional program used in this study and the program used by Cottam and Savenye (2008) is that the current program included a time limit for completion. In their 2008 study, time-in-program logs indicated that some participants left the program idle on a single screen for a long period of time. To remedy this situation, the current program had a time limit of

20 minutes, which is longer than all but one of the participants spent in the program during the 2008 study. At the end of 20 minutes the program automatically advanced the participant to the posttests and survey and they were not able to return to the listening activity screens. The use of a time limit on listening activities in studies of this type is not unprecedented. Jones and Plass (2002) and Plass and Jones (2005) used short time limits on the listening activities and assessments in their studies as well.

Upon advancing to a new listening screen, an image representing the topic of the segment appeared on screen and the audio narrative began to play once students clicked on the prominent play button. Each screen contained audio player controls which allowed the student to play, pause, stop and replay the narration.

Along with the audio controls, the left side of the screen included the seven key words that were heard within the segment. As seen in Figure 2, on mouse-over the selected key words highlighted, indicating that each was an active hyperlink to more information. Upon clicking a key word, an audio icon appeared which informed students that they could hear the word pronounced individually. Simultaneously, annotations of the selected key word appeared on the right side of the screen.

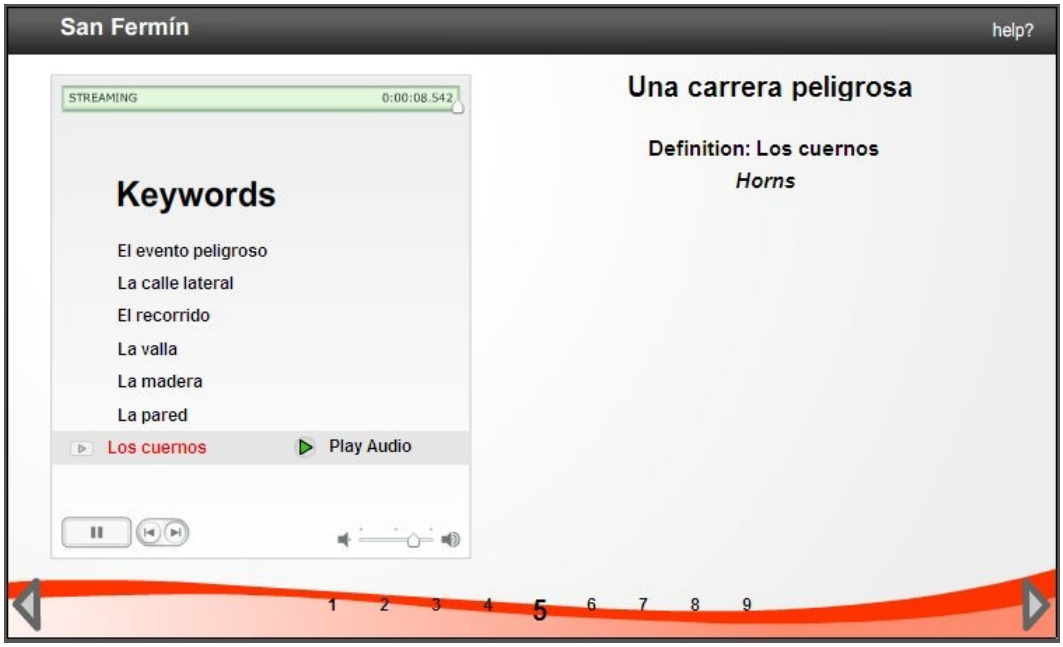


Figure 2: Listening activity screenshot for the textual definitions only treatment

There were four versions of the activity, which varied in the types of vocabulary annotations that appeared for key words of the spoken text. The variations were: 1) no annotations, 2) textual definitions only, 3) visual illustrations only, and 4) a combination of both types of annotations. Textual annotations consisted of simple English translations while visual illustrations were all photographic representations of the key words. As needed, portions of the photographs contained arrows or circles to indicate precisely which part of the photo represented the key word. For example, upon selecting a key word in the textual definitions program, the keyword and its definition appeared on the right side of the screen (Figure 2). In contrast, accessing the same key word in the visual illustrations only treatment displayed a photograph with the key word (Figure 3) and in the text and visual combination treatment the textual definition appeared along with an illustration (Figure 4). A more complete set of

screenshots illustrating the combination-annotations treatment is included in Appendix A.

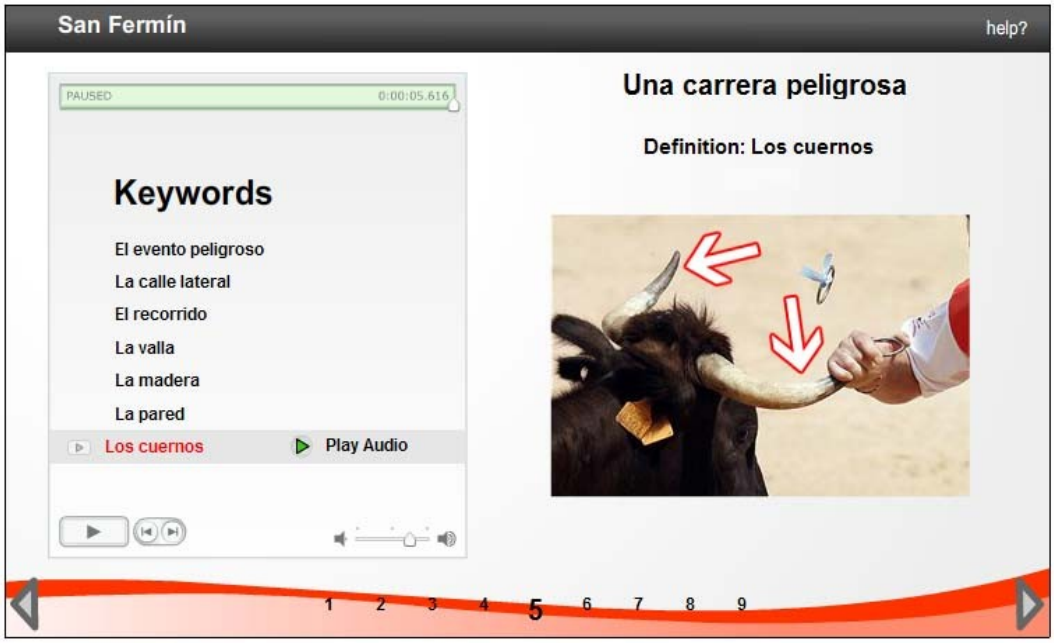


Figure 3: Listening activity screenshot for the visual illustrations only treatment

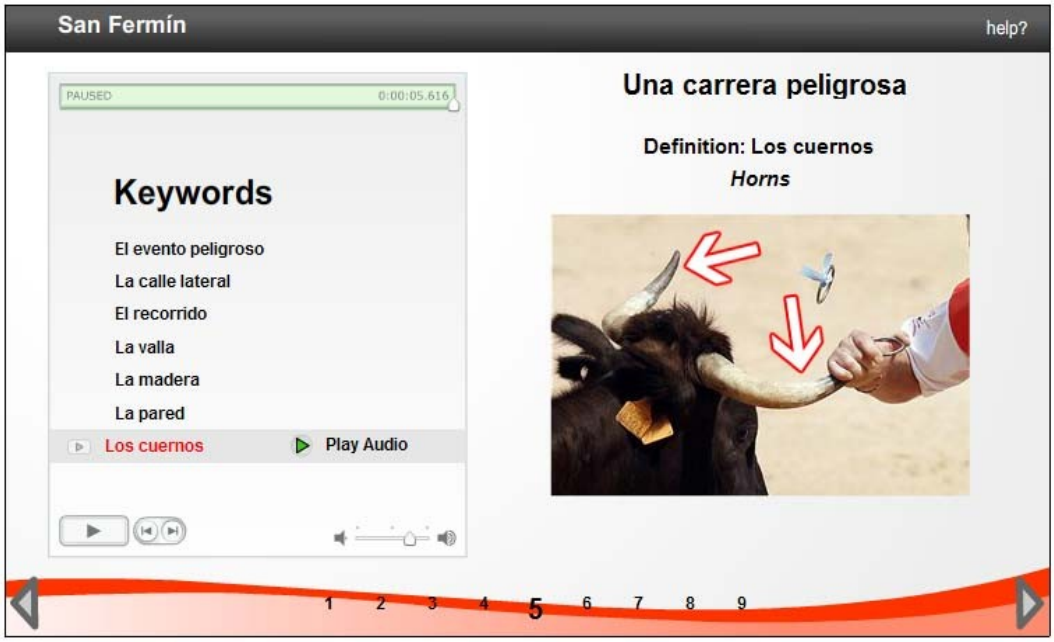


Figure 4: Listening activity screenshot for the combination textual definitions and visual illustrations treatment

After they viewed the final listening screen, participants advanced to the online posttest using on-screen navigation buttons built into the assessment tool. Once they left the listening activity screens, participants were not able to return to those screens during the posttest or survey.

Measures

Measures for the study included a comprehension and vocabulary posttest, a cognitive-load survey and an attitude survey.

The posttest consisted of one open-ended comprehension question and 25 multiple-choice vocabulary questions. Jones and Plass (2002) used similar types of listening-comprehension and vocabulary-recognition assessments in their study. All assessment items are identical to those used by Cottam and Savenye (2008). A complete copy of the comprehension posttest is included in Appendix C.

The comprehension question asked participants to "Please summarize what you have learned about San Fermines and the Running of the Bulls. Include everything you can remember and write in English. The more facts you remember, the higher your comprehension score will be." Because participants were beginning-level Spanish students, English was used to assess comprehension so that their limited Spanish language proficiency and writing ability would not interfere with the measurement of their comprehension. The researcher had previously identified 32 distinct propositions in the content of the listening activity and participant responses were evaluated according to the number of propositions identified. A similar method of assessing comprehension was used by Chun and Plass (1996), Jones and Plass (2002), Plass, Chun and Leutner (2003), Jones (2004) and Cottam and Savenye (2008).

The vocabulary posttest consisted of 25 multiple-choice items. See Appendix D for a full list of the questions on the vocabulary posttest. The question stems provided a key word in Spanish and asked participants to select the correct English translation from a set of four possible answers. The translations were identical to the textual annotations provided in two of the treatments. A sample question follows:

Select the correct translation: el herido

a. belt
b. balcony
c. injury (correct answer)
d. horns

Following the vocabulary quiz, students were presented with a series of survey questions. There were five questions to assess cognitive load, 14 general

attitude questions, and three open-ended attitude questions. Appendix E contains a full list of the survey items.

The five cognitive load measurement questions were identical to those used in the NASA-TLX assessment, originally developed by Hart and Staveland (1988) to measure cognitive load. The NASA-TLX measure was selected because it is the most commonly-used measure of cognitive load (Noyes, Garland, & Robbins, 2004) and because it has a good record of validity and reliability (Hill, Iavecchia, Byers, Bittner, Zaklad, & Christ, 1992). Gerjets, Scheiter and Catrambone (2004) and Scheiter, Gerjets, and Catrambone (2006) successfully used a similarly modified version of the NASA-TLX to measure cognitive load in their research and Su (2007) followed their model in her dissertation study.

Only the delivery of the questions in the online format was different from that used in the original NASA-TLX measurement; all five questions are word-for-word duplicates of the original questions. Each question prompted participants to rate an aspect of their perceived cognitive load on a 21-point scale. The questions addressed task demand, hard work, feeling of success, ease of navigation, and stress level. For example, the fifth question on the survey addressed the students' perceived stress level. It was displayed with the following text: "How much mental and physical activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching etc.)? That is, was the learning task easy (simple, forgiving) or demanding (exacting or unforgiving)?" Participants then selected one of 21 radio buttons that were arranged from left to right on the screen to indicate how easy or demanding the program was. As noted

earlier, a complete copy of this assessment, including all five cognitive load questions, is included in Appendix E.

Although all students were presented with the five cognitive load questions, due to a programming error that was not discovered until late in data collection, only 36 of the 95 student responses to this section of the survey were recorded properly and analyzed.

The attitude portion of the questionnaire followed the cognitive load questions on a separate screen. It included 14 items designed to elicit general reactions to the listening activity. Each was scored on a 5-point Likert-type response scale. Participants were asked for their opinions about the activity's organization, relevance, interest, ease of use, and its ease of navigation. For example, participants were asked to rate the statement, "The program was well designed and organized," by selecting one of five radio buttons labeled "strongly agree," "agree," "neither agree nor disagree," "disagree," or "strongly disagree."

Additionally, depending on the treatment group, participants were asked to indicate how hearing individual key words, reading translations and seeing illustrations helped them to understand the Sanfermines description and to learn new vocabulary. The text translation and illustration questions were only asked of participants in the corresponding treatment groups. For example, participants in all treatment groups were asked "Hearing the keywords pronounced alone helped me to LEARN the new words." However, only those in the textual-annotations-included treatment were presented with the statement, "Reading the English translations of keywords helped me to UNDERSTAND the story." All questions

in this section were rated on the "strongly agree" to "strongly disagree" 5-point scale.

Three open-ended questions followed, which asked participants how the vocabulary annotations could be made more effective, what they liked best about the activity and what could be done to improve it. For example, participants were asked, "How could we make vocabulary definitions, pronunciation aids and illustrations more effective for you?" and then were presented with an empty text box where they typed their answers. As noted earlier, a full list of questions is included in Appendix E.

Participants' time-in-program was logged by the instructional program as well. Time was recorded for the listening activity separately from time on any of the quizzes or surveys so that time-in-program could be analyzed for any effect it may have had on subsequent student performance.

Data Analysis

A posttest-only, experimental two (visual annotations excluded and included) by two (textual annotations excluded and included) factorial design was used in this study. Posttest and survey results data were extracted from the online assessment program and entered into SPSS for analysis. Separate 2x2 Analyses of Variance (ANOVA) were performed to evaluate the effects of textual definitions and visual illustrations on listening comprehension, vocabulary acquisition and cognitive load. Additional ANOVAs were calculated for time-in-program. A factor analysis was performed for the attitudes section of the survey. A MANOVA was then conducted on those factors.

CHAPTER III

RESULTS

Results for the listening comprehension posttest, vocabulary posttest, cognitive load survey and attitude survey are presented below in the same order as the three research questions. The first research question related to the effects of textual and visual annotation on listening comprehension and on vocabulary acquisition. Results from the comprehension and vocabulary measures will be presented separately.

Listening Comprehension

A one-way ANOVA was conducted initially to determine if the different locations where the study was offered and the different classes (SPA101 and SPA102) had an effect on listening comprehension scores. The ANOVA indicated no significant effect for either location or class.

The means and standard deviations for listening comprehension performance by textual definitions (excluded and included) and visual illustrations (excluded and included) are presented in Table 1. The overall mean score for all participants seems quite low at 10.32 (SD=5.16) out of a possible 32 propositions in the listening activity; however scores in this range are typical for a proposition-recall assessment of this type. The mean score for participants in the textual definitions excluded treatment was 9.16 (SD=4.56), while the overall mean for the textual definitions included treatment was 11.54 (SD=5.53). The overall mean for the visual illustrations excluded treatment was 9.76 (SD=4.60) and the overall mean score for the visual illustrations included treatment was

about one point higher, at 10.84 (SD=5.64). Participants in the no-annotations treatment (textual definitions and visual illustrations excluded) achieved a mean score of 9.27 (SD=4.86) while students in the visual illustrations only (textual definitions excluded) treatment scored a mean of 9.07 (SD=4.39). The mean for the textual definitions only treatment (visual illustrations excluded) was 10.21 (SD=4.39) which is a couple of points lower than the mean score of 13.00 (SD=6.33) for the combination treatment of textual definitions and visual illustrations.

Table 1

Means and standard deviations for the listening-comprehension measure

Textual Definitions Treatment		Visual Illustrations Treatment		
		Visual Illustrations Excluded	Visual Illustrations Included	Total
Textual Definitions Excluded	*M*	9.27	9.07	9.16
	SD	4.86	4.39	4.56
	n	22	27	49
Textual Definitions Included	*M*	10.21	13.00	11.54
	SD	4.39	6.33	5.53
	n	24	22	46
Total	*M*	9.76	10.84	10.32
	SD	4.60	5.64	5.16
	n	46	49	95

Note: The maximum score was 32 comprehension propositions recalled.

A 2 x 2 Analysis of Variance (ANOVA) was conducted to examine the effects of the visual-illustrations treatment and the textual-definitions treatment on listening comprehension posttest scores. Table 2 provides a summary of the ANOVA scores. The results indicated a main effect for the textual-illustrations treatment, $F(1, 91)=5.56$, $p<.05$. Effect sizes of .01 are considered small, .06 is considered medium, and .14 is considered high (Pierce, Block and Aguinis, 2004). The partial η^2 was .058, which is considered a medium effect size. There was no significant effect for the visual-definitions treatment and there were no significant interactions between treatments.

Table 2

ANOVA summary table for comprehension posttest achievement scores by textual definition and visual illustration conditions

Source	*df*	*F*	Partial η^2	*p*
Textual Definitions Treatment	1	5.56	.058	.021*
Visual Illustrations Treatment	1	1.58	.017	.212
Text x Visual	1	2.10	.023	.151
Error	91	(25.079)		

Note: $p<.05$, Values enclosed in parentheses represent mean square errors.

Vocabulary Acquisition

A one-way ANOVA was conducted initially to determine if the different locations where the study was offered and the different classes (SPA101 and SPA102) had an effect on vocabulary acquisition scores. The ANOVA indicated no significant effect for either location or class.

The mean scores and standard deviations for vocabulary acquisition posttest performance by textual definitions (excluded and included) and visual illustrations (excluded and included) are presented in Table 3. The overall mean score for all participants was 21.27 (SD=3.42) out of 25 possible points. The mean score for participants in the textual definitions excluded treatment was 21.10 (SD=3.45), while the overall mean for the textual definitions included treatment was 21.46 (SD=3.41). The overall mean for the visual illustrations excluded treatment was 20.65 (SD=3.28) and the overall mean score for the visual illustrations included treatment was 21.86 (SD=3.48). Participants in the no-annotations treatment (textual definitions and visual illustrations excluded) achieved a mean score of 20.50 (SD=2.70) while students in the visual illustrations only (textual definitions excluded) treatment scored a mean of 21.59 (SD=3.94). The mean for the textual definitions only treatment (visual illustrations excluded) was 20.79 (SD=3.79) which contrasts with a mean score of 22.18 (SD=2.86) for the combination treatment of textual definitions and visual illustrations.

A 2 x 2 Analysis of Variance (ANOVA) was conducted to test the effects of the visual-illustrations treatment and the textual-definitions treatment on vocabulary acquisition posttest scores. Table 4 provides a summary of the ANOVA scores. The ANOVA yielded no significant difference for any of the variables, nor were there any interaction effects.

Table 3

Means and standard deviations for the vocabulary acquisition posttest

Textual Definitions Treatment		Visual Illustrations Treatment: Visual Illustrations Excluded	Visual Illustrations Treatment: Visual Illustrations Included	Total
Textual Definitions Excluded	*M*	20.50	21.59	21.10
	SD	2.70	3.94	3.45
	n	22	27	49
Textual Definitions Included	*M*	20.79	22.18	21.46
	SD	3.79	2.86	3.41
	n	24	22	46
Total	*M*	20.65	21.86	21.27
	SD	3.28	3.48	3.42
	n	46	49	95

Note: The maximum score was 25 points.

Table 4

ANOVA summary table for vocabulary acquisition posttest achievement scores by textual definition and visual Illustration conditions

Source	*df*	*F*	Partial η^2	*p*
Textual Definitions Treatment	1	.39	.004	.53
Visual Illustrations Treatment	1	3.12	.033	.08
Text x Visual	1	.05	.000	.83
Error	91	(11.64)		

Note: Values enclosed in parentheses represent mean square errors.

Cognitive Load

The Cognitive Load measure was an adaptation of the NASA-TLX measure (Hart and Staveland, 1988) and consisted of five questions designed to address various aspects of cognitive load. Again, due to a programming error, only a subset (n=36) of the participants' responses to these five questions were recorded correctly. All questions were scored on a 21-point scale and each is presented separately below.

Task demand. The question, "How much mental and physical effort was required? Was the learning task easy or demanding?" was rated on a scale of one to twenty, from "easy" to "demanding." The mean scores and standard deviations for this question by textual definitions (excluded and included) and visual illustrations (excluded and included) are presented in Table 5. The overall mean rating for all participants was 10.33 (SD=4.90). The mean rating for participants in the textual-definitions-excluded treatment was 11.13 (SD=4.48), while the mean for the textual-definitions-included treatment was 8.92 (SD=5.47). The overall mean for the visual-illustrations-excluded treatment was 11.13 (SD=5.46) and the overall mean rating for the visual-illustrations-included treatment was 9.76 (SD=4.50). Participants in the no-annotations treatment (textual-definitions and visual-illustrations-excluded) rated the question with a mean of 12.50 (SD=5.40) while students in the visual-illustrations-only treatment (textual definitions excluded) responded with a mean rating of 10.08 (SD=3.48). The mean for the textual definitions only treatment (visual illustrations excluded) was

8.40 (SD=4.98) compared with a mean rating of 9.25 (SD=6.06) for the combination treatment of textual definitions and visual illustrations.

A 2 x 2 Analysis of Variance (ANOVA) was conducted to test the effects of the visual-illustrations treatment and the textual-definitions treatment on the student ratings on the "demanding" question. Table 6 provides a summary of the ANOVA results. The results indicate that there were no main effects for either variable and there were no interaction effects.

Table 5

Means and standard deviations for ratings of the "task demand" cognitive load question
(1=easy, 21=demanding)

Textual Definitions Treatment		Visual Illustrations Treatment		
		Visual Illustrations Excluded	Visual Illustrations Included	Total
Textual Definitions Excluded	*M*	12.50	10.08	11.13
	SD	5.40	3.48	4.48
	n	10	13	23
Textual Definitions Included	*M*	8.40	9.25	8.92
	SD	4.98	6.06	5.47
	n	5	8	13
Total	*M*	11.13	9.76	10.33
	SD	5.46	4.50	4.90
	n	15	21	36

Table 6

ANOVA summary table for the "task demand" cognitive load question

Source	*df*	*F*	Partial η^2	*p*
Textual Definitions Treatment	1	2.03	.06	.16
Visual Illustrations Treatment	1	.21	.01	.65
Text x Visual	1	.89	.03	.35
Error	35	(23.879)		

Note: Values enclosed in parentheses represent mean square errors.

Hard work. The mean scores and standard deviations for the question, "How hard did you have to work to understand the contents of the learning environment?" are presented in Table 7 by textual definitions (excluded and included) and visual illustrations (excluded and included). As the other cognitive load questions, this question was rated on a 21-point scale, from "not hard at all" to "very hard." The mean rating for all participants was 10.39 (SD=5.24). The overall mean rating for participants in the textual definitions excluded treatment was 10.74 (SD=5.22), and the overall mean for the textual definitions included treatment was 9.77 (SD=5.42). The overall mean for the visual illustrations excluded treatment was 9.87 (SD=6.19) while the overall mean score for the visual illustrations included treatment was 10.76 (SD=4.57). Participants in the no-annotations treatment (textual definitions and visual illustrations excluded) gave a mean rating of 9.90 (SD=6.49) while students in the visual illustrations only (textual definitions excluded) treatment responded with a mean of 11.38

(SD=4.17). The mean for the textual definitions only treatment (visual illustrations excluded) was 9.80 (SD=6.26) compared with a mean rating of 9.75 (SD=5.29) for the combination treatment of textual definitions and visual illustrations.

A 2 x 2 Analysis of Variance (ANOVA) was conducted to test the effects of the visual-illustrations treatment and the textual-definitions treatment on student ratings on the "hard work" question. Table 8 provides a summary of the ANOVA scores. The results indicate that there were no main effects for either variable and there were no interaction effects.

Table 7

Means and standard deviations for ratings of the "hard work" cognitive load question
(1=Not hard at all, 21=Very hard)

Textual Definitions Treatment		Visual Illustrations Treatment		
		Visual Illustrations Excluded	Visual Illustrations Included	Total
Textual Definitions Excluded	*M*	9.90	11.38	10.74
	SD	6.49	4.17	5.22
	n	10	13	23
Textual Definitions Included	*M*	9.80	9.75	9.77
	SD	6.26	5.29	5.42
	n	5	8	13
Total	*M*	9.87	10.76	10.39
	SD	6.19	4.57	5.24
	n	15	21	36

Table 8

ANOVA summary table for the "hard work" cognitive load question

Source	*df*	*F*	Partial η^2	*p*
Textual Definitions Treatment	1	.20	.01	.66
Visual Illustrations Treatment	1	.14	.00	.71
Text x Visual	1	.16	.01	.69
Error	35	(29.384)		

Note: Values enclosed in parentheses represent mean square errors.

Feeling of success. The mean scores and standard deviations for the question, "How successful do you think you were in your attempt to understand the contents of the learning environment?" are presented in Table 9 by textual definitions (excluded and included) and visual illustrations (excluded and included). This question was also rated on a 21-point scale, from "not successful" to "very successful." The mean rating for all participants was 13.36 (SD=4.14). The overall mean rating for participants in the textual definitions excluded treatment was 12.13 (SD=3.67), while the overall mean for the textual definitions included treatment was 15.54 (SD=4.16). The overall mean for the visual illustrations excluded treatment was 12.80 (SD=4.26) and the overall mean rating for the visual illustrations included treatment was 13.76 (SD=4.11). Participants in the no-annotations treatment (textual definitions and visual illustrations excluded) responded with a mean rating of 11.70 (SD=3.06) while students in the visual illustrations only (textual definitions excluded) treatment recorded a mean of 12.46 (SD=4.18). The mean for the textual definitions only treatment (visual

illustrations excluded) was 15.00 (SD=5.79) which contrasts with a mean rating of 15.88 (SD=3.18) for the combination treatment of textual definitions and visual illustrations.

Table 9

Means and standard deviations for ratings of the "feeling of success" cognitive load question (1=not successful, 21=very successful)

Textual Definitions Treatment		Visual Illustrations Treatment		
		Visual Illustrations Excluded	Visual Illustrations Included	Total
Textual Definitions Excluded	*M*	11.70	12.46	12.13
	SD	3.06	4.18	3.67
	n	10	13	23
Textual Definitions Included	*M*	15.00	15.88	15.54
	SD	5.79	3.18	4.16
	n	5	8	13
Total	*M*	12.80	13.76	13.36
	SD	4.26	4.11	4.14
	n	15	21	36

A 2 x 2 Analysis of Variance (ANOVA) was conducted to test the effects of the visual-annotations treatment and the textual-annotations treatment on student responses to the “feel successful” question. Table 10 provides a summary of the ANOVA scores. The results indicated a main effect for the textual definitions treatment, $F(1, 35)=5.77$, $p<.05$. The partial η^2 of .15 is considered a high effect size. Ratings were significantly higher for participants who received

textual translations as compared to those who did not. There was no main effect for the visual-illustrations treatment and there were no interaction effects.

Table 10

ANOVA summary table for the "feeling of success" cognitive load question

Source	*df*	*F*	Partial η^2	*p*
Textual Definitions Treatment	1	5.77	.15	.02*
Visual Illustrations Treatment	1	.34	.01	.56
Text x Visual	1	.00	.00	.97
Error	35	(15.569)		

Note: *p>.05, Values enclosed in parentheses represent mean square errors.

Navigation effort. The mean scores and standard deviations for the question, "How much effort did you have to invest to navigate the learning environment?" are presented in Table 11 by textual definitions (excluded and included) and visual illustrations (excluded and included). This question was scored on a 21-point scale, from "low effort" to "high effort." Therefore, the lower the score, the easier participants felt it was to navigate the program and the lower their level of reported extrinsic cognitive load. The mean rating for all participants on this question was 5.28 (SD=5.59) on the 21-point scale, indicating that all participants found the program relatively easy to navigate. The overall mean rating for participants in the textual definitions excluded treatment was 5.13 (SD=5.91), while the overall mean for the textual definitions included treatment was 5.54 (SD=5.21). The overall mean for the visual illustrations excluded treatment was 3.73 (SD=4.50) and the overall mean rating for the visual

illustrations included treatment was 6.38 (SD=6.13). Participants in the no-annotations treatment (textual definitions and visual illustrations excluded) responded with a mean rating of 2.00 (SD=1.05) while students in the visual illustrations only (textual definitions excluded) treatment rated this question with a mean of 7.54 (SD=6.98). The mean for the textual definitions only treatment (visual illustrations excluded) was 7.20 (SD=6.76) which contrasts with a mean of 4.50 (SD=4.14) for the combination treatment of textual definitions and visual illustrations.

Table 11

Means and standard deviations for ratings of the "navigation effort" cognitive load question (1=low effort, 21=high effort)

		Visual Illustrations Treatment		
Textual Definitions Treatment		Visual Illustrations Excluded	Visual Illustrations Included	Total
Textual Definitions Excluded	*M*	2.00	7.54	5.13
	SD	1.05	6.98	5.91
	n	10	13	23
Textual Definitions Included	*M*	7.20	4.50	5.54
	SD	6.76	4.14	5.21
	n	5	8	13
Total	*M*	3.73	6.38	5.28
	SD	4.50	6.13	5.59
	n	15	21	36

A 2 x 2 Analysis of Variance (ANOVA) was conducted to test the effects of the visual-annotations treatment and the textual-annotations treatment on student responses to the "navigation effort" question. Table 12 provides a summary of the ANOVA scores. The results indicated no main effect for either treatment. However, there was an interaction effect, $F(1, 35)=4.82$, $p<.05$.

Table 12

ANOVA summary table for the "navigation effort" cognitive load question

Source	*df*	*F*	Partial η^2	*p*
Textual Definitions Treatment	1	.33	.01	.57
Visual Illustrations Treatment	1	.57	.02	.46
Text x Visual	1	4.82	.12	.04
Error	35	(28.063)		

Note: Values enclosed in parentheses represent mean square errors.

Stress levels. The mean scores and standard deviations for the question, "How stressed did you feel during the learning task?" are presented in Table 13 by textual definitions (excluded and included) and visual illustrations (excluded and included). Participants rated the question on a 21-point scale, from "not at all" to "extremely." The mean rating for all participants was 7.86 (SD=5.48), indicating a relatively low-stress level. The overall mean rating for participants in the textual definitions excluded treatment was 9.04 (SD=5.25), while the overall mean for the textual definitions included treatment was about four points lower, at 5.77 (SD=5.45). The overall mean for the visual illustrations excluded treatment

was 6.33 (SD=4.79) and the overall mean rating for the visual illustrations included treatment was 8.95 (SD=5.79). Participants in the no-annotations treatment (textual definitions and visual illustrations excluded) responded with a mean rating of 7.00 (SD=5.12) while students in the visual illustrations only (textual definitions excluded) treatment rated this question with a mean of 10.62 (SD=4.98). The mean for the textual definitions only treatment (visual illustrations excluded) was 5.00 (SD=4.24) which contrasts with a mean of 6.25 (SD=6.32) for the combination treatment of textual definitions and visual illustrations.

Table 13

Means and standard deviations for the "stress levels" cognitive load question (1=not at all, 21=extremely)

Textual Definitions Treatment		Visual Illustrations Treatment		
		Visual Illustrations Excluded	Visual Illustrations Included	Total
Textual Definitions Excluded	*M*	7.00	10.62	9.04
	SD	5.12	4.98	5.25
	n	10	13	23
Textual Definitions Included	*M*	5.00	6.25	5.77
	SD	4.24	6.32	5.45
	n	5	8	13
Total	*M*	6.33	8.95	7.86
	SD	4.79	5.79	5.48
	n	15	21	36

A 2 x 2 Analysis of Variance (ANOVA) was conducted to test the effects of the visual-illustrations treatment and the textual-definitions treatment on student responses to the "stress" question. Table 14 provides a summary of the ANOVA scores. The results indicated no main effects for either treatment and no interaction effects.

Table 14

ANOVA summary table for the "stress levels" cognitive load question

Source	*df*	*F*	Partial η^2	*p*
Textual Definitions Treatment	1	2.92	.08	.10
Visual Illustrations Treatment	1	1.71	.05	.20
Text x Visual	1	.40	.01	.53
Error	35	(27.643)		

Note: Values enclosed in parentheses represent mean square errors.

Participant attitude survey scores by item

A set of Likert-type questions was used to measure student perceptions of how well the program was designed. Table 15 contains a full list of questions along with mean scores and standard deviations for participant responses. Questions are presented in the order in which they appeared to study participants. For ease of data presentation, the questions are numbered from 1 to 14 although in the survey the items were numbered differently in the actual survey. The response scale was from one to five, from strongly disagree to strongly agree, thus a higher number indicates stronger agreement with the given statement. The range of

mean scores was from 3.48 to 4.65, and the overall mean score for all questions was 4.20 (SD=.93).

Table 15

Overall mean scores and standard deviations for attitude survey questions

Item	Mean	SD
1. The program was well designed and organized.	4.34	0.71
2. The topic of the program was relevant to my Spanish study.	4.17	0.83
3. The story was interesting to me.	4.21	0.90
4. Instructions within the program were clear and easy to follow.	4.52	0.76
5. Navigation within the program was easy to understand. (n=94)	4.60	0.68
6. The listening activity helped me to learn new vocabulary.	4.05	1.11
7. The listening activity helped me to learn about the cultural topic.	4.29	0.84
8. I would like to have more listening activities of this type to help me understand spoken Spanish.	4.12	0.98
9. Hearing the keywords pronounced alone helped me to LEARN the new words. (all groups, n=95)	3.67	1.03
10. Hearing the keywords pronounced alone helped me to UNDERSTAND the story. (all groups, n=95)	3.48	1.14
11. Reading the English translations of keywords helped me to LEARN the new words. (textual definitions only and combination groups, n=46)	4.65	0.53
12. Reading the English translations of keywords helped me to UNDERSTAND the story. (textual definitions only and combination groups, n=46)	4.65	0.53

13. Seeing the graphics illustrating keywords helped me to LEARN the new words. (visuals only and combination groups, n=49)	4.33	0.77
14. Seeing the graphics illustrating keywords helped me to UNDERSTAND the story. (visuals only and combination groups, n=49)	4.41	0.73

Note: *All items were rated on a five-point scale from strongly disagree to strongly agree. N=95 except as noted.*

The first eight questions were designed to elicit participant attitudes towards the overall design of the instructional program. In response to the design and organization question participants gave a mean rating of 4.34 (SD=.71). Participants also rated the relevance of the program positively, resulting in a mean score of 4.17 (SD=.83). The interest of the story was rated favorably with a mean score of 4.21 (SD=.90). Participants rated the clarity of the instructions with a mean of 4.52 (SD=.76) and ease of navigation with a mean of 4.60 (SD=.68). The statement that the program helped them learn new vocabulary was rated slightly lower, with a mean of 4.05 (SD=1.11). Participants indicated that the program helped them learn about the cultural topic with a mean score of 4.29 (SD=.84). The statement that they would like to have more activities of this type in their regular Spanish class scored a mean of 4.12 (SD=.98).

Questions 9 through 14 in Table 15 were related to the effectiveness of the vocabulary annotations participants accessed within the program. Questions 9 and 10 were given to participants in all treatments, while questions 11 and 12 were only given to participants in the textual-definitions-included treatment and questions 13 and 14 were offered only to participants in the visual-illustrations-included treatment.

Questions 9 and 10 were the two lowest-rated items in the survey. Question 9, "Hearing the keywords pronounced alone helped me to learn the new words" had a mean score of 3.67 (SD=1.03). Participants rated question 10, "Hearing the keywords pronounced alone helped me to understand the story" with a mean of 3.48 (SD=1.14).

The textual-definitions questions, numbers 11 and 12, were rated more highly, however. Question 11, "Reading the English translations of keywords helped me to learn the new words" had a mean rating of 4.65 (SD=.53) on the five-point scale. Question 12, "Reading the English translations of keywords helped me to understand the story" had an identical mean score of 4.65 (SD=.53).

Likewise, visual-illustration questions, numbers 13 and 14, were rated highly. "Seeing the graphics illustrating keywords helped me to learn the new words," question 13 had a mean score of 4.33 (SD=.77). Question 14, "Seeing the graphics illustrating keywords helped me to understand the story" received the second-highest score of the survey with a mean of 4.41 (SD=.73).

A factor analysis was performed to determine appropriate variables for further statistical analysis. After finding eigenvalues and examining a scree plot, question one did not load onto any specific factor and was removed from the analysis. For the first factor, the eigenvalue was 3.93, the second was 3.41, the third was 3.56, the fourth was 1.20, and the fifth was 1.16. These five factors accounted for 87% of the total item variance. A varimax rotation was used to better interpret the factor pattern. The rotated factor loadings and communities for the five extracted factors are reported in Table 16.

After the first question was eliminated, questions 2, 3, 4 and 5 had the highest loading in factor 1. These questions comprised two themes; one about relevance and interest, the other about clarity and ease of navigation. Factor 1 was labeled "relevance and clarity" for further analysis. The next three questions in the survey, items 6, 7, and 8, all loaded onto the second factor. These three questions related to perceived learning of vocabulary, comprehension and overall learning. Thus, factor 2 was labeled "learning." The last two items, numbers 13 and 14, loaded onto factor number 3. These two questions were related to the visual illustrations and their effects on learning vocabulary and on listening comprehension. Factor 3 was labeled "attitude towards visual annotations." Item numbers 9 and 10 were both about the effects of hearing key words pronounced individually and they were the only two questions that loaded onto factor 4. Factor 4 was labeled "attitude towards hearing keywords." Finally, factor 5 was comprised of two items, numbers 11 and 12. These two items related to the perceived effects of textual annotations on vocabulary learning and on listening comprehension. Thus, factor 5 was labeled "attitude towards textual annotations."

Table 16

Summary of factor loadings of five-factor solution for student attitudes

Item	Factors 1	2	3	4	5
1. The program was well designed and organized. (excluded from analysis)					
2. The topic of the program was relevant to my Spanish study.	**.596**	.517	-.019	-.148	-.077
3. The story was interesting to me.	**.700**	.268	.470	-.101	-.113
4. Instructions within the program were clear and easy to follow.	**.981**	.051	.062	-.061	-.165
5. Navigation within the program was easy to understand. (n=94)	**.901**	-.011	-.016	.019	-.009
6. The listening activity helped me to learn new vocabulary.	-.028	**.915**	-.175	.290	.054
7. The listening activity helped me to learn about the cultural topic.	.139	**.838**	-.122	.197	-.130
8. I would like to have more listening activities of this type to help me understand spoken Spanish.	.134	**.793**	.142	.205	-.053
9. Hearing the keywords pronounced alone helped me to LEARN the new words. (all groups, n=95)	.052	.319	-.284	**.802**	.173
10. Hearing the keywords pronounced alone helped me to UNDERSTAND the story. (all groups, n=95)	-.205	.306	.025	**.929**	.008

11. Reading the English translations of keywords helped me to LEARN the new words. (textual definitions only and combination groups, n=46)	-.078	-.127	.080	.166	**.971**
12. Reading the English translations of keywords helped me to UNDERSTAND the story. (textual definitions only and combination groups, n=46)	-.163	.007	.354	-.030	**.679**
13. Seeing the graphics illustrating keywords helped me to LEARN the new words. (visuals only and combination groups, n=49)	.063	-.124	**.977**	-.019	.158
14. Seeing the graphics illustrating keywords helped me to UNDERSTAND the story. (visuals only and combination groups, n=49)	.111	-.029	**.904**	-.171	.247

Note: *Bold type indicates highest factor loadings.*

From these five identified factors, five new variables were created for further analysis. Means were calculated for the items that made up each factor and subsequent analyses were performed. The factors "relevance and clarity," "learning," and "hearing keywords" had participants from all treatments, so a 2x2 multivariate analysis of variance (MANOVA) was calculated to check for significant differences among treatments. Factors "attitude towards visual annotations" and "attitude towards textual annotations" contained a variable with only one level. The "attitude towards visual annotations" questions were only presented to those in the visual-annotations-included groups. The "attitude

towards textual annotations" questions were only offered to those in the textual-annotations-included treatments. Therefore one-way ANOVAs were performed on these factors.

Results for the relevance and clarity, learning, and hearing keywords attitude factors. Means and standard deviations for these three factors are found in Table 17. The highest mean ratings were given for the factor "relevance and clarity" (M=4.38, SD=.62) slightly lower ratings were given to the "learning" factor (M=4.15, SD=.85) and the lowest ratings were given for the factor "hearing keywords" (M=3.56, SD=.99)

Students in all groups rated the "relevance and clarity" of the program positively. The textual-annotations-excluded participants rated it with a mean score of 4.33 (SD=.61), the textual-annotations-included group rated it 4.43 (SD=.62), the visual-annotations-excluded group rated it 4.36 (SD=.59), and the visual-annotations-included participants rated it with a mean of 4.40 (SD=.65).

Table 17

Means and standard deviations for student attitude factors "relevance and clarity," "learning," and "hearing keywords" (1=strongly disagree, 5=strongly agree)

	Factors for Student Attitudes Survey					
	Relevance and Clarity		Learning		Hearing Keywords	
Group	M	SD	M	SD	M	SD
Textual annotations						
Excluded	4.33	.61	4.07	.89	3.42	1.08
Included	4.43	.62	4.23	.81	3.72	.87
Visual annotations						
Excluded	4.36	.59	4.19	.75	3.49	1.02
Included	4.40	.65	4.11	.95	3.64	.95
Totals	4.38	.62	4.15	.85	3.56	.99

Likewise, the "learning" factor received positive ratings from all groups, although means were slightly lower than for those on the "relevance and clarity" factor. The textual-annotations-excluded participants rated it with a mean score of 4.07 (SD=.89), the textual-annotations-included group rated it 4.23 (SD=.81), the visual-annotations-excluded group rated it 4.19 (SD=.75), and the visual-annotations-included participants rated it with a mean of 4.11 (SD=.95).

Participants rated the "hearing keywords" factor neutrally. The textual-annotations-excluded participants rated it with a mean score of 3.42 (SD=1.08), the textual-annotations-included group rated it 3.72 (SD=.87), the visual-

annotations-excluded group rated it 3.49 (SD=1.02), and the visual-annotations-included participants rated it with a mean of 3.64 (SD=.95).

A multivariate analysis of variance (MANOVA) was conducted to determine the effect of textual and visual annotations on students perceptions of the three factors, "relevance and clarity," "learning," and "hearing keywords." The results of the analysis revealed no significant main effects or interactions among the variables.

Results of the "attitude towards textual annotations" factor. The mean scores and standard deviations for the "attitude towards textual annotations" factor are presented in Table 18 by visual illustrations (included and excluded). Participants rated the questions that make up this factor on a five-point scale, from “strongly disagree” to “strongly agree.” Again, only participants in the textual-annotations-included group were presented with these questions. The mean rating of the effectiveness of textual annotations for all participants who accessed them was 4.65 (SD=.49). Participants who accessed who accessed both textual and visual annotations rated this factor higher (M=4.82, SD=.36) than those who accessed textual annotations, but did not see visual annotations (M=4.50, SD=.55)

A one-way analysis of variance (ANOVA) was conducted to test the effects of the visual-illustrations treatment on student responses to the survey questions that make up the “attitude toward textual annotations” factor. The ANOVA was significant, $F(2, 44) = 5.34$, $p=.03$. This indicates that participants in the both-annotations group felt significantly more positive towards the presence

of textual annotations than did those participants in the textual-annotations-only group.

Table 18

Means and standard deviations for factor "attitude towards textual annotations" (1=strongly disagree, 5=strongly agree)

		Visual Illustrations Treatment		Totals
		Visual Illustrations Included	Visual Illustrations Excluded	
Textual Definitions Included	*M*	4.82	4.50	4.65
	SD	.36	.55	.49
	N	24	22	46

Table 19

One-way ANOVA summary table for the effect of visual illustrations on "attitude towards textual annotations"

Source	*SS*	*MS*	*F(2, 44)*	*p*	Partial η^2
Between groups	1.16	1.16	5.34	.03*	.11
Within groups	9.77	.22			

Note: *p>.05

Results of the "attitude toward visual annotations" factor. The mean scores and standard deviations for the factor "attitude toward visual annotations" are presented in Table 20 by textual definitions (included and excluded). Participants who accessed visual annotations rated the questions that make up this factor on a five-point scale, from “strongly disagree” to “strongly agree.” The mean rating for all participants was 4.37 (SD=.72). Participants in the textual-annotations-included group rated the factor with a mean of 4.57 (SD=.58) and

those in the textual-annotations-excluded group gave this factor a lower rating of 4.20 (SD=.79).

Table 20

Means and standard deviations for factor "attitude toward visual annotations" (1=strongly disagree, 5=strongly agree)

		Textual Definitions Treatment		Totals
		Textual Definitions Included	Textual Definitions Excluded	
Visual Illustrations Included	*M*	4.57	4.20	4.37
	SD	.58	.79	.72
	n	22	27	49

A one-way analysis of variance (ANOVA) was conducted to test the effects of the textual-annotations treatment on student responses to the survey questions that make up the “visual annotations” factor. Table 21 provides a summary of the ANOVA scores. The results indicated no significant effects.

Table 21

One-way ANOVA Summary Table for effect of visual illustrations on perceptions of "attitude toward visual annotations"

Source	*SS*	*MS*	*F(2, 44)*	*p*	Partial η^2
Between groups	1.16	1.16	3.25	.08	.07
Within groups	23.28	.50			

Open-ended survey question responses

Participants responded to three open-ended questions at the end of the attitude survey. Results for each question are presented in turn below, followed by participant responses from each treatment group.

Vocabulary aid improvements. The first open-ended question was "How could we make vocabulary definitions, pronunciation aids and illustrations more effective for you?" Table 22 contains a summary of responses to this question. We will describe the overall results first and then summarize responses by treatment group.

The 95 participants in the study wrote a total of 115 different responses to this question. The two most common responses for all groups, with twenty-eight responses each, was that the vocabulary aids were acceptable in their current form and that they would like more or better textual definitions or translations. The second most common response, with nineteen respondents, was that they would like to see more or better visual illustrations. Ten responses indicated a desire for video illustrations and ten others indicated that the quality or the pace of the narrative could be improved. Other respondents indicated a desire for the keywords to be highlighted on the screen as they were spoken in the narrative (eight responses), a desire for improved functionality in the way annotations were displayed (seven responses), a desire for a full transcript of the audio to read (three responses). Two additional responses indicated a need for vocabulary practice activities.

Twenty-six responses were entered by the no-annotations group. Thirteen of those twenty-six responses indicated that they would like more or better textual definitions or translations. Four wanted visual illustrations, four wanted the quality or pace of the narration to be better. Two wanted video illustrations and two more would like the keywords highlighted as the narrative plays. One response indicated that the vocabulary aids were acceptable as they were.

Twenty-six comments were entered by participants in the textual–definitions-only group. Eight of those responses indicated that the vocabulary aids were acceptable as they were. Six wanted more or better visual illustrations and four wanted better textual definitions. Three of these participants indicated that they would like the annotations to be displayed differently. Two wanted video illustrations, two more wanted a full transcript of the audio and one wanted keywords highlighted as the audio played.

Students in the visual-illustrations-only group were the most verbose in their comments, with a total of thirty-five responses. Ten of those indicated a desire for more or better textual definitions. Seven thought the visual annotations they saw were acceptable and seven more thought there should be more or better visual annotations. Five comments included a desire for video illustrations. Two wanted the quality or pace of the narrative of the program to be better and two more commented that they would like the functionality to be improved. One participant mentioned highlighting keywords as they were spoken and one more wanted vocabulary practice activities.

Table 22

Vocabulary aid improvements: Summary of participant responses

Item	Responses by annotation treatment				
	None	Text	Visual	Both	Total
How could we make the vocabulary pronunciation aids more effective for you?					
Vocabulary aids were acceptable	1	8	7	12	28
(More/better) textual definitions or translations	13	4	10	1	28
(More/better) visual illustrations	4	6	7	2	19
Video illustrations	2	2	5	1	10
Quality or pace of the narrative	4	0	2	4	10
Highlight keywords as they occur in the narration	2	1	1	4	8
Functionality of the program or how annotations are displayed	0	3	2	2	7
Full transcript or translations of audio	0	2	0	1	3
Vocabulary practice activities	0	0	1	1	2
Totals Number of responses	26	26	35	28	115

Note: *Annotation treatment group names in the responses columns refer to no-annotations, textual-definitions-only, visual-illustrations-only, and both-annotation types, respectively. Many individuals made more than one comment for each question.*

The both-annotations group made a total of twenty-eight comments in response to this question. Twelve of the comments, the most of any group, indicated that students liked the vocabulary aids as they were. Four wanted the quality or pace of the narrative to improve, and four more wanted the keywords highlighted as they were heard. Two commented that they would like better visual illustrations and two others indicated that they would like the way annotations were displayed to be improved. One expressed a desire for better textual definitions, one wanted video illustrations, another wanted a full transcript and one more wanted vocabulary practice activities.

Program likes. The second open-ended question was, "What did you like best about the program?" A summary of participant responses is found in Table 23. Again, we will describe the overall results and then discuss responses from each treatment group in turn.

There were 141 responses to this question from the 95 participants in the study, indicating that many participants made more than one comment about what they liked in the program. Thirty-six comments indicated that the students liked the cultural topic or story. Thirty liked the visual illustrations, twenty-three liked the program design and ease-of-use. Twenty-two more liked the keywords and ten liked the quality of the narration. Eight mentioned the pronunciation aids, six mentioned the textual definitions and four more liked the pace of the audio. There were two other comments that do not fit into the above categories. One student liked the fact that it could be completed at home and another liked the fact that it was all in Spanish.

Table 23

Program likes: Summary of participant responses

Item	Responses				
	None	Text	Visual	Both	Total
What did you like best about the program?					
Cultural topic or story	10	9	11	6	36
Visual illustrations	8	1	11	10	30
Program design, ease of use	7	5	5	6	23
Presence/effectiveness of keywords	3	8	3	8	22
Quality of narration	2	3	1	4	10
Pronunciation aids	0	2	3	3	8
Textual definitions	0	1	0	5	6
Pace of the audio	1	1	0	2	4
Other	0	1	1	0	2
Total responses	31	31	35	44	141

Note: *Annotation treatment group names in the responses columns refer to no-annotations, textual-definitions-only, visual-illustrations-only, and both-annotation types, respectively.*

Thirty-one of the comments came from the no-annotations group. Ten of them liked the cultural topic. Eight liked the visual illustrations they saw, which were general illustrations about the event, not specific to vocabulary. Seven liked the program design and three liked the keywords. Two liked the quality of the narration and one mentioned that the pace of the audio was good.

The textual-definitions-only group provided thirty-one responses as well. Nine of them mentioned the cultural topic and eight mentioned the effectiveness of keywords. Five liked the program design, three liked the quality of the narration and two liked the pronunciation aids. One comment indicated the visual illustrations were helpful, one mentioned the textual definitions, another mentioned the pace of the audio and one more mentioned that it was all in Spanish.

There were thirty-five responses from the visual-illustrations-only group. Eleven liked the cultural topic and eleven more liked the visual illustrations. Five liked the design of the program, three liked the keywords and three more liked the pronunciation aids. One mentioned the quality of narration and the one liked the fact that it could be completed at home.

The most comments, forty-four, came from the both-annotations group. Ten mentioned the visual illustrations, eight liked the keywords, six liked the cultural topic, and six more mentioned the program design as a positive. Five liked the textual definitions, four liked the quality of the narration. Three mentioned the pronunciation aids and two liked the pace of the audio.

Program improvements. The final open-ended question was, "What could be done to improve the program?" A summary of participant responses is found in Table 24. Once again, we will describe the overall results and then discuss responses from each treatment group in turn.

There was a total of 116 responses to this question. Twenty-three of the comments mentioned textual definitions and twenty-one mentioned wanting

slower audio. Sixteen mentioned a desire for more control over how annotations appeared. Twelve wanted improved visual illustrations and twelve more wanted video illustrations. Eight mentioned vocabulary practice activities, and seven wanted emphasized or highlighted keywords with the narration. Seven more mentioned that the program was good as it is. Five wanted more information about the topic and five wanted the opposite, less information or smaller chunks of information.

There were twenty-six responses from the no-annotations group. Ten of those responses mentioned the need for more textual definitions. Four wanted visual illustrations and three wanted slower audio. Two mentioned design improvements, two wanted vocabulary practice. Two others thought the program was fine as it was. One wanted keywords highlighted with the narration and one more wanted less information or smaller chunks of information.

The textual-definitions-only group provided twenty-eight comments. Several items had four responses: better textual definitions, slower audio, design, visual illustrations, video illustrations, and a desire for more information on the topic. Two comments mentioned vocabulary practice, one indicated satisfaction with the program as it was and another indicated a desire for less information.

Table 24

Program improvements: Summary of participant responses

Item	Responses				
	None	Text	Visual	Combo	Total
What could be done to improve the program?					
(More/better) textual definitions or translations	10	4	7	2	23
Slower audio	3	4	9	5	21
Design, control of annotations	2	4	5	5	16
(More/better) visual illustrations	4	4	3	1	12
Video illustrations	1	4	5	2	12
Vocabulary practice activities	2	2	2	2	8
Nothing	2	1	2	2	7
Emphasize/highlight keywords along with narration	1	0	2	4	7
Additional information about topic and keywords	0	4	1	0	5
Less information or smaller chunks of information	1	1	1	2	5
Total responses	26	28	37	25	116

Note: *Annotation treatment group names in the responses columns refer to no-annotations, textual-definitions-only, visual-illustrations-only, and both-annotation types, respectively.*

There were twenty-five comments from the both-annotations group. Five wanted slower audio and five more wanted more control over annotations. Four wanted keywords highlighted. There were two comments each in the categories better textual definitions, video illustrations, vocabulary practice, nothing, and less information. One more comment mentioned wanting better visual illustrations.

Participant time in program

Additional data were collected to investigate student time in-program. The time participants spent in the program was captured in a log by the assessment program. Logs recorded time participants spent in the listening activity as a whole, but did not record time spent on each screen. Time was not analyzed for the quiz or survey.

The mean scores and standard deviations for time in program are presented in Table 25 by textual definitions (excluded and included) and visual illustrations (excluded and included). The mean time in program for all participants was 11:03 (SD=5:06). The overall mean time for participants in the textual definitions excluded treatment was 11:38 (SD=5:21), while the overall mean time for the textual definitions included treatment was 10:26 (SD=4:49). The overall mean for the visual illustrations excluded treatment was 10:22 (SD=5:01) and the overall mean time for the visual illustrations included treatment was 11:41 (SD=5:10). Participants in the no-annotations treatment (textual definitions and visual illustrations excluded) spent a mean time of 11:19 (SD=4:50) while students in the visual illustrations only (textual definitions

excluded) treatment spent a mean time of 11:53 (SD=5:49). The mean time for the textual definitions only treatment (visual illustrations excluded) was 9:30 (SD=5:07) which contrasts with a mean of 11:27 (SD=4:21) for the combination treatment of textual definitions and visual illustrations.

Table 25

Means and standard deviations for time in program

Textual Definitions Treatment		Visual Illustrations Treatment: Visual Illustrations Excluded	Visual Illustrations Treatment: Visual Illustrations Included	Total
Textual Definitions Excluded	*M*	11:19	11:53	11:38
	SD	4:50	5:49	5:21
	n	22	27	49
Textual Definitions Included	*M*	9:30	11:27	10:26
	SD	5:07	4:21	4:49
	n	24	22	48
Total	*M*	10:22	11:41	11:03
	SD	5:01	5:10	5:06
	n	46	49	95

A 2 x 2 Analysis of Variance (ANOVA) was conducted to test the effects of the visual-illustrations treatment and the textual-definitions treatment on student time in program. Table 26 provides a summary of the ANOVA scores. The results indicated no main effects for either treatment and no interaction effects.

Table 26

ANOVA Summary table for time in program

Source	*df*	*F*	Partial η^2	*p*
Textual Definitions Treatment	1	1.13	.01	.29
Visual Illustrations Treatment	1	1.43	.02	.23
Text x Visual	1	.43	.01	.51
Error	94	(94028.75)		

Note: Values enclosed in parentheses represent mean square errors.

CHAPTER IV

DISCUSSION

This research study was designed to investigate the effects of visual and textual annotations in a multimedia listening activity on listening comprehension, vocabulary acquisition, and cognitive load. It was designed similarly to the Cottam and Savenye (2008) and included more participants in order to clarify potential effects that were observed in that study. Furthermore, because there has been more multimedia research on reading comprehension in languages than there has been on listening comprehension (Plass and Jones, 2005), the current study was intended to extend the understanding of how multimedia annotations affect comprehension and vocabulary acquisition in this particular type of activity. Also due to the difference in research volume between reading and listening comprehension, in this discussion we will reference reading comprehension as frequently as listening comprehension research in order to explore the research findings of the current study.

Participants in first-year Spanish classes at a community college and a university in the southwestern United States completed one of four randomly assigned Spanish listening activities with different types of visual and textual vocabulary annotations, which served as extralinguistic cues to make the Spanish that students heard more comprehensible. Participants then completed a listening comprehension posttest, a vocabulary posttest, and a survey that included cognitive load measures and attitude assessments.

In this chapter we will discuss the results related to each of the three research questions in turn, starting with the effects of textual and visual annotations on listening comprehension and vocabulary acquisition, followed by a discussion of the effects of textual and visual annotations on cognitive load. We will also include a discussion of student attitudes and time-in-program and will conclude with implications for design and suggestions for future research.

Listening comprehension

The first half of research question one, "What are the effects of textual and visual annotations on listening comprehension and vocabulary acquisition?" relates to how well students understood the story in Spanish under the different annotation conditions. Listening comprehension was measured by asking students to provide a summary in English of everything they heard in the activity. As in previous reading-comprehension studies by Chun and Plass (1996) and Plass, Chun, Mayer and Leutner (2003), and listening-comprehension studies by Jones and Plass (2002), Jones (2004), and Cottam and Savenye (2008), a recall protocol in English was used so that students' limited ability to express themselves in the Spanish language would not interfere with the assessment of their comprehension. This type of assessment consistently results in relatively low scores for participants. Therefore, although the mean score of 10.23 out of the 32 propositions in the listening comprehension posttest in the current study appears low, it is typical for reading and listening comprehension assessments of this type.

In the current study, data indicated that textual annotations had a significant positive effect on listening comprehension scores. Students who

accessed textual annotations of the key words identified significantly more of the narrative's propositions than did students who did not see the textual annotations. This result supports the findings of researchers such as Chun and Plass (1996a, 1996b), and Plass, Chun, Mayer, and Leutner (1998) who found a positive effect of textual annotations on reading comprehension tasks. Jones and Plass (2002), and Plass and Jones (2005) conducted studies similar to the current research and also found a positive effect for textual annotations on listening comprehension. Interestingly, Cottam and Savenye (2008) used the same materials and measures as the current study, and found no significant effect for textual annotations on listening comprehension.

Although the current study results indicated a positive effect for textual annotations on listening comprehension, they did not indicate a statistically significant effect for visual annotations on listening comprehension. Comprehension scores were higher for participants who accessed visual annotations than for those who did not see them, but the difference was not statistically significant. These results not only differ from the findings of Cottam and Savenye (2008), but they also differ from findings in other studies that used similar materials and listening comprehension measures. For example, Jones and Plass (2002), and Jones (2004) found a significant positive effect for visual annotations on listening comprehension. Their studies employed a listening activity with keywords and visual annotations presented in a design visually similar to the activity that was used in the current study. They also used a

proposition-recall measure of listening comprehension similar to the current study.

Although the above mentioned studies indicated a positive effect for visual annotations on listening comprehension, Ercetin (2004) and Ariew and Ercetin (2004) found a significant negative effect of visual annotations on reading comprehension in their studies. Their studies used a reading comprehension activity rather than a listening activity and they measured reading comprehension with a different type of measure. Their reading comprehension measure was comprised of a combination of short-answer, multiple-choice, and open-ended questions. Although students in their studies indicated a preference for visual annotations of keywords, those annotations had a negative effect on their comprehension scores (Ercetin, 2004; Ariew and Ercetin, 2004).

There may be a few reasons that the listening comprehension findings in the current study differ from previous studies. In addition to the treatment and measurement differences described here, the environment in which the research was conducted was vastly different. With the exception of the Cottam and Savenye (2008) study, previous listening comprehension and reading comprehension studies were conducted with intact classes in structured language classes or language lab environments, whereas the current study was offered to students in various online and in-person classes as an optional activity. It was completed on the students' own time, outside of class, in whatever environment they chose.

Thus, the instructional environment of the Cottam and Savenye (2008) study and the current study more closely approach the way that students might participate in a listening activity independently in an online language class. Perhaps in this environment, the presence of visual annotations has a less significant effect on listening comprehension than it does in a laboratory or classroom environment. Nonetheless, more research is required to investigate if there is a difference between student behaviors in a class or lab and a completely online environment.

Vocabulary acquisition

The second half of research question one, "What are the effects of textual and visual annotations on listening comprehension and vocabulary acquisition?" relates to how well students recalled vocabulary items under the different annotation conditions. The results of the data analysis indicated that there was no significant difference among students' vocabulary posttest scores. There was a trend towards higher achievement for students who saw visual annotations ($p=.08$), but the difference did not meet the .05 significance level.

Curiously, the no-significant-difference findings of the current study contrast with those in several other previous reading and listening comprehension studies. Chun and Plass (1996a, 1996b), Plass, Chun, Mayer, Leutner (1998), Yoshii and Flaitz (2002) each found that visual annotations had a positive effect on vocabulary acquisition in their reading comprehension studies. The listening comprehension research by Jones and Plass (2002), and Jones (2004) also found significant positive effects for visual annotations on a vocabulary recall task.

Furthermore, Yeh and Wang (2003) studied the effects of three types of annotations (text, text + picture, and text + picture + sound) on vocabulary acquisition in a word-learning activity. Their study differed from others discussed here in the nature of the task; that is, they presented students with individual words rather than words in the context of a story or narrative. When presenting isolated words, they found that text + picture annotations had a significant positive effect on vocabulary acquisition as compared to either text alone, or text + picture + sound annotations.

One reason the visual annotations may not have made a difference for vocabulary recall is that the multiple-choice assessment may have been too easy for students. The overall average score on the vocabulary posttest was 21.27 out of 25 possible points. As Jones (2004) notes in her research, vocabulary recognition tasks, such as multiple choice questions, are easier for learners than vocabulary recall tasks, such as free-writing a translation of a given word.

Similar to the current study, Jones and Plass (2002) used a recognition-type, multiple-choice measure and found significant differences in the effect of visual and textual annotations on vocabulary acquisition. However, in many previous studies in which significant differences were found for textual and visual annotations, the vocabulary acquisition assessments were the more demanding recall tasks (Chun and Plass, 1996; Dubois and Vial, 2000; Jones, 2004). In fact Yoshii and Flaitz (2002), and Yoshii (2006) conducted vocabulary acquisition studies in the context of reading comprehension tasks that included a vocabulary-production measure and a vocabulary-recognition measure. Results in both studies

indicated a significant positive effect for visual annotations on the vocabulary production measure, but not on the recognition measure. Therefore, it is possible that a more rigorous measure of vocabulary acquisition would have resulted in significantly different results in the current study.

Cognitive load

The second research question, "What are the effects of textual and visual annotations on cognitive load?" led us to analyze how hard participants had to work to understand the Spanish language and vocabulary in the listening activity. Previous researchers, such as Plass, Chun and Leutner (2003), suggested that cognitive load may vary for different levels of vocabulary annotations; however, none of the previous language acquisition studies have attempted to measure cognitive load. Jones and Plass (2002) and Jones (2004) encouraged research in the area of cognitive load. To address this gap in the research, the Cottam and Savenye (2008) study as well as the current study included five cognitive load questions based on the NASA-TLX measure originally developed by Hart and Staveland (1988). Unfortunately, due to a programming error that was not discovered until late in the data collection process, valid responses to the cognitive load questions were only recorded for 35 of the 95 participants in the current study.

The results of the current study are similar to Cottam and Savenye (2008), which is expected due to the similar research design and similar number of participants. Cottam and Savenye (2008) speculated that the measure used in their study may not have been sensitive enough to detect differences in cognitive

load due to the fact that the cognitive load questions were rated on a 10-point scale rather than the original NASA-TLX 21-point scale. However, the findings of the current study further reinforce that the scale may not have had any impact on the previous results. Similar to Cottam and Savenye (2008), only one of the five cognitive load questions, "feeling of success," resulted in significant difference among treatment groups. None of the other questions, which related to task demand, hard work, navigation ease, or stress levels resulted in significant differences among participant groups. Each of the questions will be discussed in the order they appeared to students.

Task demand. Although not explicitly stated by other researchers, task demand relates to the intrinsic dimension of cognitive load since the question measured how naturally complex the learning task was for students. Although there were no significant differences among treatment groups, the results of the task-demand question indicated that accessing annotations of any type tended to reduce the effort the task required. The both-annotations group rated the task demand lowest, and the group that received no annotations rated it more demanding.

In Cottam and Savenye (2008), the visual-illustrations-only group rated the task more difficult than the other three groups. The researchers reviewed survey responses and hypothesized that some of the images used in the activity may not have been clear enough for students and may have made the task more difficult. All images that participants commented about in Cottam and Savenye (2008) were modified and clarified for the current study. It appears that the

changes may have made a difference since the visual-annotations-included group rated this task-demand question lower than the visual-annotations-excluded group.

Hard work. The cognitive load question asking students to rate how hard they worked most closely aligns with the notion of germane cognitive load. The question was designed to measure the amount of effort a student had to put into understanding the content of the listening activity. Analysis revealed no significant difference among treatments for this question. Nonetheless, it appears that the students who had to work the hardest on the activity were those who only accessed visual annotations. The students who accessed both types of annotations rated this question slightly lower than all other groups. However, the ratings are so close that no trend is evident among the three other treatments.

Feeling of success. The only one of the five cognitive load questions that resulted in a significant difference among treatment groups was the question, "How successful do you think you were in your attempt to understand the contents of the learning environment?" Participants who accessed textual annotations rated their level of success higher than participants who did not see textual annotations at all. The highest rating was given by the students in the both-annotations group, closely followed by the textual-annotations-only group although there was no significant difference between those two groups. However, unlike the Cottam and Savenye (2008) study, there was no significant effect found for visual annotations.

Navigation effort. The question about how easy it was to navigate the program resulted in a significant interaction effect in the current study, although there were no main effects for either textual or visual annotations. The participants in the no-annotations group rated their navigation effort much lower than participants in other groups. The both-annotations treatment also rated the navigation effort as low, but students in the textual-annotations-only group and those in the visual-annotations-only group rated the navigation effort slightly higher. It appears that having no annotations made the program simpler to navigate as there was less to look at and access on the screen. Likewise, having access to both types of annotations seems to have made the navigation appear less difficult. However, when only one type of annotation was present, students may have felt the navigation was not as clear. This is a result that was not evident in the Cottam and Savenye (2008) pilot study, in which no significant effects were found for this question.

This finding is interesting because the actions a participant would take on each screen to navigate through the program were identical for all treatments. The difference was that clicking on key words on the left of the screen in different treatments resulted in different content being displayed on the right side of the screen. Actual navigation within the screen and from screen to screen was not different at all for the four treatments.

Stress levels. There was no significant difference among participants' responses to the question, "How stressed did you feel during the learning task?" in the Cottam and Savenye (2008) pilot, nor in the current study. It appears that

students who only accessed visual annotations rated their stress levels higher than the rest of the groups, but again, it was not significantly different from the others. The lowest stress level was reported by students in the text-annotations-only group, followed by the both-annotations group and then the no-annotations group. The trend, although not significant statistically, seems to be towards higher stress levels when there are only visual annotations on-screen and no text to explain what they mean. This trend was not evident at all in Cottam and Savenye (2008) either, in which participants rated all four treatments almost identically to one another.

Student attitudes

The third research question, "What are the effects of textual and visual annotations on student attitudes?" relates to what students thought about the quality and effectiveness of the program and its different features. Participants rated most of the questions in the survey very positively. Only two of the fourteen questions were rated lower than a 4 on a 5-point scale, indicating that the design and features of the program were perceived as relevant, interesting and effective in many ways. This result confirms initial findings of Cottam and Savenye (2008). In fact, the same two questions regarding the helpfulness of individual word pronunciation were rated lower than all other questions in the 2008 study and in the current study. This is not surprising, because both studies used similar materials.

However, two questions that were rated lower than a four in the 2008 study were rated higher than a four in the current study. Those two statements

were, “The activity helped me to learn new vocabulary” and “I would like to have more listening activities of this type to help me understand spoken Spanish.” It appears that the small changes that were made to the visual annotations in the current study, clarifying some keywords that were previously ambiguous, may have made a difference in student perceptions.

Analyses in the current study confirmed that participants were consistently positive about the program. There were no significant differences among the groups on four of the five factors that were analyzed in the factor analysis. Only the “attitudes toward textual annotations” factor resulted in a significant difference among treatments. Those students who accessed both textual and visual annotations rated the effectiveness of the textual annotations higher than the students who accessed textual annotations alone. Perhaps seeing the visual annotations with the text helped students to visualize the vocabulary items and the narrative more effectively.

Time in program

In addition to the three major research questions, data were also collected on how long participants spent in the listening activity. It was anticipated that time-in-program might differ and might impact posttest and survey results. However, statistical analysis indicated no significant differences among treatment groups. This was somewhat surprising, because having more annotations, either textual or visual, was likely to have prompted participants to spend more time on each listening screen. However, this was not the case. No matter what annotations they received, students spent about the same amount of time in the program.

Perhaps students who did not have the assistance of vocabulary annotations ended up spending more time trying to figure out the words, thus equalizing the time in-program among the groups.

Implications for instructional design

Results of the current study have some implications for the design of multimedia listening activities in online language classes.

The current research confirms that students like the addition of annotations in listening comprehension tasks. Textual annotations, in particular help students to feel more successful in a listening activity of this type. The survey responses indicate that students feel they should have more multimedia annotations in activities of this type. In fact, students consistently asked for more or improved annotations, regardless of the treatment group they were in. Students who accessed no annotations at all felt that they were missing. Students who saw one type of annotation wanted the other annotation type as well. Students who accessed both types of annotations still wanted more visuals, more keywords, or videos. Perhaps the judicious inclusion of multimedia into listening activities will prove motivational for students and may decrease their perception of cognitive load.

It also appears that offering students vocabulary annotations in listening tasks may help them to understand the spoken language. Although the annotations did not have a significant effect on all of the measures in the current study, there is evidence that annotations will not hurt and will probably help students understand language and remember key words. In the online environment where

students cannot see their instructors, students may hear language out of context and may not be able to understand it without extralinguistic cues that textual and visual annotations provide. In order to make the language students hear more comprehensible to them, words that are not familiar to them should be annotated in some way.

Future research

The current study has extended our understanding of design elements that will help students improve listening comprehension and learn new vocabulary in language classes. However, there is a need for further research in a number of areas.

The Cottam and Savenye (2008) study and the current study are unique among multimedia listening comprehension research studies in that they were conducted in the setting of an online class, with students working independently. With the current growth of online language courses, the need for language acquisition research in this new and ever-changing environment is becoming more urgent. Listening comprehension, specifically, is a language skill that has not received adequate attention in the multimedia research. Research in the area was urged by Plass and Jones (2005), yet outside of the current research, little progress has been made. The results of the current study support the need for further research since significant effects found in the classroom and laboratory were not completely confirmed.

Cognitive load measurement continues to be a challenge for researchers (Paas, Tuovinen, Tabbers, and Van Gerven, 2001; Brunken, Plass and Leutner,

2003). Cognitive load is most commonly measured through survey instruments, however there are other methods being investigated. For example, Antonenko, Niederhauser, and Thompson (2007) used Electroencephalography (EEG) in their research into the effects of hypertext leads on cognitive load. They found that that EEG was a useful measure of cognitive load during continuous reading activities, although neither the participants' self-reports, nor the EEG showed differences among their experimental treatments. Until biometrics such as EEG are investigated further in the measurement of cognitive load, it appears that self-reports such as the NASA-TLX will continue to be used in cognitive load measurement. Yet, such biometric instruments will be limited to use in the laboratory setting. Further research into cognitive load measures in the online environment would help to clarify the effects of activities such as the listening activity in the current study on cognitive load.

Another area of research has been prompted by the students' responses to the open-ended survey questions. They expressed a desire for other annotation types, such as video and animation, in the listening comprehension activity. Some researchers, such as Chun and Plass (1996), Al-Seghayer (2001) have studied the effects of video on reading comprehension and have found it effective for vocabulary acquisition. This same research has not been extended to the skill of listening comprehension as yet, however.

Multimedia annotations and video in particular, may play an important role in motivating students and focusing their attention on the learning task. Astleitner and Wiesner (2004) have proposed that motivation should be added to

the purely cognitive theory of multimedia learning that Mayer (2001) has described. Astleitner and Wiesner (2004) propose a model of multimedia learning that includes motivational processing and mental resource management. They propose that motivation may have effects on cognitive load as well. Further research is needed to investigate their model by measuring motivation and cognitive load in a multimedia learning environment such as the one used in the current study.

Conclusion

Language study opportunities in the online environment continue to grow and there are a myriad of possible activities in which students may engage as they try to acquire new language skills. The ubiquity of multimedia language learning resources online is something that students and instructors have come to expect. It is imperative, therefore, that research into how to make those materials effective for students continue to advance.

REFERENCES

Al-Seghayer, K. (2001). The effect of multimedia annotation modes on L2 vocabulary acquisition: A comparative study. *Language Learning and Technology. 5(1)*, 202-232.

Allen, I.E. & Seaman, J. (2009). Staying the course: Online education in the United States, 2009. Needham, MA: Sloan Consortium.

American Council on the Teaching of Foreign Languages (2001). Standards for foreign language learning in the 21st century, 3rd ed. Executive Summary, Retrieved December 2, 2008 from http://www.actfl.org/files/public/StandardsforFLLexecsumm_rev.pdf

Antonenko, P., Niederhauser, D.S., & Thompson, A. (2007). Optimization of cognitive load in conceptually rich hypertext: Effect of leads. In D.S. McNamara & J.G. Trafton (Eds.), *Proceedings of the 29th Annual Cognitive Science Society* (pp. 1707-1709). Austin, TX: Cognitive Science Society.

Ariew, R. & Ercetin, G. (2004). Exploring the potential of hypermedia annotations for second language reading. *Computer Assisted Language Learning, 17*(2), 237–259.

Astleitner, H. & Wiesner, C. (2004). An integrated model of multimedia learning and motivation. *Journal of Educational Multimedia and Hypermedia, 13*(1), 3-21.

Baddeley, A. (1986). Working Memory. New York: Oxford University Press.

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

Baddeley, A. (1999). The essentials of human memory. Hove, UK: Psychology Press Ltd.

Brünken, Plass & Leutner (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist. 38*(1), 53-61.

Chun, D.M. & Plass, J.L. (1996a). Effects of multimedia annotations on vocabulary acquisition. *Modern Language Journal*, *80*, 183-198.

Chun, D.M. & Plass, J.L. (1996b). Facilitating reading comprehension with multimedia. *System. 24*(4), 503-519.

Cottam, M.E. & Savenye, W. (2008, November). *The effects of visual and textual annotations on Spanish listening comprehension, incidental vocabulary acquisition and cognitive load*. Poster presented at the annual conference of the Association of Educational Communication and Technology, Orlando, FL

Dubois, M. & Vial, I. (2000). Multimedia Design: The Effects of Relating Multimodal Information. *Journal of Computer Assisted Learning, 16*, 157-165.

Gass, S.M. & Selinker, L. (1994). Second Language Acquisition. New Jersey: Lawrence Erlbaum Associates.

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.

Gevins, Smith, Leong, McEvoy, Whitfield, Du & Rush (1998). Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. *Human Factors, 40*, 79-91.

Ginther, A. (2002). Context and content visuals and performance on listening stimuli. *Language Testing, 19*(2), 133-167.

Krashen, S. (1982). Principles and Practice in Second Language Acquisition. London: Pergamon.

Krashen, S. (1985). The Input Hypothesis: Issues and Implications. New York: Longman.

Jones, L.C. & Plass, J.L. (2002). Supporting listening comprehension and vocabulary acquisition in french with multimedia annotations. *The Modern Language Journal, 86*, 546-561.

Jones, L. (2004). Testing L2 Vocabulary Recognition and Recall Using Pictorial and Written Test Items. *Language Learning and Technology, 8*(3), 122-143.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): results of experimental and theoretical research. In P. A. Hancock, & N. Meshkati (Eds.), Human mental workload. 139–183. Amsterdam: North-Holland.

Hill S. G., Iavecchia H. P., Byers J. C., Bittner A. C., Zaklad A. L. & Christ R. E. (1992) Comparison of four subjective workload rating scales. *Human Factors*, *34*, 429–439.

Lafford, B. A. & Salaberry, R. (2003). *Spanish Second Language Acquisition: State of the Science.* Washington D.C.: Georgetown University Press.

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.

Low, R. & Sweller, J. (2005). The modality principle in multimedia learning. In R. Mayer, (Ed.), *The Cambridge handbook of multimedia learning*, pp. 19-30. New York: Cambridge University Press.

Mayer, R. (2001). Multimedia learning. New York: Cambridge University Press.

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

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.

Moreno, R. & Mayer, R.E. (2000). Engaging students in active learning: The case for personalized multimedia messages. *Journal of Educational Psychology*, *92*(4), 724-733.

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

Paas, F., & Van Merrienboer, J. (1994). Variability of worked examples and transfer of geometrical problem solving skills: A cognitive-load approach. *Journal of Educational Psychology,* 86, 122-133.

Paas, F., Tuovinen, J.E., Tabbers, H. & Van Gerven, P.W.M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*(1), 63-71.

Penny, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory and Cognition, 17*, 398-422.

Pierce, C. A., Block, C. A. & Aguinis, H. (2004). Cautionary note on reporting eta-squared values from multifactor ANOVA designs. Educational and Psychological Measurement 64(6).

Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second language multimedia learning environment. *Journal of Educational Psychology, 90*(1), 25-36.

Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (2003). Cognitive load in reading a foreign language text with multimedia aids and the influence of verbal and spatial abilities. *Computers in Human Behavior, 19*(2), 221-243.

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

Sakar, A. & Ercetin, G. (2005). Effectiveness of hypermedia annotations for foreign language reading. *Journal of Computer Assisted Learning*. 21, 28-38.

Scheiter, K., Gerjets, P., & Catrambone, R. (2006). Making the abstract concrete: Visualizing mathematical solution procedures. *Computers in Human Behavior*, *22*(1), 9-25.

Su, Y. (2007). The impact of scaffolding type and prior knowledge in a hypermedia, problem-based learning environment. (Doctoral dissertation, Arizona State University, 2007). *Dissertation Abstracts International, 68,* 11.

Sweller, J. (1999). Instructional design in technical areas. ACER Press, Victoria, Australia.

Sweller, J. (2003). Evolution of human cognitive architecture. In B. Ross (Ed.), *The Psychology of Learning and Motivation, 43*, 215-266. San Diego: Academic Press.

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. Mayer, (Ed.), *The Cambridge handbook of multimedia learning*, pp. 19-30. New York: Cambridge University Press.

Yeh, Y. & Wang, C. (2003). Effects of multimedia vocabulary annotations and learning styles on vocabulary learning. *CALICO Journal. 21*(1), 131-144.

Yoshii, M. (2006). L1 and L2 glosses: Their effects on incidental vocabulary learning. *Language Learning & Technology, 10*(3). 85-101.

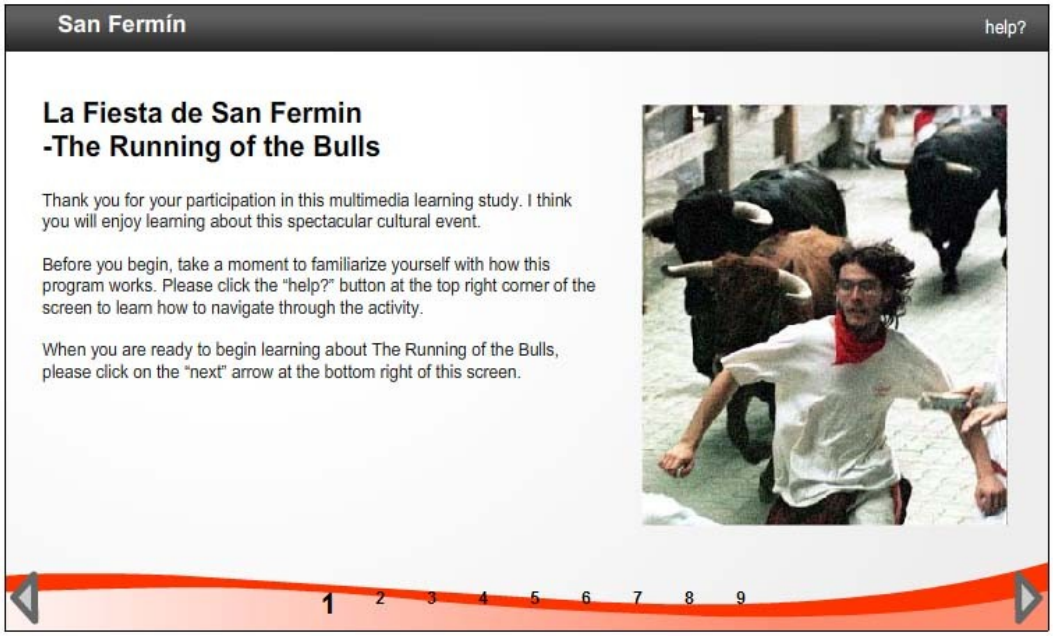
Yoshii, M. & Flaitz, J. (2002) Second language incidental vocabulary retention: The effect of text and picture annotation types. *CALICO Journal, 20*(1), 22-58.

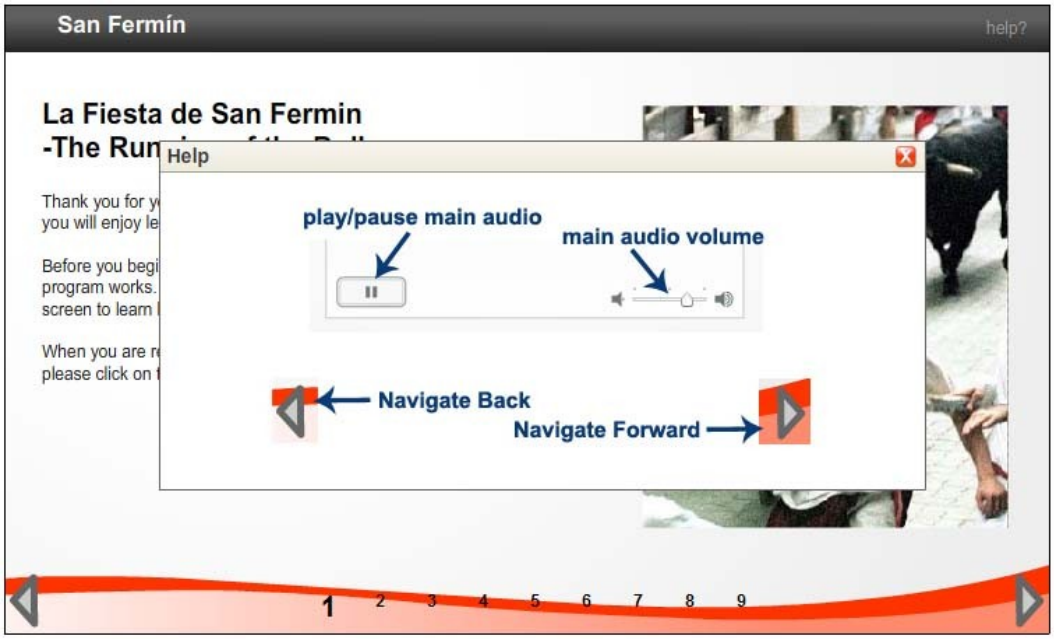
White, C. (2003). Language learning in distance education. New York: Cambridge University Press.

APPENDIX A

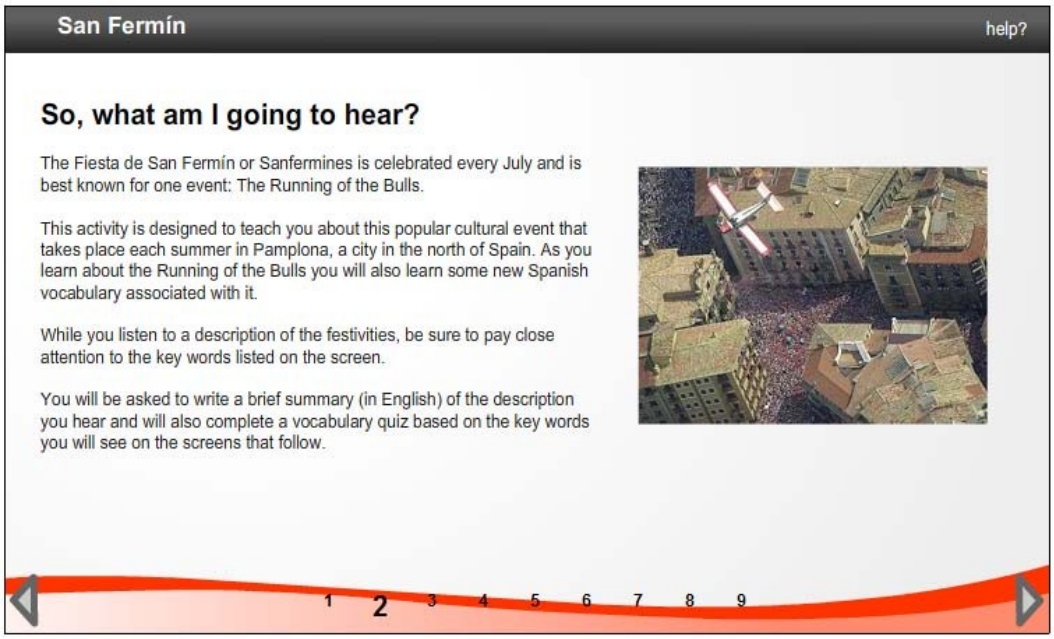
SCREENSHOTS OF THE BOTH-ANNOTATIONS TREATMENT

The following screenshots are from the both-annotation treatment of the *Sanfermines* Spanish listening activity program. Screens are arranged in order of appearance within the program. The three introductory screens are first, followed by each of the listening screens. Participants access five listening screens, but each screen is interactive, allowing students to control the audio and to access vocabulary annotations as they choose. To illustrate this interactivity, this appendix includes screenshots in the order students would encounter them if they were to access words in the list from top to bottom. The audio script of what students will hear upon accessing each of the five segments is included under the first listening screenshot for each segment. That initial screenshot is followed by individual screenshots illustrating each of the vocabulary annotations.





(Note: The screen above appears when the "help?" button in the upper right corner of the screen is pressed at any time throughout the program.)



San Fermín help?

First, some background…

Pamplona is an ancient city with narrow cobblestone streets and traditional plazas dating back to the middle ages. Many of the multi-storied houses lining the streets have open balconies to let in cooler summer breezes.

This historic setting provides a backdrop to some of the most exciting and dangerous moments any tourist would want to experience.

Thousands of people from all over the world visit Pamplona during the Sanfermines festival to experience the Running of the Bulls. Some stand safely behind street barracades or perch in balconies along the bulls' course. Others decide to run with the bulls, knowing that at any moment the massive animals could gore or trample them.

Listen as you hear more about this exciting event in Spanish on the next few screens.

1 2 3 4 5 6 7 8 9

San Fermín help?

PAUSED 0:00:08.228

Keywords

Click on each keyword for additional information.

El encierro
El espectáculo
La carrera
Los mozos
El toro
La calle estrecha
La ciudad

Pamplona y el encierro

1 2 3 4 5 6 7 8 9

Audio Script for Screen 4

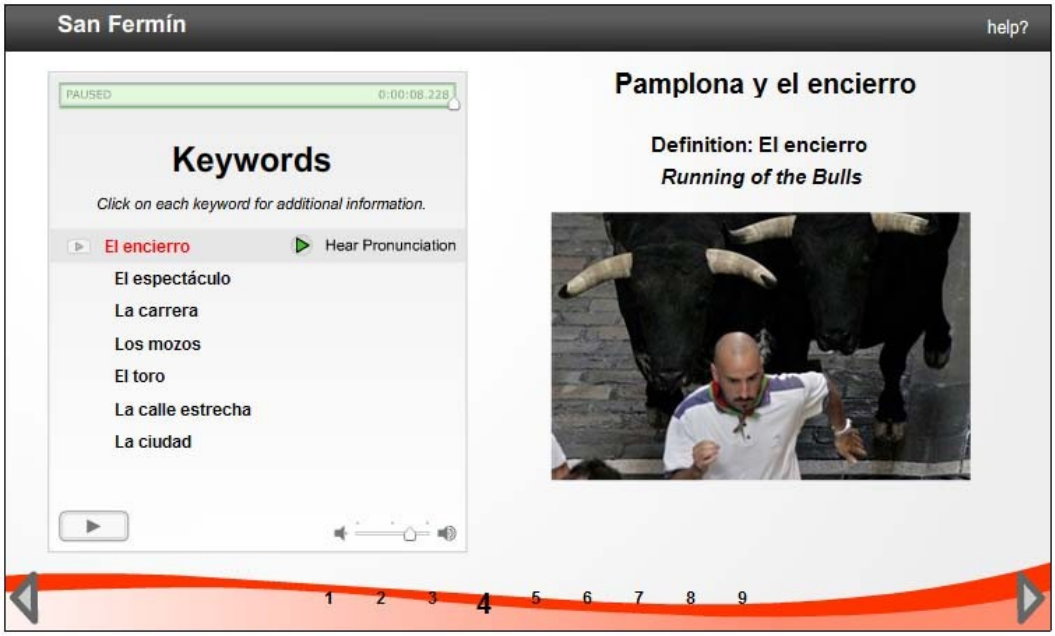
(Note: Script does not appear on screen at any time.)

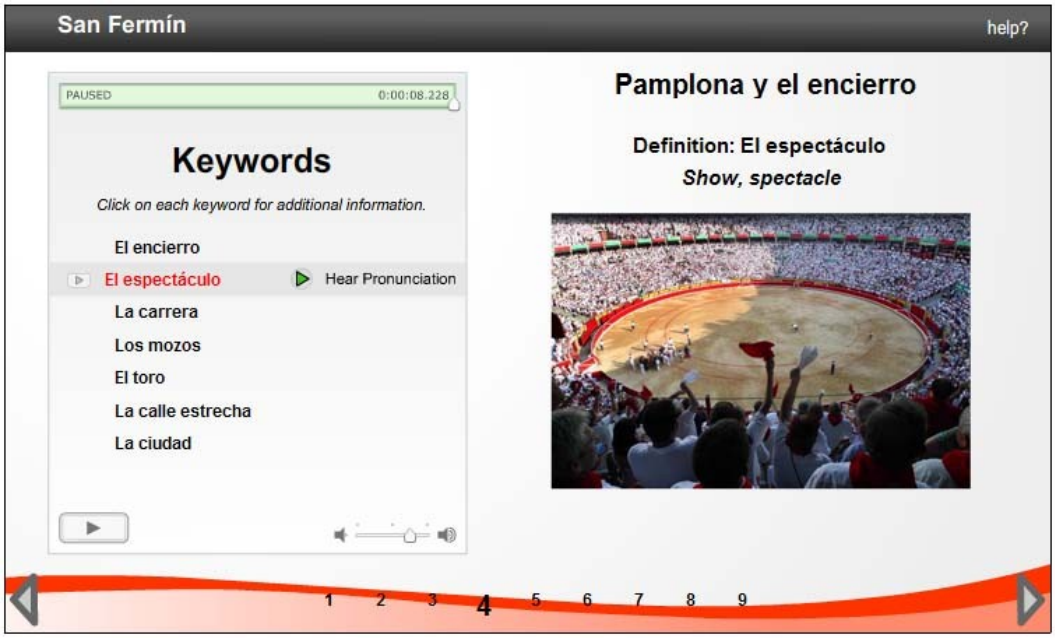
Pamplona y el encierro

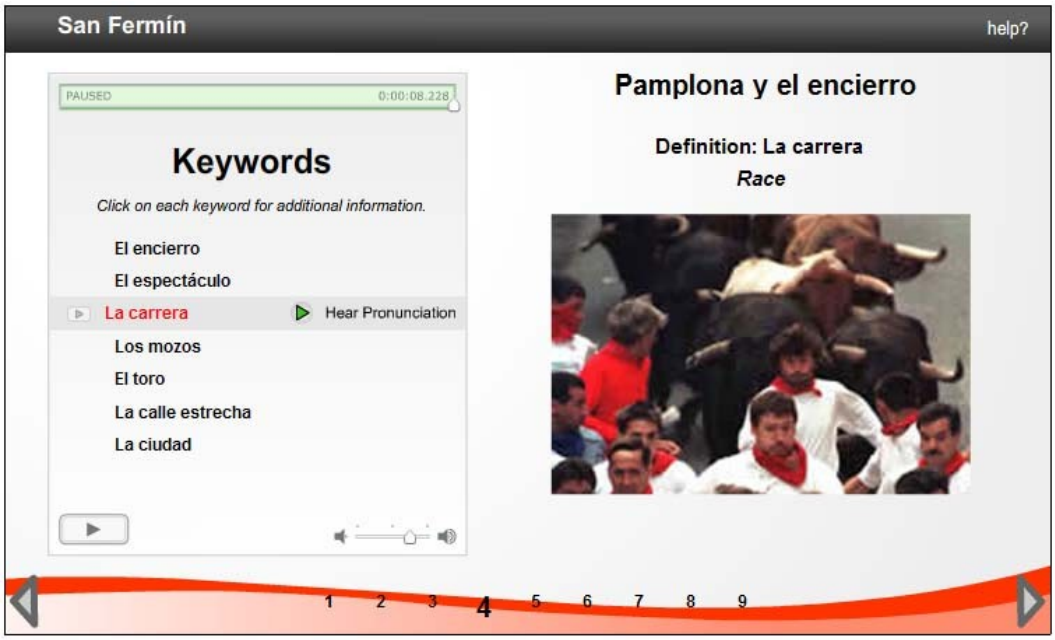
Pamplona, España es conocido por una fiesta, Sanfermines, y por un evento especial, el encierro. Es un espectáculo que dura solamente dos o tres minutos, pero son unos minutos vividos en colores blancos y rojos.

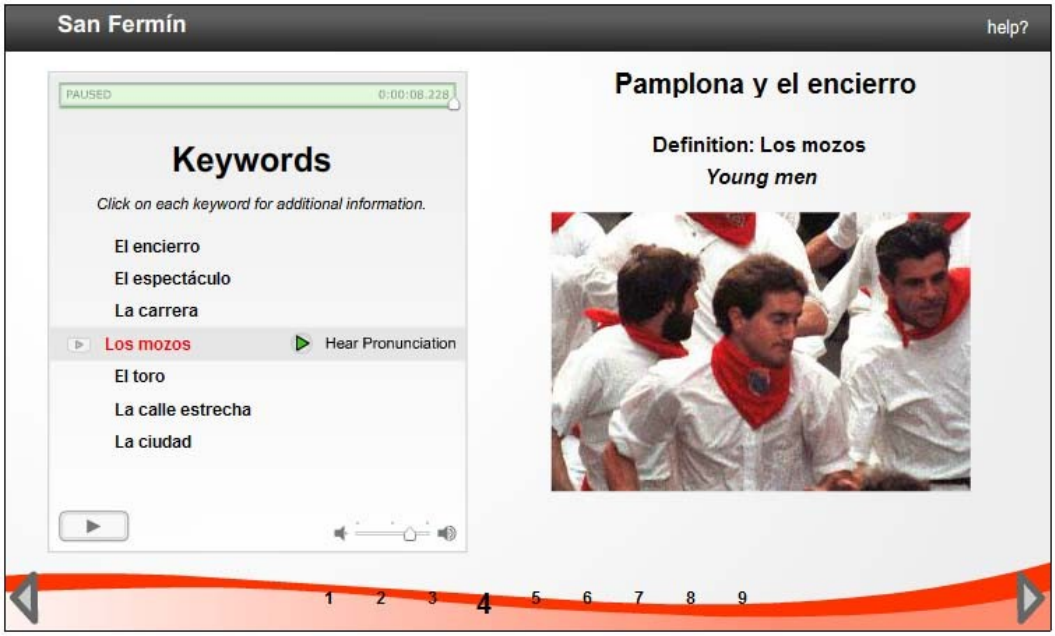
Cada año la fiesta de Sanfermines empieza el día siete de julio y termina el catorce de julio. El encierro es una parte importante de la fiesta que ocurre temprano por la mañana cada día de la fiesta.

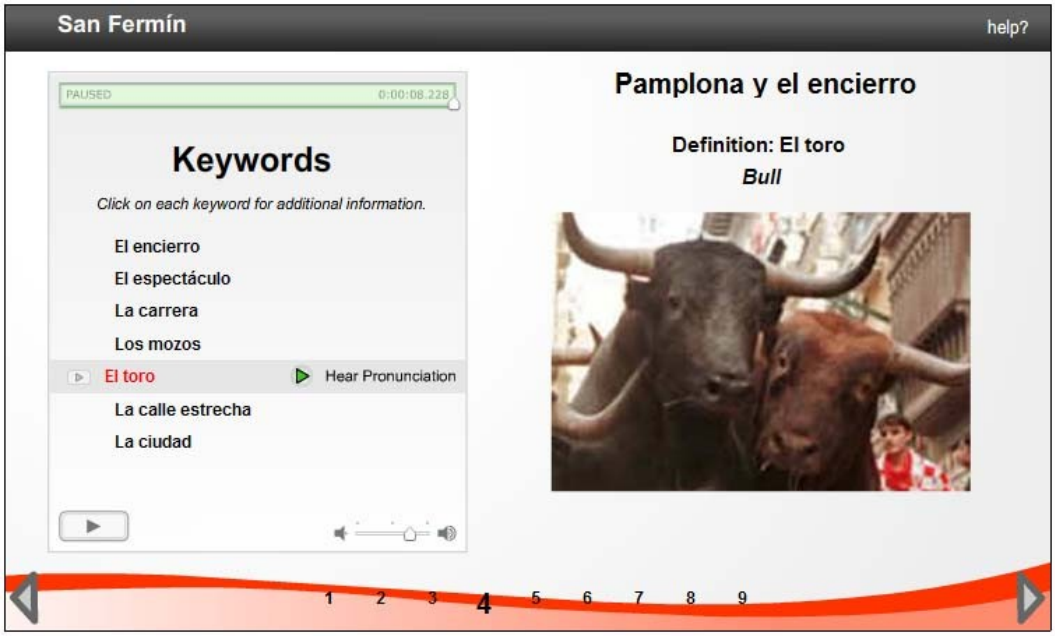
Es una carrera entre mozos y toros, corriendo por las calles estrechas de la ciudad, desde los corrales hasta la plaza de toros.

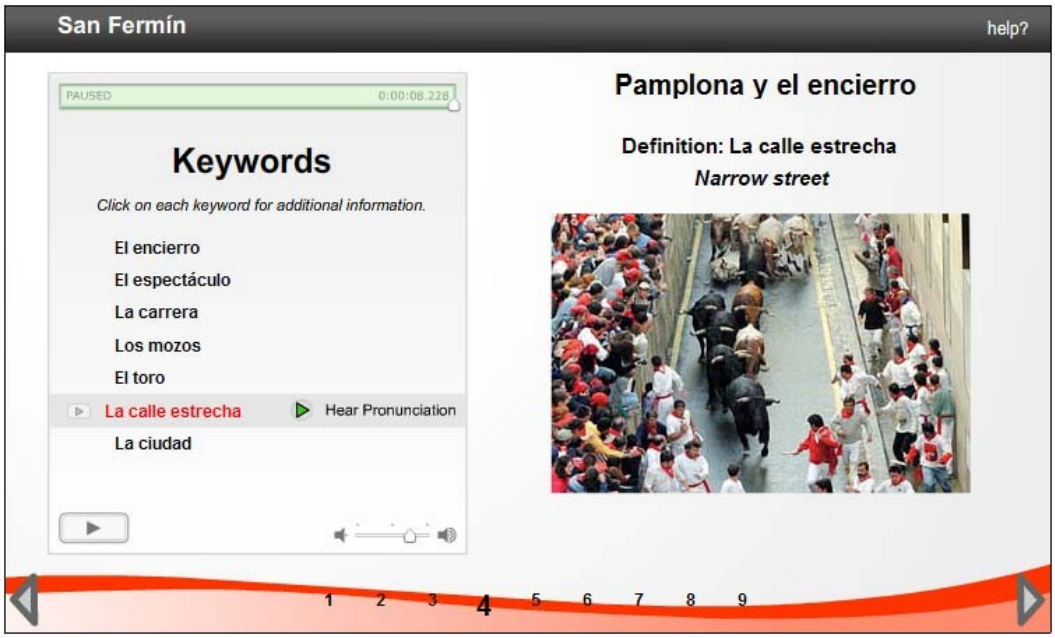

San Fermín
help?
PAUSED
0:00:08.228
Keywords
Click on each keyword for additional information.
El encierro
Hear Pronunciation
El espectáculo
La carrera
Los mozos
El toro
La calle estrecha
La ciudad
Pamplona y el encierro
Definition: El encierro
Running of the Bulls
1 2 3 4 5 6 7 8 9

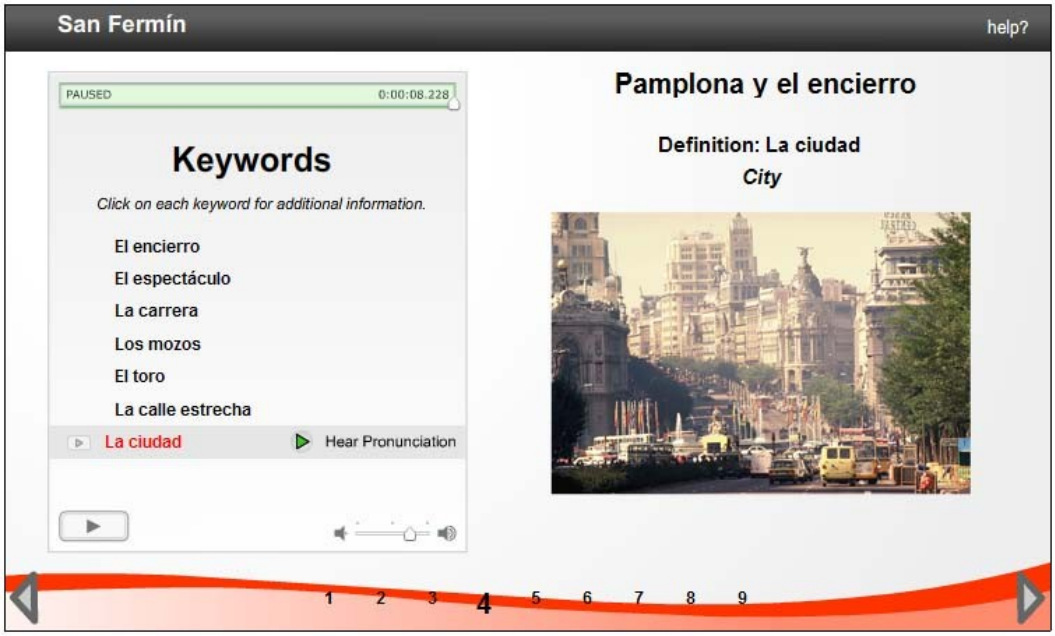

San Fermín
help?
PAUSED
0:00:08.228
Keywords
Click on each keyword for additional information.
El encierro
El espectáculo
Hear Pronunciation
La carrera
Los mozos
El toro
La calle estrecha
La ciudad
Pamplona y el encierro
Definition: El espectáculo
Show, spectacle
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:08.228
Keywords
Click on each keyword for additional information.
El encierro
El espectáculo
La carrera
Hear Pronunciation
Los mozos
El toro
La calle estrecha
La ciudad
Pamplona y el encierro
Definition: La carrera
Race
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:08.228
Keywords
Click on each keyword for additional information.
El encierro
El espectáculo
La carrera
Los mozos
Hear Pronunciation
El toro
La calle estrecha
La ciudad
Pamplona y el encierro
Definition: Los mozos
Young men
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:08.228
Keywords
Click on each keyword for additional information.
El encierro
El espectáculo
La carrera
Los mozos
El toro
Hear Pronunciation
La calle estrecha
La ciudad
Pamplona y el encierro
Definition: El toro
Bull
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:08.228
Keywords
Click on each keyword for additional information.
El encierro
El espectáculo
La carrera
Los mozos
El toro
La calle estrecha
Hear Pronunciation
La ciudad
Pamplona y el encierro
Definition: La calle estrecha
Narrow street
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:08.228
Keywords
Click on each keyword for additional information.
El encierro
El espectáculo
La carrera
Los mozos
El toro
La calle estrecha
La ciudad
Hear Pronunciation
Pamplona y el encierro
Definition: La ciudad
City
1 2 3 4 5 6 7 8 9

Audio script for Screen 5

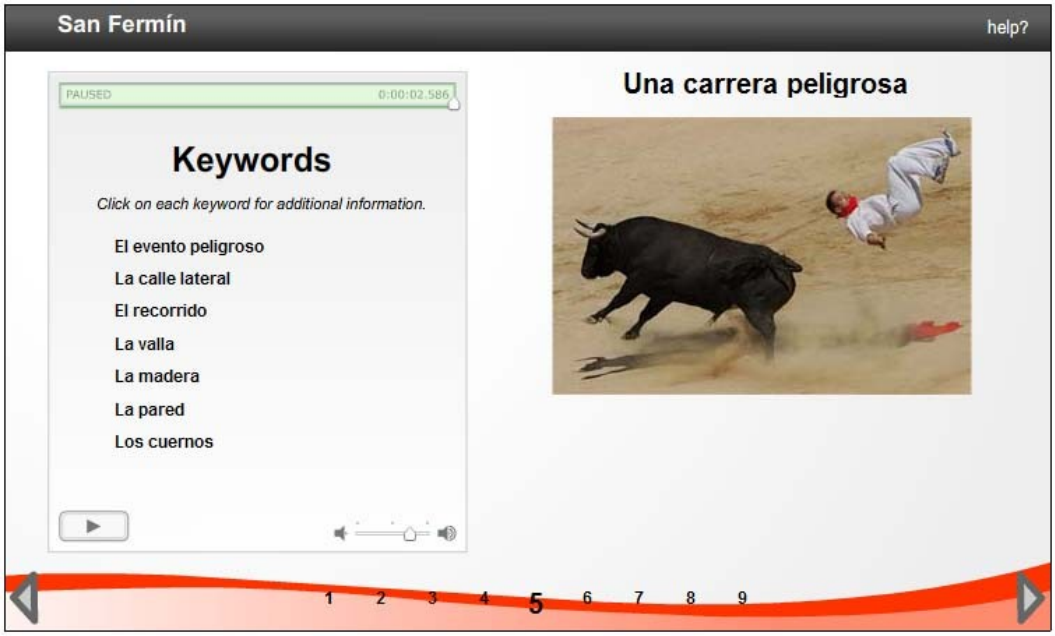
(Note: Script does not appear on screen at any time.)

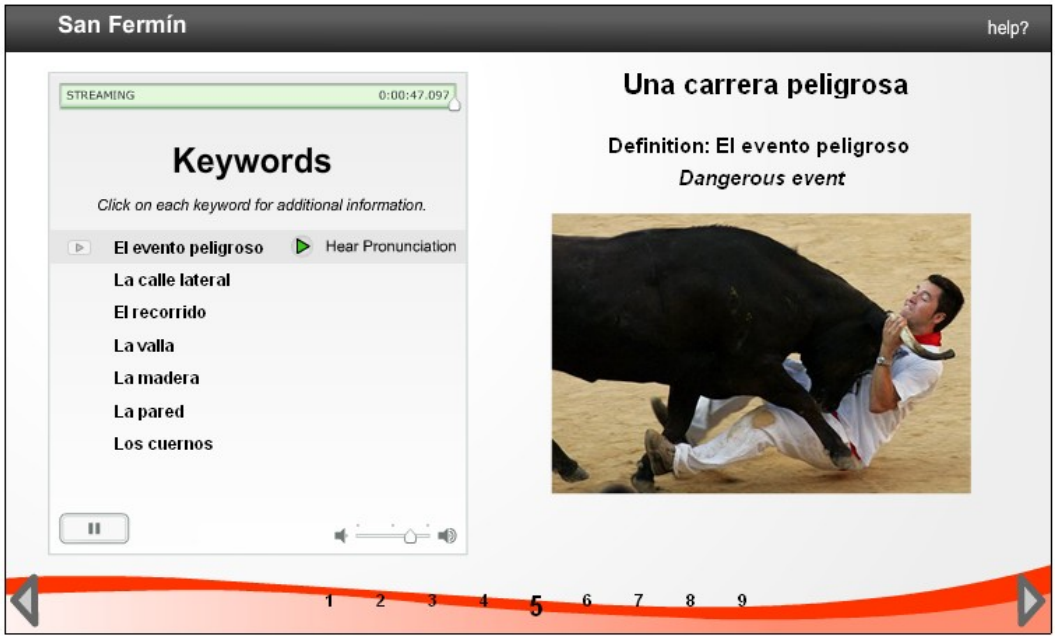
Una carrera peligrosa

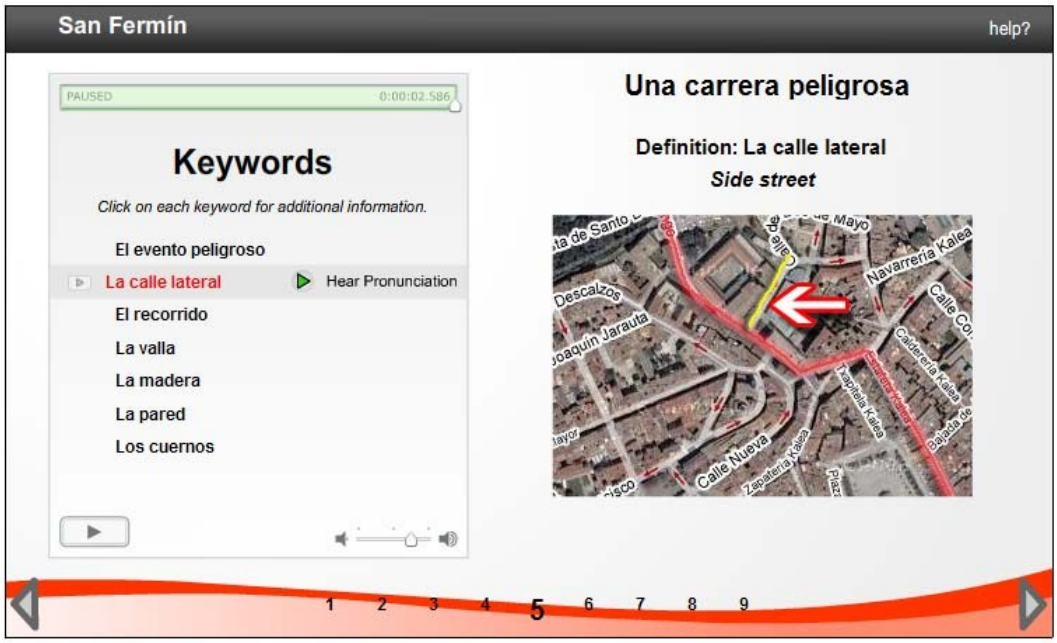
El encierro es un evento peligroso para los participantes. Cada mañana hay seis toros y cientos de personas, la mayoría hombres jóvenes, que participan en la carrera.

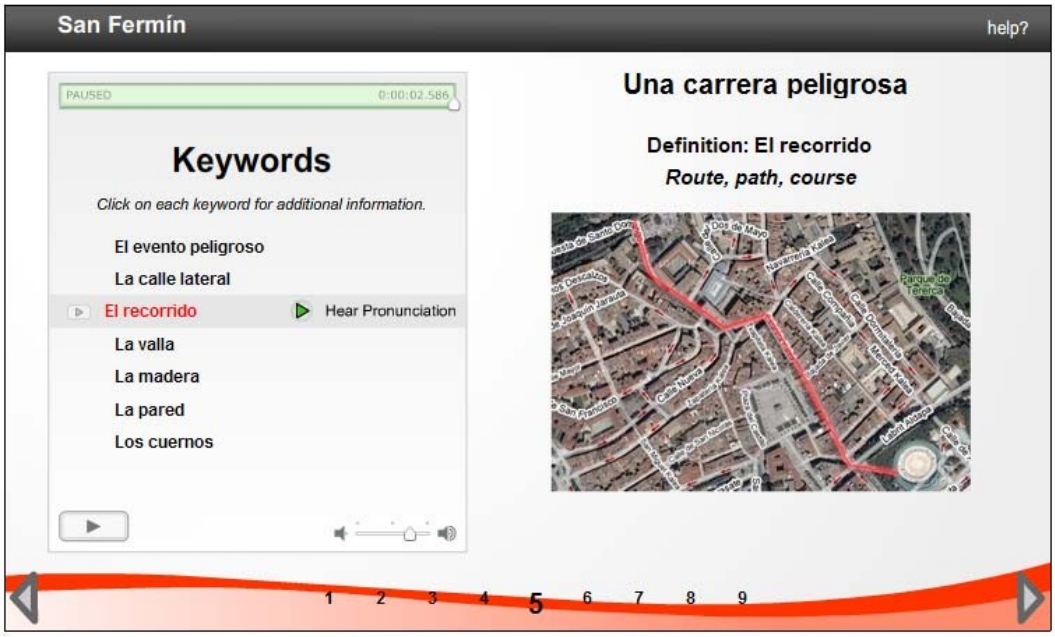
Antes del encierro la policía corta las calles laterales con unas vallas de madera para dirigir a los participantes y toros y para proteger a los espectadores.

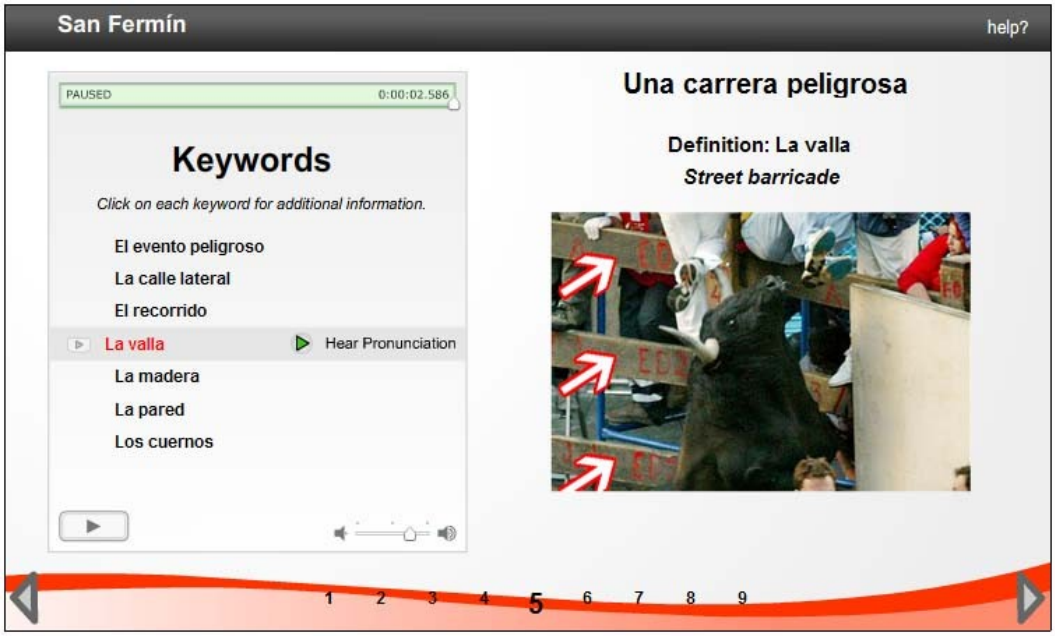
Los mozos que corren con los toros están en una calle estrecha entre paredes y vallas y no pueden salir fácilmente. Los toros son enormes y tienen cuernos largos y apuntados.

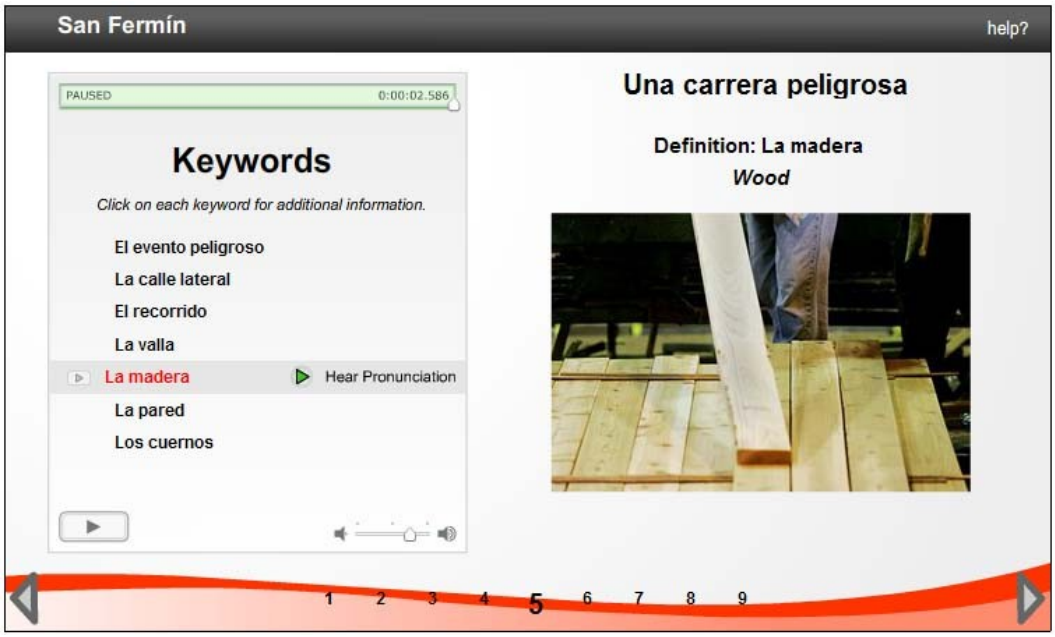

San Fermín
help?
PAUSED
0:00:02.586
Keywords
Click on each keyword for additional information.
El evento peligroso
La calle lateral
El recorrido
La valla
La madera
La pared
Los cuernos
Una carrera peligrosa
1 2 3 4 5 6 7 8 9

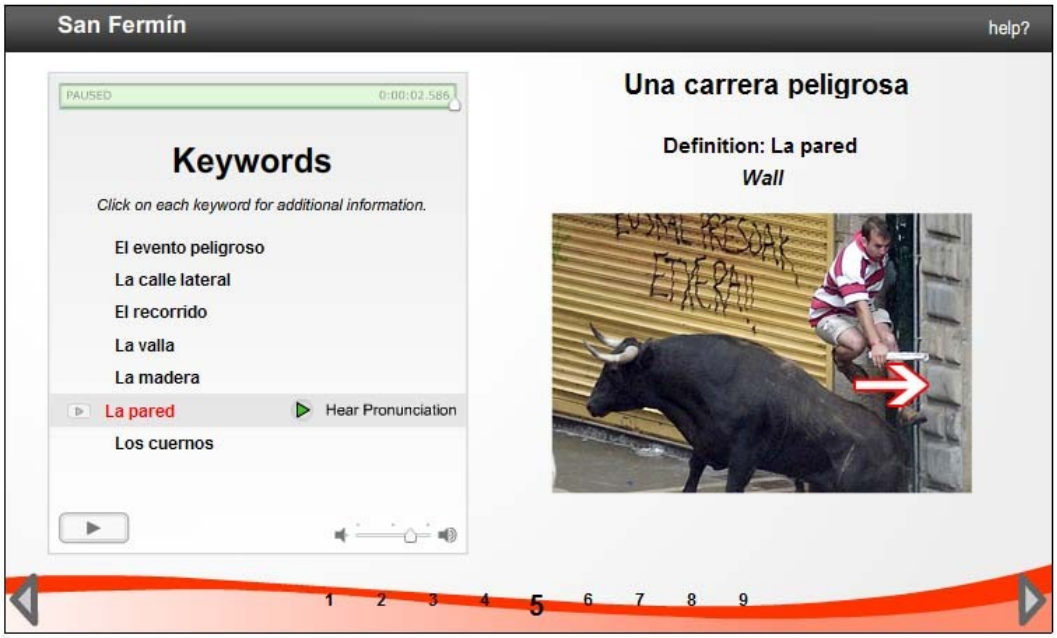

San Fermín
help?
STREAMING
0:00:47.097
Keywords
Click on each keyword for additional information.
El evento peligroso
Hear Pronunciation
La calle lateral
El recorrido
La valla
La madera
La pared
Los cuernos
Una carrera peligrosa
Definition: El evento peligroso
Dangerous event
1 2 3 4 5 6 7 8 9

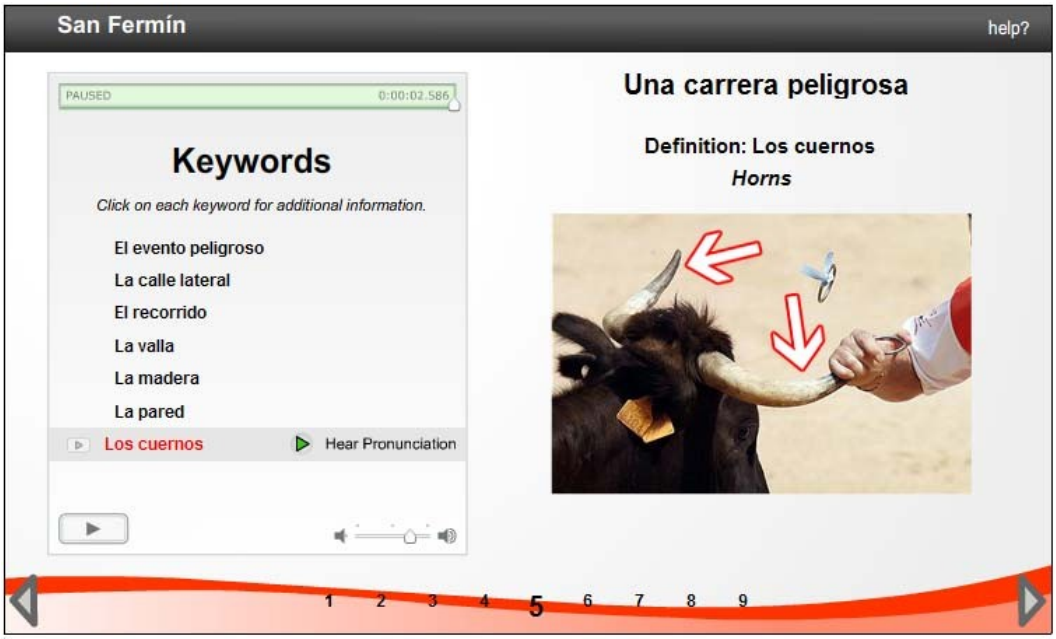

San Fermín
help?
PAUSED
0:00:02.586
Keywords
Click on each keyword for additional information.
El evento peligroso
La calle lateral
Hear Pronunciation
El recorrido
La valla
La madera
La pared
Los cuernos
Una carrera peligrosa
Definition: La calle lateral
Side street
Navarreria Kalea
Descalzos
Calle Nueva
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:02.586
Keywords
Click on each keyword for additional information.
El evento peligroso
La calle lateral
El recorrido
Hear Pronunciation
La valla
La madera
La pared
Los cuernos
Una carrera peligrosa
Definition: El recorrido
Route, path, course
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:02.586
Keywords
Click on each keyword for additional information.
El evento peligroso
La calle lateral
El recorrido
La valla
Hear Pronunciation
La madera
La pared
Los cuernos
Una carrera peligrosa
Definition: La valla
Street barricade
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:02.586
Keywords
Click on each keyword for additional information.
El evento peligroso
La calle lateral
El recorrido
La valla
La madera
Hear Pronunciation
La pared
Los cuernos
Una carrera peligrosa
Definition: La madera
Wood
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:02.586
Keywords
Click on each keyword for additional information.
El evento peligroso
La calle lateral
El recorrido
La valla
La madera
La pared
Hear Pronunciation
Los cuernos
Una carrera peligrosa
Definition: La pared
Wall
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:02.586
Keywords
Click on each keyword for additional information.
El evento peligroso
La calle lateral
El recorrido
La valla
La madera
La pared
Los cuernos
Hear Pronunciation
Una carrera peligrosa
Definition: Los cuernos
Horns
1 2 3 4 5 6 7 8 9

Audio script for Screen 6

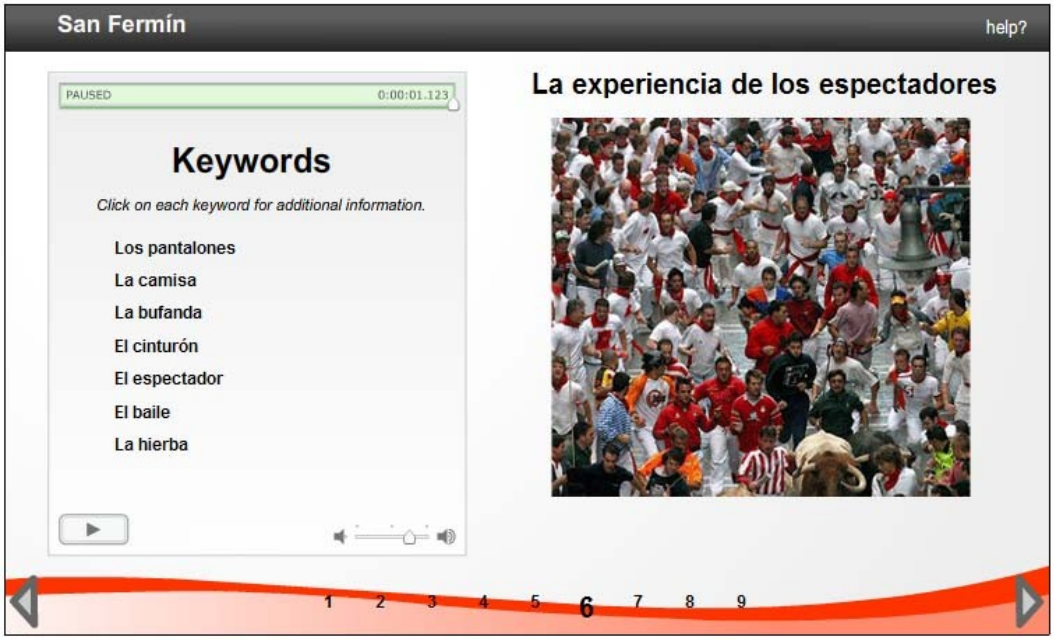
(Note: Script does not appear on screen at any time.)

La experiencia de los espectadores

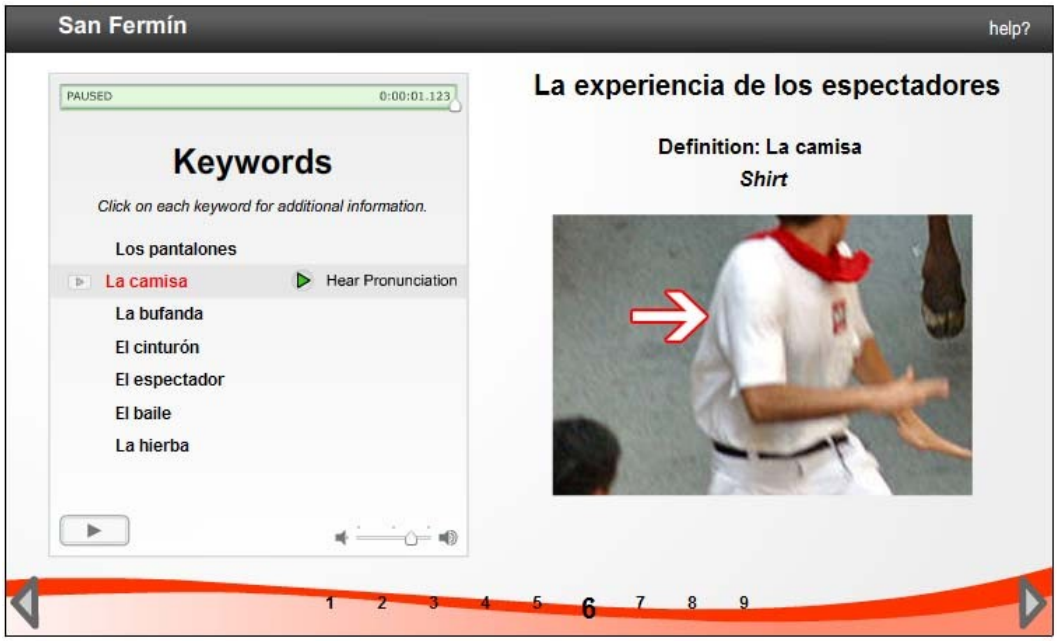
Miles de personas viajan a Pamplona simplemente para ver el espectáculo, no para correr. Como parte de la fiesta, todos se visten de blanco y rojo. Generalmente llevan pantalones y camisas blancos y bufandas y cinturones rojos.

Los espectadores festejan con baile y bebida toda la noche en los bares y clubes. Hay hoteles, pensiones y campings para alojarse. Pero algunas personas duermen en sus coches en la calle o en la hierba en el parque por dos o tres horas.

Por la mañana todos están cansados y muchos están borrachos, pero están animados para el encierro.

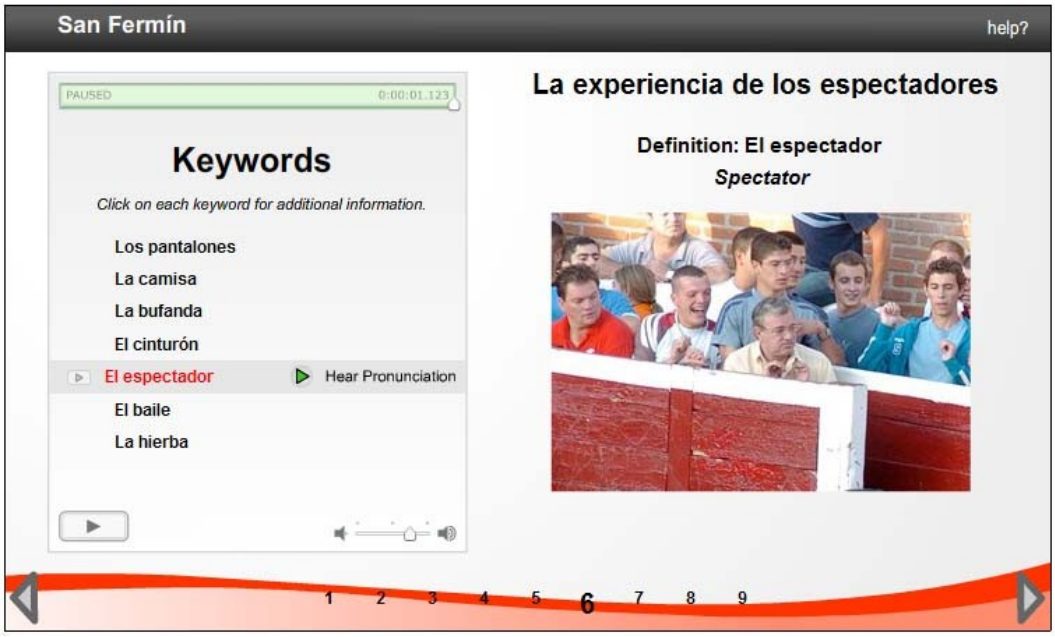

San Fermín
help?
PAUSED
0:00:01.123
Keywords
Click on each keyword for additional information.
Los pantalones
La camisa
La bufanda
El cinturón
El espectador
El baile
La hierba
La experiencia de los espectadores
1 2 3 4 5 6 7 8 9

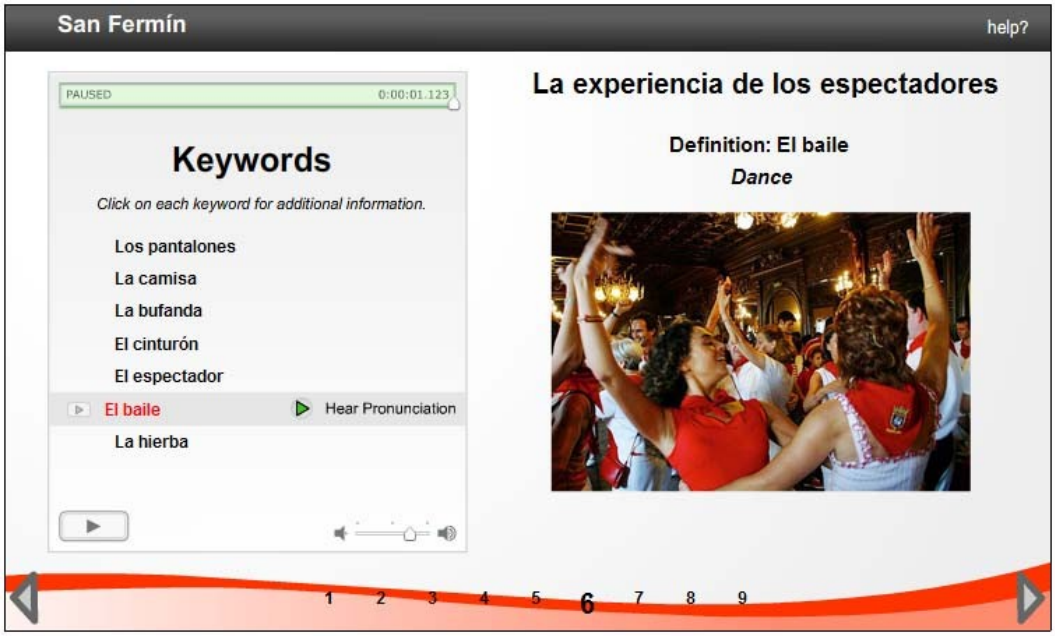

San Fermín
help?
PAUSED
0:00:01.123
Keywords
Click on each keyword for additional information.
Los pantalones
Hear Pronunciation
La camisa
La bufanda
El cinturón
El espectador
El baile
La hierba
La experiencia de los espectadores
Definition: Los pantalones
Pants
1 2 3 4 5 6 7 8 9

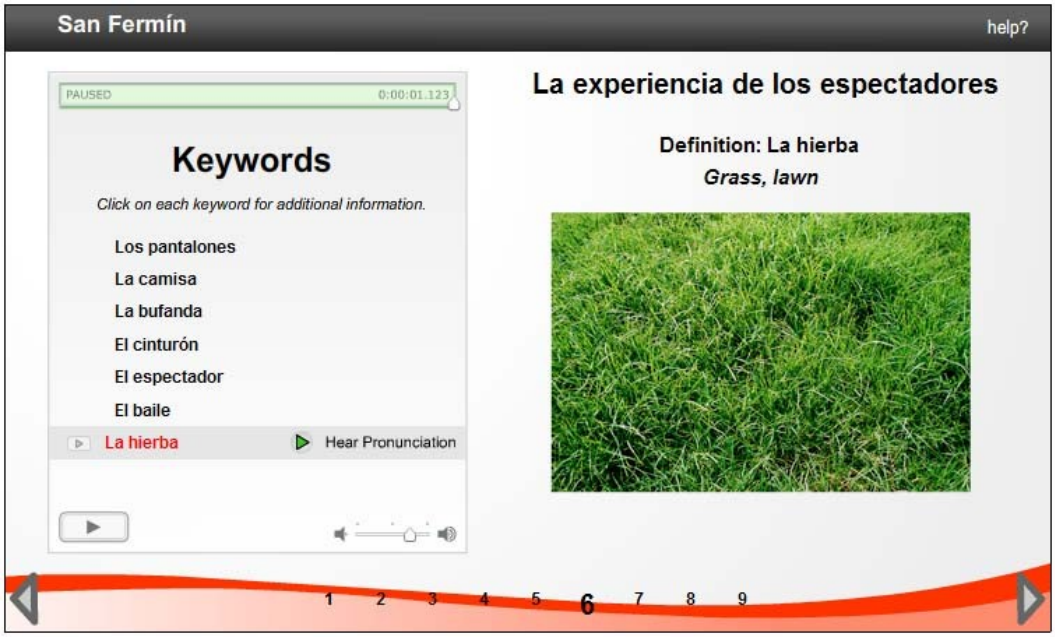

San Fermín
help?
PAUSED
0:00:01.123
Keywords
Click on each keyword for additional information.
Los pantalones
La camisa
Hear Pronunciation
La bufanda
El cinturón
El espectador
El baile
La hierba
La experiencia de los espectadores
Definition: La camisa
Shirt
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:01.123
Keywords
Click on each keyword for additional information.
Los pantalones
La camisa
La bufanda
Hear Pronunciation
El cinturón
El espectador
El baile
La hierba
La experiencia de los espectadores
Definition: La bufanda
Scarf
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:01.123
Keywords
Click on each keyword for additional information.
Los pantalones
La camisa
La bufanda
El cinturón
Hear Pronunciation
El espectador
El baile
La hierba
La experiencia de los espectadores
Definition: El cinturón
Belt
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:01.123
Keywords
Click on each keyword for additional information.
Los pantalones
La camisa
La bufanda
El cinturón
El espectador
Hear Pronunciation
El baile
La hierba
La experiencia de los espectadores
Definition: El espectador
Spectator
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:01.123
Keywords
Click on each keyword for additional information.
Los pantalones
La camisa
La bufanda
El cinturón
El espectador
El baile
Hear Pronunciation
La hierba
La experiencia de los espectadores
Definition: El baile
Dance
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:01.123
Keywords
Click on each keyword for additional information.
Los pantalones
La camisa
La bufanda
El cinturón
El espectador
El baile
La hierba
Hear Pronunciation
La experiencia de los espectadores
Definition: La hierba
Grass, lawn
1 2 3 4 5 6 7 8 9

Audio script for Screen 7

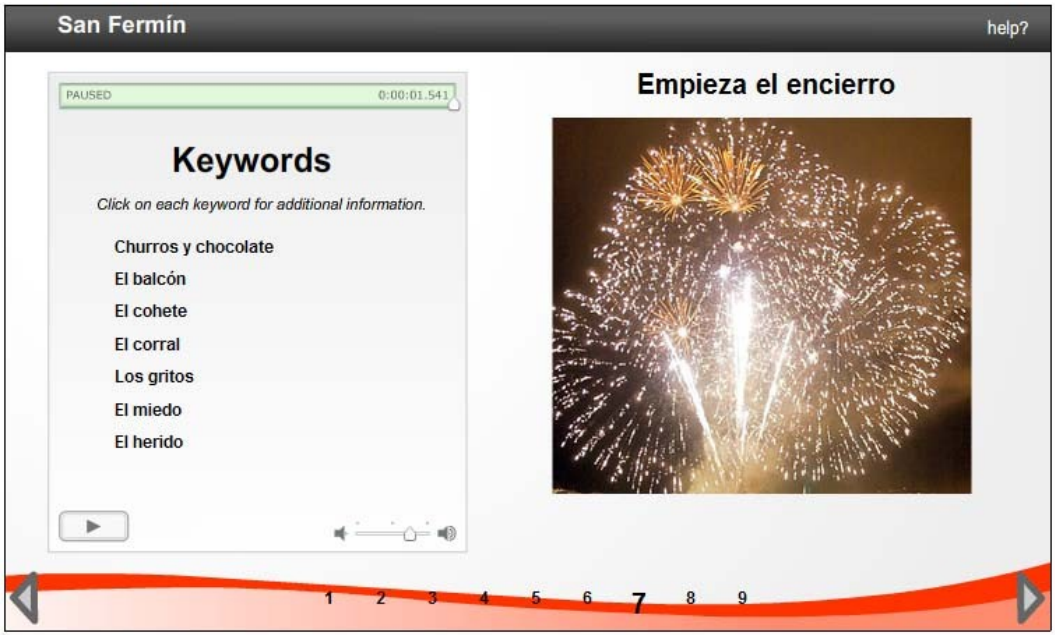
(Note: Script does not appear on screen at any time.)

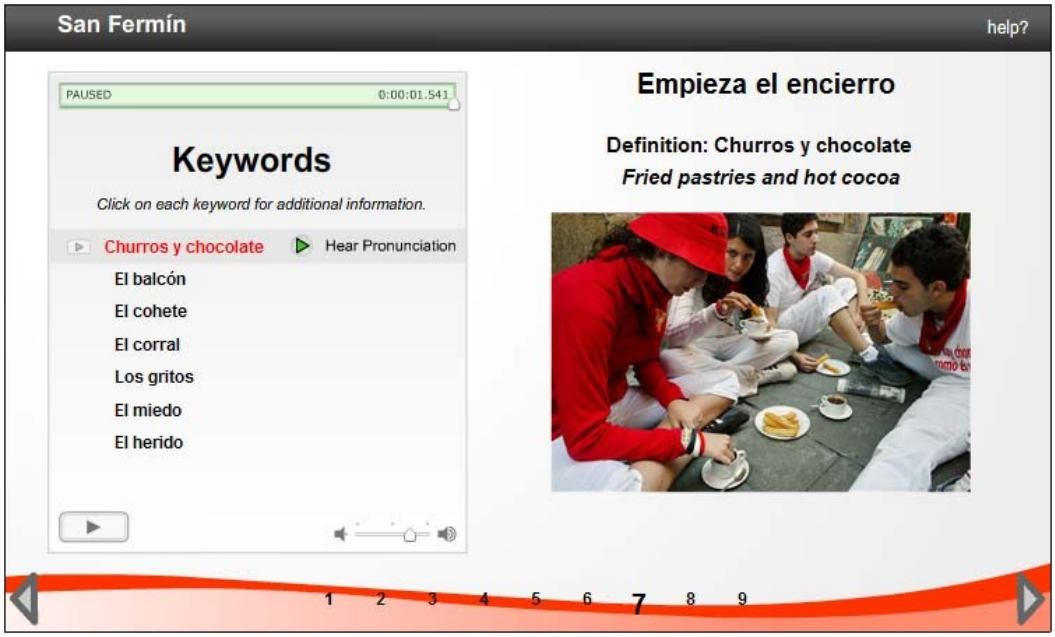
Empieza el encierro

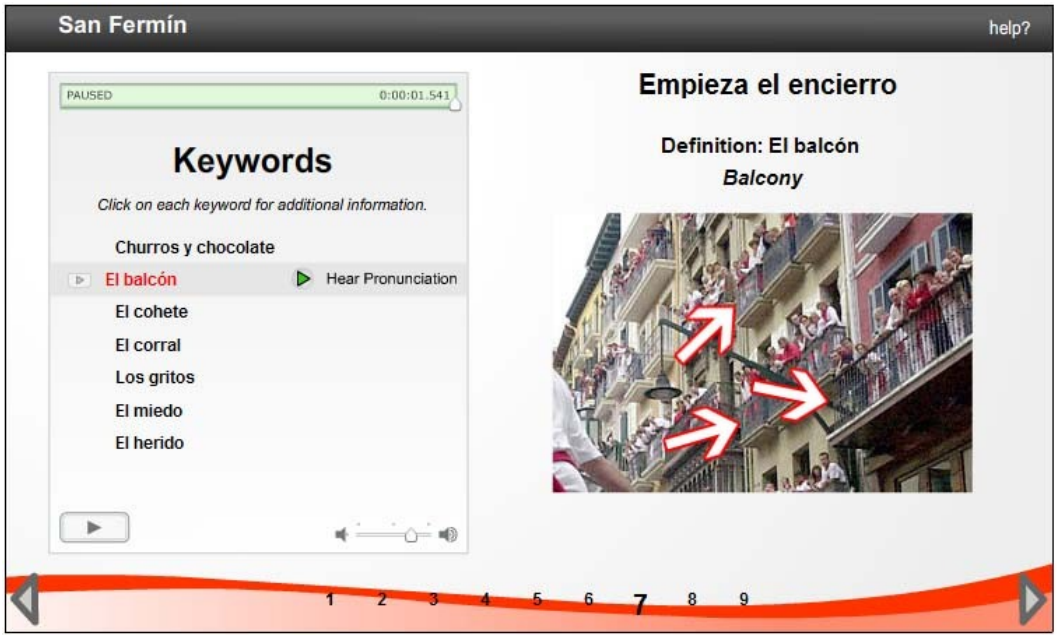
Temprano por la mañana muchos toman churros y chocolate, un desayuno típico durante las fiestas. Todos van hacia el recorrido del encierro para ver a los mozos y toros correr.
Se ponen detrás de vallas o en los balcones de las casas alrededor de la ruta.

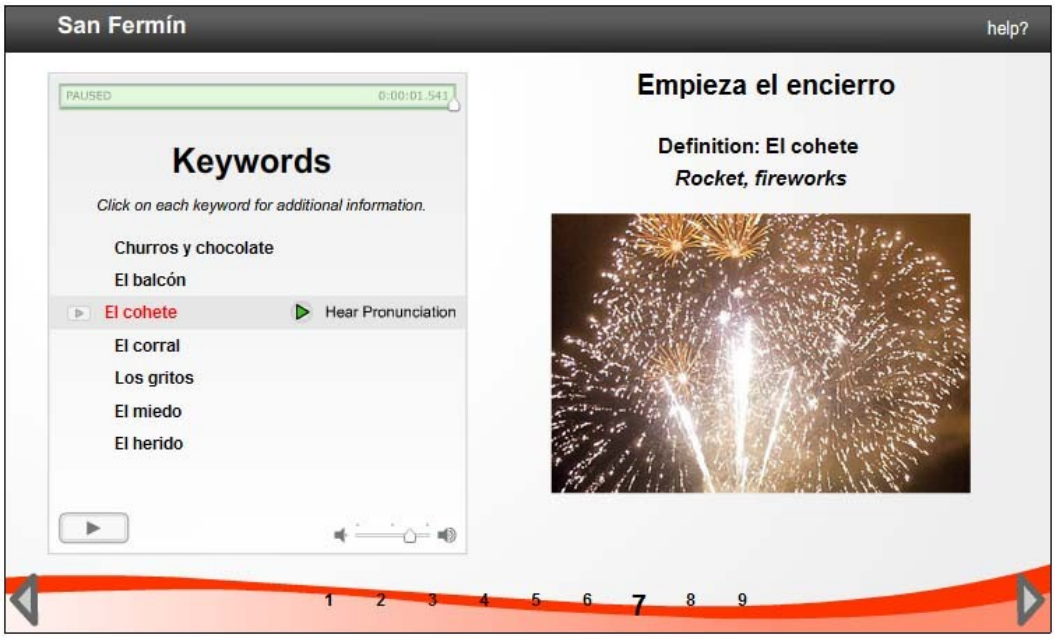
A las ocho lanzan un cohete, libran a los toros de los corrales y todos empiezan a correr. Se oyen los gritos del público de los balcones, el ruido de los toros y cientos de personas corriendo, muchos gritando en voz alta.

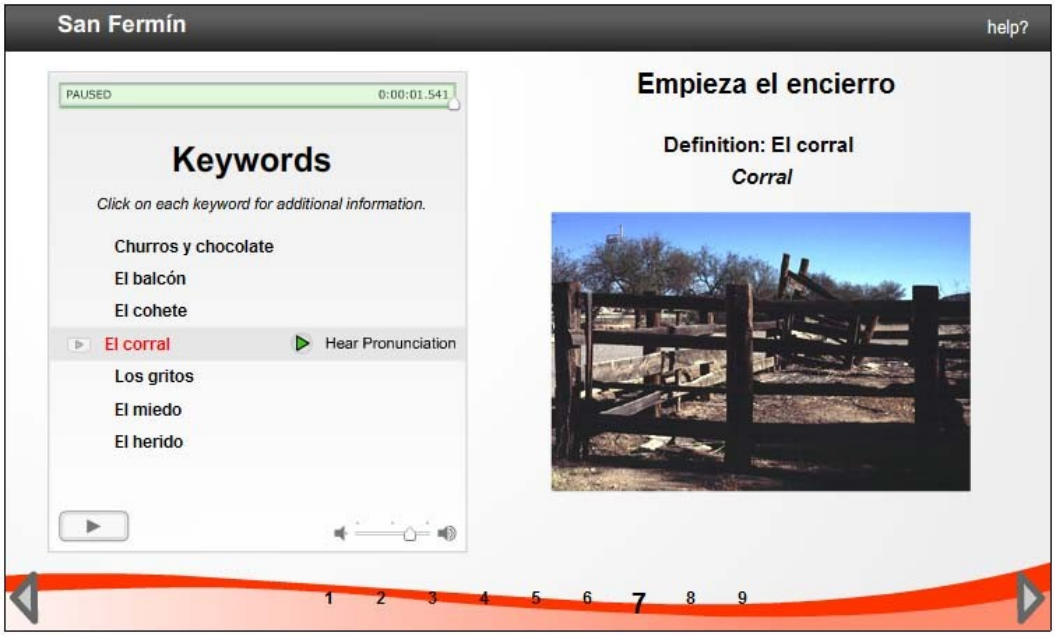
El miedo es evidente en las caras de los mozos pero por lo menos están corriendo y en pocos minutos se acaba. Todos esperan escapar del recorrido sin heridos.

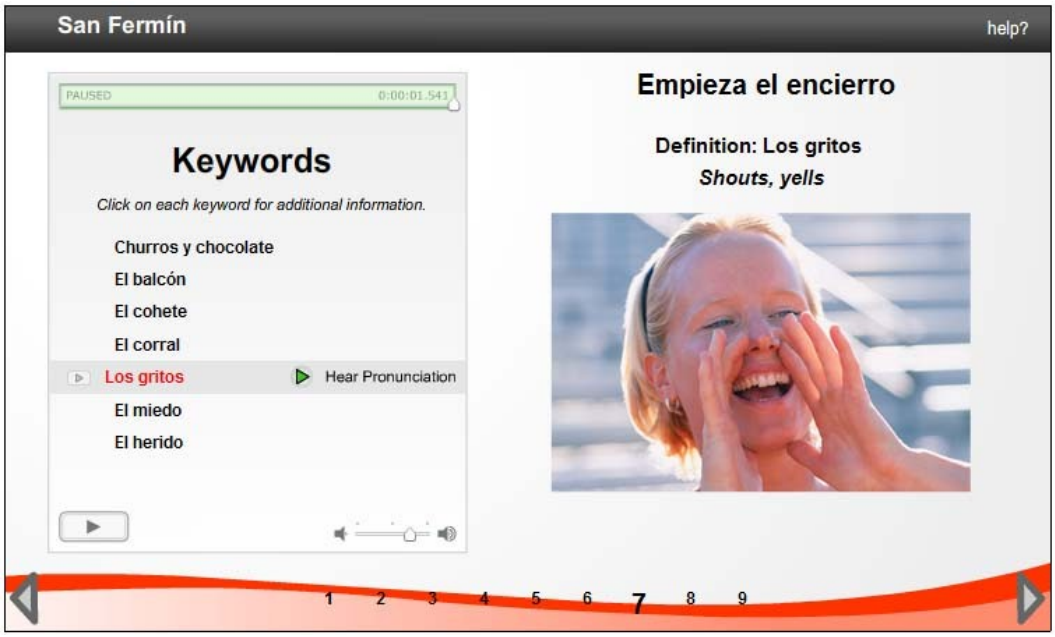

San Fermín
help?
PAUSED
0:00:01.541
Keywords
Click on each keyword for additional information.
Churros y chocolate
El balcón
El cohete
El corral
Los gritos
El miedo
El herido
Empieza el encierro
1 2 3 4 5 6 7 8 9

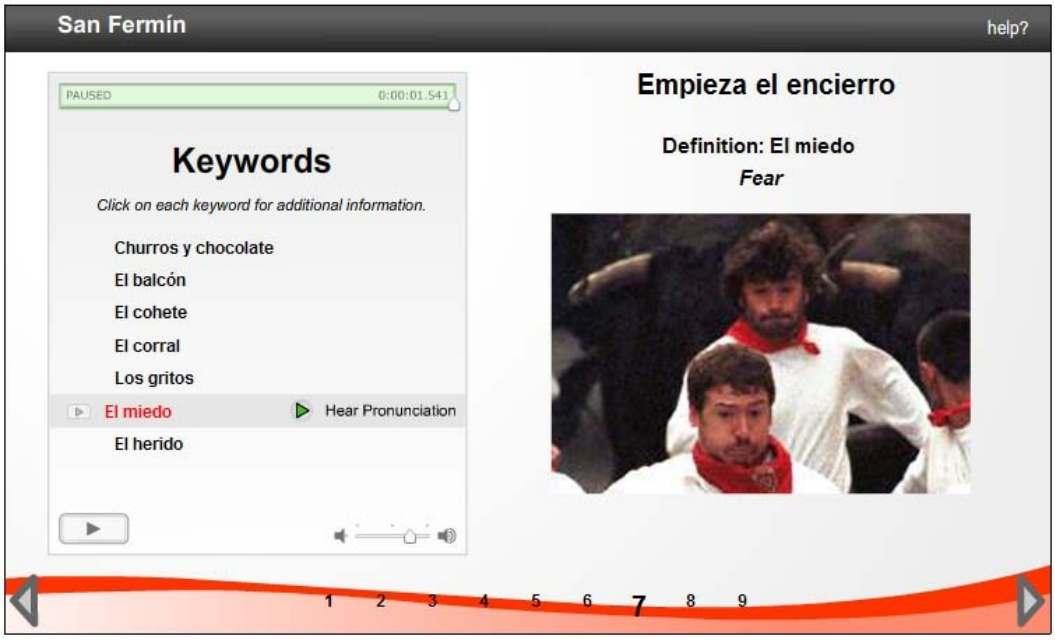

San Fermín
help?
PAUSED
0:00:01.541
Keywords
Click on each keyword for additional information.
Churros y chocolate
Hear Pronunciation
El balcón
El cohete
El corral
Los gritos
El miedo
El herido
Empieza el encierro
Definition: Churros y chocolate
Fried pastries and hot cocoa
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:01.541
Keywords
Click on each keyword for additional information.
Churros y chocolate
El balcón
Hear Pronunciation
El cohete
El corral
Los gritos
El miedo
El herido
Empieza el encierro
Definition: El balcón
Balcony
1 2 3 4 5 6 7 8 9

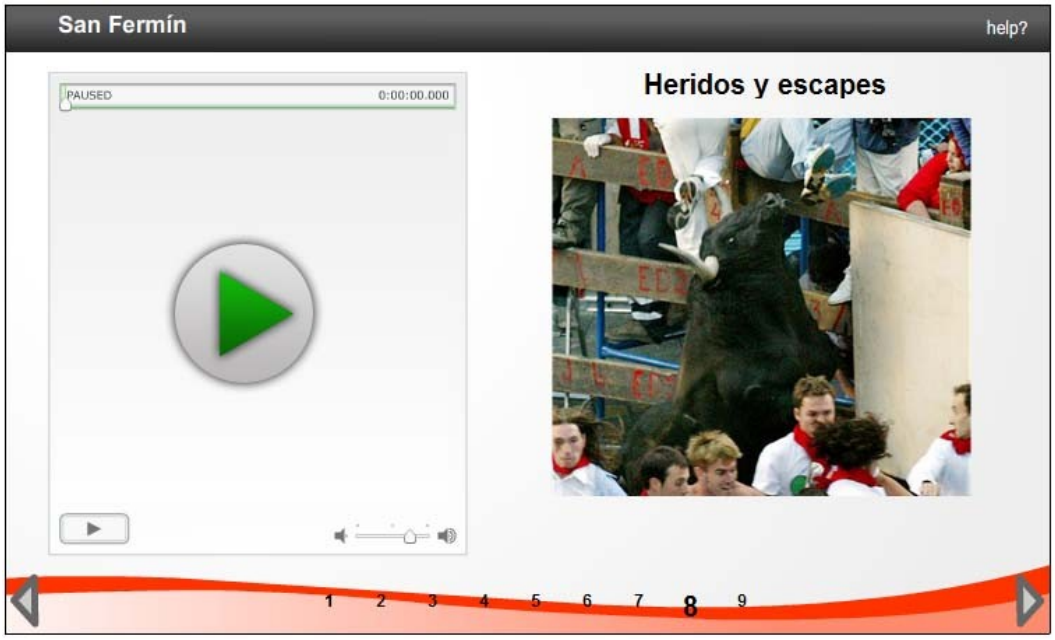

San Fermín
help?
PAUSED
0:00:01.541
Keywords
Click on each keyword for additional information.
Churros y chocolate
El balcón
El cohete
Hear Pronunciation
El corral
Los gritos
El miedo
El herido
Empieza el encierro
Definition: El cohete
Rocket, fireworks
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:01.541
Keywords
Click on each keyword for additional information.
Churros y chocolate
El balcón
El cohete
El corral
Hear Pronunciation
Los gritos
El miedo
El herido
Empieza el encierro
Definition: El corral
Corral
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:01.541
Keywords
Click on each keyword for additional information.
Churros y chocolate
El balcón
El cohete
El corral
Hear Pronunciation
Los gritos
El miedo
El herido
Empieza el encierro
Definition: Los gritos
Shouts, yells
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:01.541
Keywords
Click on each keyword for additional information.
Churros y chocolate
El balcón
El cohete
El corral
Los gritos
El miedo
Hear Pronunciation
El herido
Empieza el encierro
Definition: El miedo
Fear
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:01.541
Keywords
Click on each keyword for additional information.
Churros y chocolate
El balcón
El cohete
El corral
Los gritos
El miedo
El herido
Hear Pronunciation
Empieza el encierro
Definition: El herido
Injury
1 2 3 4 5 6 7 8 9



Audio script for Screen 8

(Note: Script does not appear on screen at any time.)

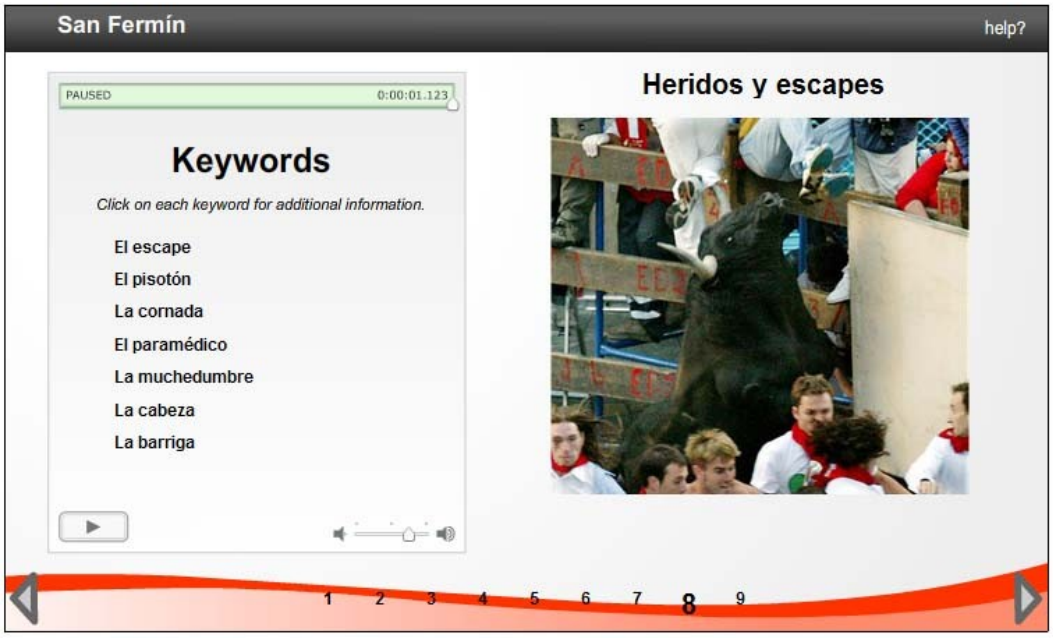
Heridos y escapes

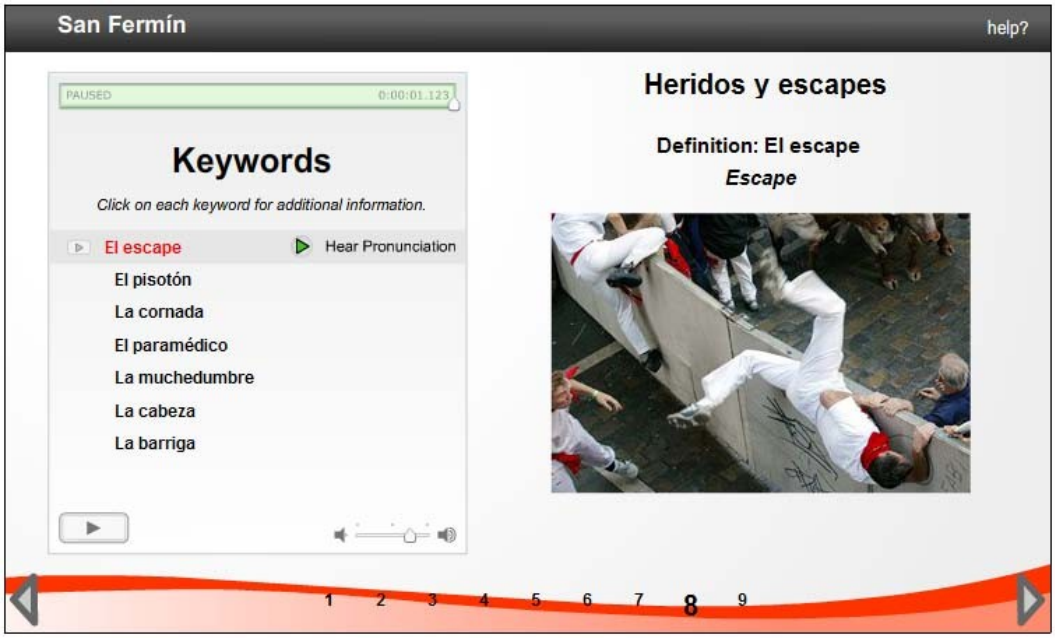
Cada año hay varios heridos. Pisotones y cornadas son comunes pero pocas personas mueren en el encierro. Siempre hay paramédicos cerca para cuidar a los heridos.

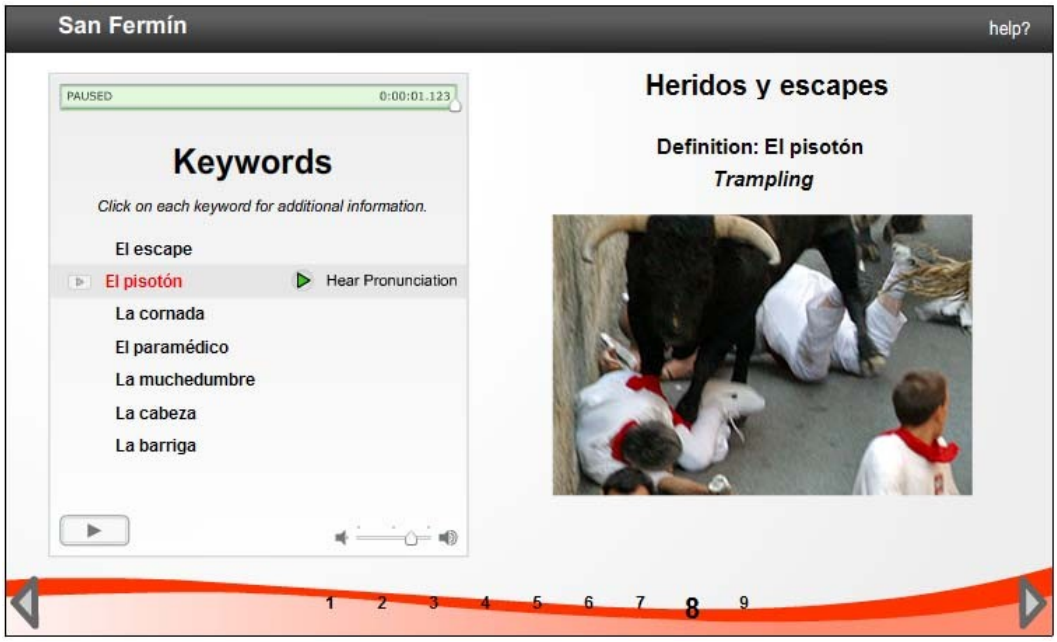
A veces un toro le da un pisotón a un mozo que se cae en la muchedumbre. Si un toro alcanza un mozo corriendo, a veces le da una cornada. Por eso cuando caen, los mozos protegen la cabeza y la barriga cuando pueden.

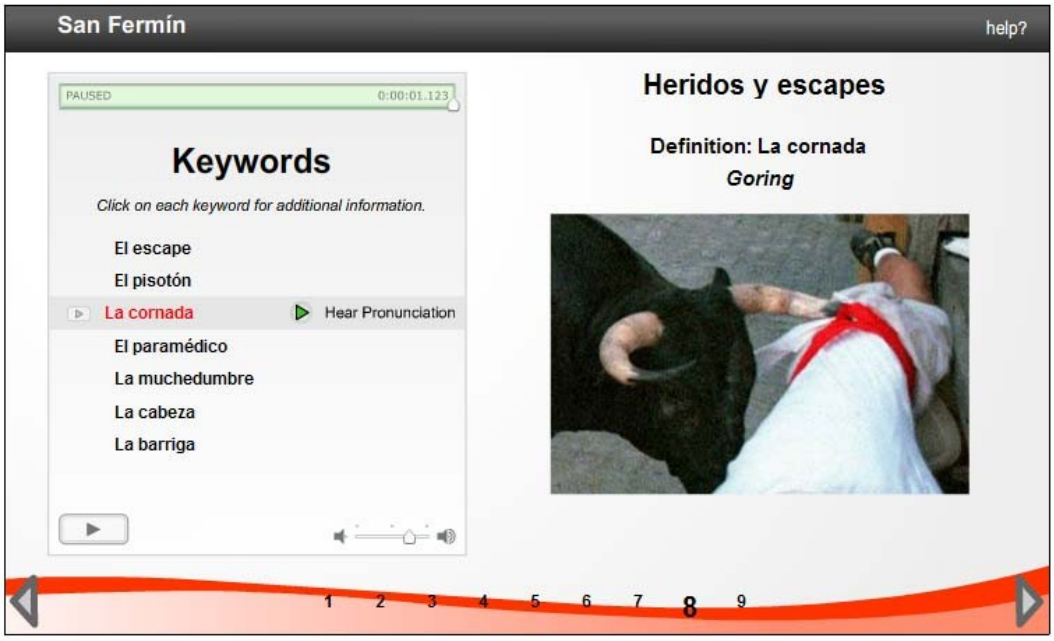
Casi todos se escapan sin heridos. Algunos se escapan encima o debajo de las vallas. Otros mozos escalan las paredes para llegar a un balcón donde los espectadores tratan de ayudar.

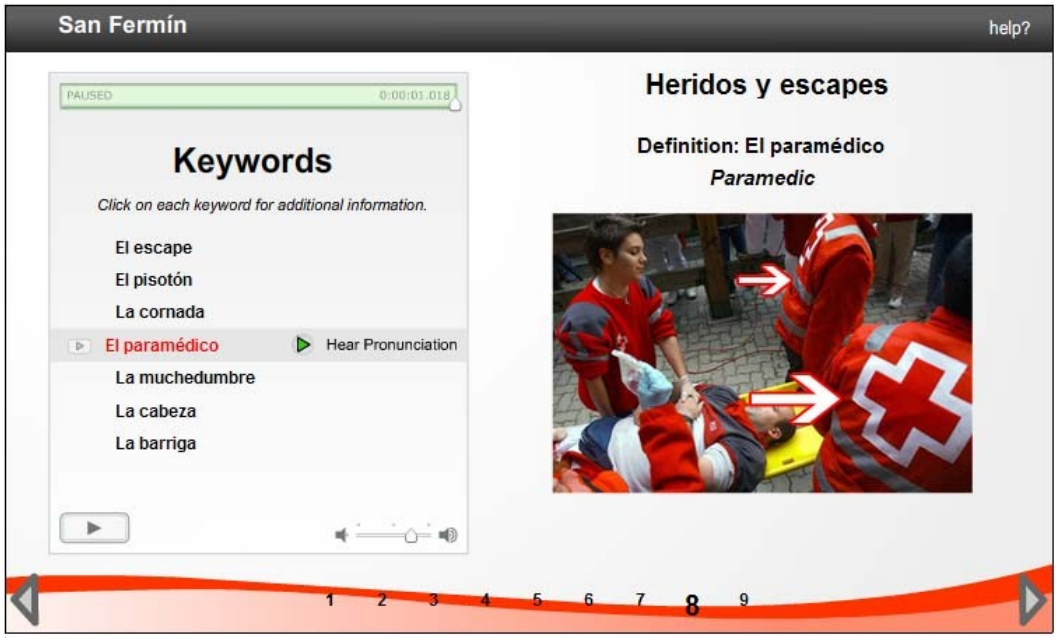
El encierro termina en dos o tres minutos cuando todos los toros están en la plaza y se oye otro cohete. Pero para los participantes y los espectadores es una experiencia inolvidable, lleno de riesgo, miedo, y al fin alivio al sobrevivir.

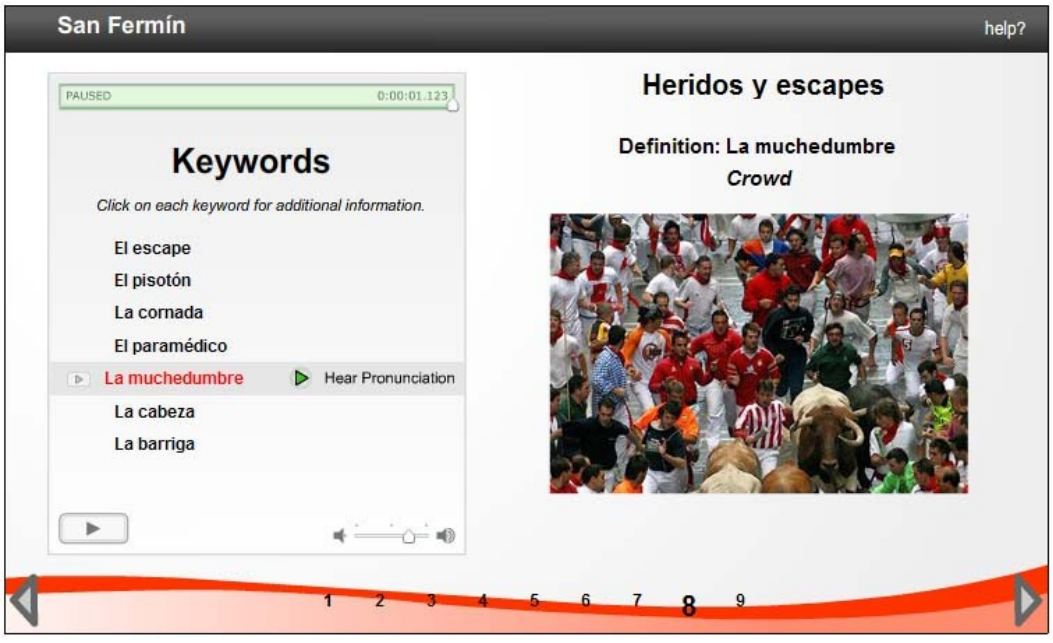

San Fermín
help?
PAUSED
0:00:01.123
Keywords
Click on each keyword for additional information.
El escape
El pisotón
La cornada
El paramédico
La muchedumbre
La cabeza
La barriga
Heridos y escapes
1 2 3 4 5 6 7 8 9

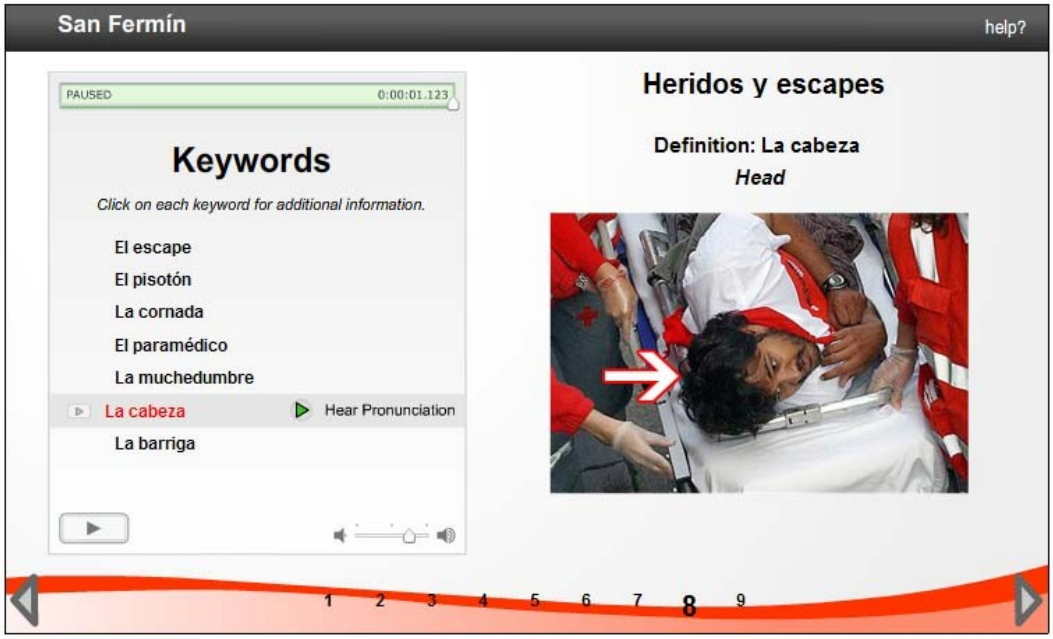

San Fermín
help?
PAUSED
0:00:01.123
Keywords
Click on each keyword for additional information.
El escape
Hear Pronunciation
El pisotón
La cornada
El paramédico
La muchedumbre
La cabeza
La barriga
Heridos y escapes
Definition: El escape
Escape
1 2 3 4 5 6 7 8 9

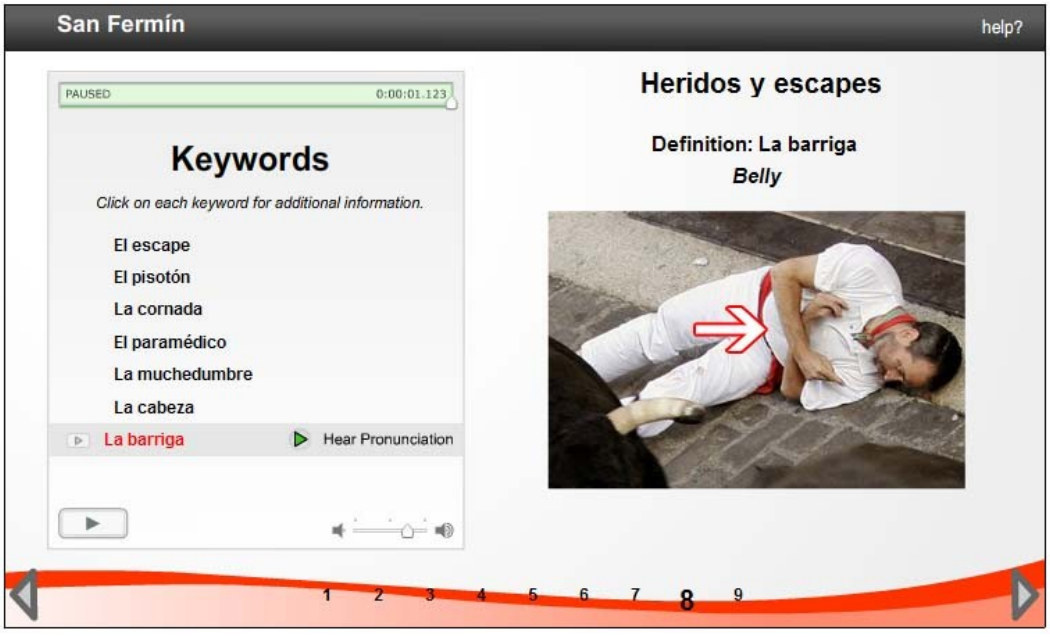

San Fermín
help?
PAUSED
0:00:01.123
Keywords
Click on each keyword for additional information.
El escape
El pisotón
Hear Pronunciation
La cornada
El paramédico
La muchedumbre
La cabeza
La barriga
Heridos y escapes
Definition: El pisotón
Trampling
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:01.123
Keywords
Click on each keyword for additional information.
El escape
El pisotón
La cornada
Hear Pronunciation
El paramédico
La muchedumbre
La cabeza
La barriga
Heridos y escapes
Definition: La cornada
Goring
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:01.018
Keywords
Click on each keyword for additional information.
El escape
El pisotón
La cornada
El paramédico
Hear Pronunciation
La muchedumbre
La cabeza
La barriga
Heridos y escapes
Definition: El paramédico
Paramedic
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:01.123
Keywords
Click on each keyword for additional information.
El escape
El pisotón
La cornada
El paramédico
La muchedumbre
Hear Pronunciation
La cabeza
La barriga
Heridos y escapes
Definition: La muchedumbre
Crowd
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:01.123
Keywords
Click on each keyword for additional information.
El escape
El pisotón
La cornada
El paramédico
La muchedumbre
La cabeza
Hear Pronunciation
La barriga
Heridos y escapes
Definition: La cabeza
Head
1 2 3 4 5 6 7 8 9


San Fermín
help?
PAUSED
0:00:01.123
Keywords
Click on each keyword for additional information.
El escape
El pisotón
La cornada
El paramédico
La muchedumbre
La cabeza
La barriga
Hear Pronunciation
Heridos y escapes
Definition: La barriga
Belly
1 2 3 4 5 6 7 8 9

San Fermín help?

What did you learn?

You have reached the end of the Running of the Bulls presentation. Until the time limit expires on the activity you may review any screen of the presentation before moving on to the posttest and survey.

Once you are ready, click on the submit button below the presentation to advance to the quiz and survey.

The quiz and survey will only take you 10-20 minutes at most to complete. Please spend sufficient time on the summary and each vocabulary question to give us your best answers and your honest feedback. Your input on this study may help to improve Spanish instruction at the college for hundreds of students each semester.

At the end of the survey you will have an opportunity to provide your name and address so that the researcher can send you the $10 study incentive in the mail. Your personal information will remain confidential and will be stripped from the study data before analysis.

Thank you for helping us improve Spanish classes at Rio Salado College!

Photos courtesy of AP Photo Archive & Jupiter Images

1 2 3 4 5 6 7 8 9

APPENDIX B

LIST OF KEYWORDS, TEXT DEFINITIONS AND ILLUSTRATIONS

Key Term in Spanish	English Text Definition	Illustration
El encierro	Running of the Bulls	
El espectáculo	Spectacle	
La carrera	Race	
Los mozos	Young men	
El toro	Bull	

La calle estrecha	Narrow street	
La ciudad	City	
El evento peligroso	Dangerous event	
La calle lateral	Side street	
El recorrido	Path, course	
La valla	Street barricade	

La madera	Wood	
La pared	Wall	
Los cuernos	Horns	
Los pantalones	Pants	
La camisa	Shirt	
La bufanda	Scarf	

El cinturón	Belt	
El espectador	Spectator	
El baile	Dance	
La hierba	Grass, lawn	
Churros y Chocolate	Churros and hot chocolate	
El balcón	Balcony	

El cohete	Rocket, fireworks	
El corral	Corral	
Los gritos	Shouts, yells	
El miedo	Fear	
El herido	Injury	
El escape	Escape	

El pisotón	Trampling	
La cornada	Goring	
El paramédico	Paramedic	
La muchedumbre	Crowd	
La cabeza	Head	
La barriga	Belly	

APPENDIX C

LISTENING COMPREHENSION POSTTEST

Listening Comprehension

Please complete all questions without referring to any outside resources. No dictionaries.

Your full effort and complete responses will help us to improve the quality of the listening activities that will be included in future course offerings.

You will first write a brief summary of what you learned in English, then answer a series of multiple choice vocabulary questions.

Following the quiz questions, you will complete a survey about the activity you have just completed. The survey asks about the effectiveness of the Sanfermines listening comprehension program. Please complete the survey in full!

Please summarize what you have learned about San Fermines and the Running of the Bulls. Include everything you can remember and **write in English**.

Type your response here.

APPENDIX D

VOCABULARY POSTTEST

1 of 25

Select the correct translation: **el herido**

- belt
- balcony
- injury
- horns

2 of 25

Select the correct translation: **el toro**

- bull
- horns
- scarf
- wood

3 of 25

Select the correct translation: **el balcón**

- wall
- city
- balcony
- barrier

4 of 25

Select the correct translation: **la barriga**

- pants
- head
- dance
- belly

5 of 25

Select the correct translation: **la muchedumbre**

- crowd
- young men
- wall
- fear

6 of 25

Select the correct translation: **el recorrido**

- course, path
- side street
- grass
- injury

7 of 25

Select the correct translation: **los pantalones**

- shirt
- pants
- wood
- trampling

8 of 25

Select the correct translation: **el evento peligroso**

- narrow street
- side street
- rockets, fireworks
- dangerous event

9 of 25

Select the correct translation: **la carrera**

- horns
- race
- young man
- city

10 of 25

Select the correct translation: **la pared**

- wall
- grass
- fear
- balcony

11 of 25

Select the correct translation: **los cuernos**

- cities
- rockets
- horns
- barriers

12 of 25

Select the correct translation: **la bufanda**

- belt
- trampling
- scarf
- dance

13 of 25

Select the correct translation: **la hierba**

- bull
- grass
- barrier
- injury

14 of 25

Select the correct translation: **los mozos**

- young men
- wood
- shirts
- walls

15 of 25

Select the correct translation: **las calles estrechas**

- side streets
- dangerous events
- cities
- narrow streets

16 of 25

Select the correct translation: **la madera**

- wood
- belt
- wall
- barrier

17 of 25

Select the correct translation: **el cinturón**

- shirt
- pants
- belt
- race

18 of 25

Select the correct translation: **el pisotón**

- trampling
- goring
- belt
- scarf

19 of 25

Select the correct translation: **el cohete**

- belt
- rocket, fireworks
- horn
- shirt

20 of 25

Select the correct translation: **el espectador**

- course, path
- fear
- spectator
- running of the bulls

21 of 25

Select the correct translation: **la ciudad**

- belt
- rocket, fireworks
- race
- city

22 of 25

Select the correct translation: **las vallas**

- pants
- barriers
- wood
- grass

23 of 25

Select the correct translation: **la camisa**

- race
- shirt
- fear
- horn

24 of 25

Select the correct translation: **el baile**

- dance
- barrier
- course, path
- goring

25 of 25

Select the correct translation: **el miedo**

- wood
- shirt
- fear
- race

APPENDIX E

COGNITIVE LOAD MEASURE AND ATTITUDE SURVEY

In the following five questions you will be asked to indicate how demanding you found the learning task you just completed with regard to each of five dimensions described in this section of the survey.

Please base your responses on the definitions stated for each of the five dimensions instead of using your own notion of what the dimensions are. All your data will be treated confidentially.

1. How much mental and physical activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching etc.)? That is, was the learning task easy (simple, forgiving) or demanding (exacting or unforgiving)?

Easy Demanding

2. How hard did you have to work in your attempt to understand the contents of the learning environment?

Not hard at all Very hard

3. How successful do you think you were in your attempt to understand the contents of the learning environment?

Not successful Very successful

4. How much effort did you have to invest in order to navigate the learning environment (e.g., for deciding between different hyperlinks, finding your way around)?

Low effort High effort

5. How stressed (insecure, discouraged, irritated, annoyed) did you feel during the learning task?

Not at all stressed Very stressed

The preceding five questions are adapted from the NASA-TLX survey instrument

(Hart & Staveland, 1988)

.

	Strongly Agree (5)	Agree (4)	Neither Agree nor Disagree (3)	Disagree (2)	Strongly Disagree (1)
The program was well designed and organized.	○	○	○	○	○
The topic of the program was relevant to my Spanish study.	○	○	○	○	○
The story was interesting to me.	○	○	○	○	○
Instructions within the program were clear and easy to follow.	○	○	○	○	○
Navigation within the program was easy to understand.	○	○	○	○	○
The listening activity helped me to learn new vocabulary.	○	○	○	○	○
The listening activity helped me to learn about the cultural topic.	○	○	○	○	○
I would like to have more listening activities of this type to help me understand spoken Spanish.	○	○	○	○	○

Please rate the effectiveness of the keyword (definitions, pronunciations, illustrations) you received in this listening activity.

Note: The translation and illustration questions below were only offered to participants in the corresponding treatments. The no-annotations group will see the first two questions only. The both-annotations group will see all six questions.

	Strongly Agree (5)	Agree (4)	Neither Agree nor Disagree (3)	Disagree (2)	Strongly Disagree (1)
Hearing the keywords pronounced alone helped me to LEARN the new words.	○	○	○	○	○
Hearing the keywords pronounced alone helped me to UNDERSTAND the story.	○	○	○	○	○
Reading the English translations of keywords helped me to LEARN the new words.	○	○	○	○	○
Reading the English translations of keywords helped me to UNDERSTAND the story.	○	○	○	○	○
Seeing the graphics illustrating keywords helped me to LEARN the new words.	○	○	○	○	○
Seeing the graphics illustrating keywords helped me to UNDERSTAND the story.	○	○	○	○	○

Note: The first question was different for each treatment group. For the no-annotations group the question only includes "pronunciation aids" while "definitions" and "illustrations" were included in corresponding treatments.

How could we make the vocabulary definitions, pronunciation aids and illustrations more effective for you?

What did you like best about the program?

What could be done to improve the program?

To receive the $10 incentive for full participation in this study, please enter your name and address below. Your personal information will remain confidential and will only be used for purposes of mailing the incentive.

Name:
Mailing Address:

Please indicate if you are a Rio Salado College or Arizona State University student by typing RSC or ASU below, and if you are in SPA101 or SPA102.

APPENDIX F

IRB APPROVAL FORM

Arizona State University
Office of Research Integrity and Assurance
PO Box 871103
Tempe, AZ 85287-1103
phone: (480) 965-6788
fax: (480) 965-7772
email: research.integrity@asu.edu

To:	Wilhelmina Savenye EDB
From:	Mark Roosa, Chair Soc Beh IRB
Date:	08/27/2009
Committee Action:	**Exemption Granted**
IRB Action Date:	08/27/2009
IRB Protocol #:	0907004179
Study Title:	The Effects of Visual and Textual Annotations on Spanish Listening Comprehension, Incidental Vocabulary Acquisition, and Cognitive Load

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.