2005

# Research into the use of help options in a multimedia listening unit

Maja Grgurović
*Iowa State University*

# Research into the use of help options in a multimedia listening unit

by

**Maja Grgurović**

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF ARTS

Major: Teaching English as a Second Language/Applied Linguistics
(Computer-Assisted Language Learning)

Program of Study Committee:
Volker Hegelheimer, Major Professor
Carol Chapelle
Denise Schmidt

Iowa State University

Ames, Iowa

2005

Graduate College
Iowa State University

This is to certify that the master's thesis of

Maja Grgurović

has met the thesis requirements of Iowa State University

Signatures have been redacted for privacy

*In the loving memory of my father Petar and my grandmother Nana*

# TABLE OF CONTENTS

# ABSTRACT

This study investigated learner behavior and performance in a CALL multimedia listening activity which offered a video and two textual help options--a lecture transcript and subtitles (L2 close captioning) in cases of comprehension breakdowns. In particular, the study examined and compared the learners' use of two help options by looking at the time and frequency of interaction with help in addition to student performance during and after the activity. The study aimed to identify, explain, and compare possible patterns of participants' interaction with help options as well as investigate possible differences between participants at two different proficiency levels. Finally, the study looked at participants' attitudes towards two help options before and after the activity.

The participants in the study were eighteen ESL college students at the intermediate level of listening proficiency who were enrolled in an academic listening class. The study took place over three weeks and was incorporated into the course curriculum. The materials given to participants included an initial listening test and a recall test, two questionnaires, and the CALL listening comprehension activity entitled *The Astronomy unit* which was the main data collection instrument. The Astronomy unit consisted of a video of an academic lecture on Astronomy divided into ten segments each followed by a comprehension question. The participants were offered help as a choice between the subtitles and the transcript when they could not answer the comprehension question correctly. Students' work on the activity was recorded using Camtasia screen capturing program and their interaction with the material was later transcribed. Selected participants were also interviewed after the activity.

The results showed that participants varied in their use of help options in terms of time on help, number of help page openings, and number of instances of useful interaction with help. Generally, the students interacted with the subtitles more frequently and for longer periods of time than with the transcript. Furthermore, the students exhibited four different patterns of interaction with help: subtitles, transcript, non-interaction, and mixed interaction pattern. While the subtitles and the transcript pattern groups showed very similar behavior, the non-interaction group differed from other groups the most. Some self-reported reasons behind four navigational patterns were personal preferences, previous help use, motivation,

and experimenting with help. The differences between the two proficiency groups (higher and lower) were found on their performance during and after the activity with the higher group having better comprehension. Additionally, the higher group spent more time on the subtitles while time and instances of useful interaction with the transcript remained very similar for both proficiency groups.

The findings obtained from the study showed that learners don't always take advantage of help features in CALL listening materials and that software design as well as teachers' use of software with students could promote interaction with help. The results suggested that in post-comprehension breakdowns, learners should be given both the subtitles and the transcript as well as a choice to skip help. Further research could look at a larger and varied learner sample in addition to the use of materials in different settings outside class to inform future software development and the most effective software use by the teachers and students alike.

# CHAPTER 1. INTRODUCTION

As multimedia language learning materials are becoming increasingly common in foreign and second language classrooms, the design of those materials is becoming an important avenue of research in computer-assisted language learning (CALL). Chapelle (2003) argues that the real challenge in applied linguistics at present is not a comparison of CALL and the classroom but a search for "evidence for the most effective ways to design software for CALL, to use software effectively in tasks, and to help learners to take advantage of the electronic resources available to them" (p. xiii). This challenge is, I believe, as tempting for CALL software designers and publishers, CALL researchers and practitioners, as for teachers and learners given the present rate of technological advancement of our world. This work is in part a small contribution to embracing the challenge since it tries to add to the understanding of learner use of multimedia listening software. In particular, it focuses on learner interaction with software help features by examining learner behavior and performance.

## Purpose of the study

This study investigates a multimedia listening activity that employs a video of an academic lecture and two help options--subtitles and the transcript. The subtitles in the study are in form of target language close captions that are synchronized with the video while the transcript is the text of the lecture. Both help options are given to participants after their comprehension of a lecture segment failed. The aim of the study is to examine how participants, ESL college students, use the two help options by looking at frequency and time of interaction. The study also examines possible patterns of interaction with help as well as differences in behavior and performance of participants at different proficiency levels. Comparisons between the use of two help options, groups following different navigational patterns, and groups at two proficiency levels are presented in addition to students' attitudes towards help.

## Rationale of the study

My interest for this research comes from my teaching of English 101 Listening, Academic Listening for non-native speakers of English, at Iowa State University. The goal of the course is to help students improve listening strategies and vocabulary necessary for them to function in their everyday academic life. ESL students in the class found some listening activities difficult so I tried to use different ways to assist them. For example, when listening in class, I provided students with the transcript of the lecture they were listening to, which they found very helpful. In addition to watching a video in class, they sometimes listened to on-line video lectures for homework assignments. Students' positive reactions to these activities motivated me to create a CD-ROM-based multimedia unit with a video of a lecture on Astronomy and two textual help options. I wanted to give students additional listening practice through the exposure to an academic lecture and comprehension exercises which are similar to the type of tasks they encounter outside ESL classes.

This research focuses on help options which I consider an essential part of multimedia materials. Indeed, help options in CALL present learners with opportunities for modified interaction (Chapelle, 2001) which can potentially create negotiation of meaning crucial for language development according to the interactionist perspective on second language acquisition (SLA). Furthermore, the research has provided evidence of the language gains by participants who made use of help. For example, it was found that both subtitles (Borras & Lafayette, 1994; Garza, 1991; Vanderplank, 1988, 1990) and transcripts (Hsu, 1994; Liou, 1997, 2000) can have beneficial effects on students' listening comprehension. However, none of these studies closely compared these two textual help options, which although having the same purpose (Pujola, 2002, p. 251), may not be used in the same manner. In addition, none of the studies investigated the behavior of participants in a controlled multimedia environment which offers help only in cases of comprehension breakdowns. These issues created additional motivation for the pursuit of this study.

## Research questions

Although the literature has offered some insights into the effectiveness of video and subtitles for listening comprehension as well as the design of help options and patterns of user interaction with multimedia, my study attempts to shed light on these issues by looking at a single multimedia listening unit and learners' interaction with it. Specifically, the study hopes to add more understanding of the participants' use of subtitles and the transcript by comparing them. Additionally, this work attempts to determine, describe, and compare possible patterns of user interaction with help options. Finally, the study examines whether proficiency level has an effect on behavior and performance of participants. In order to investigate the use of help in multimedia listening materials, the following four research questions will be addressed:

1. How frequently and for how long do participants use two help options (subtitles and the transcript) offered in a multimedia activity? How do these two help options compare in terms of frequency and time of use?
2. Are there patterns of participants' behavior and how could those patterns be explained? If there are patterns, how do groups of participants following those patterns compare in terms of comprehension, and time and frequency of help use?
3. Do participants at different proficiency levels vary in their use of help options?
4. What are participants' attitudes towards two help options before and after the activity?

## Organization of the study

The next chapter, Chapter 2, provides an overview of theoretical and practical issues involved in multimedia CALL materials especially as they pertain to listening comprehension and software help options. Chapter 3 covers materials and methods used in the study and includes the participants, data collection instruments, and methods of data analysis. Chapter 4 presents results of the study by the research question and accounts for students' use of help as well as differences between four types of navigational behavior. Finally, Chapter 5 summarizes the major findings, and addresses the limitation and

implication of results to teaching listening using CALL software and multimedia software design.

# CHAPTER 2. LITERATURE REVIEW

This chapter provides a literature review of the areas related to help options in multimedia listening programs. It opens with an overview of theoretical underpinnings relevant for the development of multimedia materials. The interactionist perspective on second language acquisition (Pica, 1994) is reviewed in connection to some of the practical considerations of CALL materials design (Chapelle, 1998). The third section of the chapter deals with the advantages and use of multimedia with a special focus on multimedia listening materials. Since the multimedia unit used for data collection in this study uses a video clip and subtitles, previous research into these areas is presented as well. In the next section, studies dealing with help functions in multimedia listening are discussed since many of them proved instrumental for this research. Finally, some of the research issues into patterns of user interaction with multimedia materials are presented and the lack of research into patterns of user interaction with help options highlighted.

## The interactionist theory and the negotiated interaction model in SLA

The theoretical base for this research came from the interactionist perspective on second language acquisition (SLA). This perspective considers interaction to be a crucial factor for language acquisition since it can promote negotiation of meaning (Long, 1996; Pica, 1994). The interaction hypothesis was originally based on the study of the negotiation of meaning between human interlocutors, but the notion of interaction in CALL can be extended to include "what takes place between a person and the computer" (Chapelle, 2003, p. 55).

Chapelle (1998, p. 23) presented the interactionist model of SLA (based on Gass, 1997) which could serve as a theoretical base for the practical design of multimedia CALL. This model was built on Krashen's (1982) theory of comprehensible input as well as noticing of salient linguistic features in the input (Schmidt, 1990). According to the model, not everything that learners hear/read will necessarily help their language development. Only the input that is noticed, or apperceived, can become beneficial and that is why instructional materials should contain features that enhance input through modifications and elaborations

(Chapelle, 2003, p. 40). The next step in the model represents comprehension of the input which, after being semantically and syntactically processed, can become intake. Finally, the intake has the potential to be integrated into the learner's linguistic system whose development can be monitored through the output, or linguistic production.

## The interactionist model in SLA and multimedia CALL development

Following the interactionist model of SLA outline above and work of Long (1996) and Pica (1994), Chapelle (1998) looked at hypotheses about ideal conditions for SLA in terms of hypotheses application to CALL. Chapelle's seven suggestions relevant for the development of multimedia CALL are as follows:

1. Salience of target language input
2. Help in input comprehension
3. Opportunities for language output
4. Error noticing in output
5. Correction of linguistic output
6. Negotiation of meaning
7. Engagement in language tasks

In proposing those, Chapelle considered CALL software as a participant in the language task. It is believed that CALL software can create conditions that facilitate second language acquisition through its input, interaction, and feedback to the learner.

This study is most interested in the comprehension part of the interactionist model as well as Chapelle's second (p. 24) and seventh suggestion (p. 27). The second suggestion proposes that learners "should receive help in comprehending semantic and syntactic aspects" of the input (p. 24). This help comes through a form of modification of the input mode (p. 27) since, in case of the listening unit used in the study, the audio form of the input delivered through the video changes into textual form in the transcript and subtitles. Suggestion seven advocates learner engagement in the task so that opportunities for interaction can be maximized (p. 25). Making allowances for interaction in a multimedia CALL program for listening through the design of help options (like subtitles and transcript)

is the first step in prompting learner interaction with the material. Whether learners use opportunities to engage in that interaction is the issue of investigation in this and many other CALL studies.

Although SLA theory argues that the comprehensible oral input (which can reach learners through listening) is a starting point for the process of the acquisition of the second language, it does not specify the exact input delivery forms in instructional materials that could be beneficial for SLA. Similarly, although Chapelle's (1998) suggestions for multimedia CALL development call for help in input comprehension, more research is necessary to determine which type of textual help given to listeners in multimedia environments is valuable for SLA. While previous studies have found mixed results on the correlation between the use of transcripts and the improvement in listening comprehension (Hsu, 1994; Liou, 1997), this study will try to offer more insight into beneficial effects of both the transcript and the subtitles.

## Multimedia and listening

Multimedia software is "a program that integrates text, graphics, images, video and audio into only one media" (Keng-Soon, 1999, p. 299). According to Brett (1995), the advantages of multimedia for language learning are its huge data storing capacities, excellent quality of content data that does not deteriorate with time, high degree of learner control, and quick and easy access to materials. All media combined into one may satisfy many of the components of the negotiated interaction model. For example, authentic digital video can be used a source of input which learners can control using functions such as pause and rewind. Subtitles in the video can help with the decoding of the aural message and add textual redundancy to the audio and visual message. Furthermore, immediate feedback can help check and confirm the understanding of the message. Other help options in multimedia learning materials such as transcripts, glossaries, and dictionaries can additionally aid comprehension of the message by facilitating learner interaction with the material and consequently the negotiation of the input.

A literature review of studies that employed multimedia software revealed that a lot has been written about different characteristics of multimedia which can enhance reading,

vocabulary acquisition, and even speaking. For example, using college students learning German, Chun and Plass (1997) investigated whether video advanced organizers and different media annotations had an effect on reading comprehension and vocabulary acquisition. Their results showed that integration of visual and verbal elements in multiple media can facilitate comprehension and acquisition. Similarly, Plass, Chun, Mayer & Leutner (1998) found that the use of multimedia for visual and verbal annotations of vocabulary words assisted vocabulary acquisition. Along the same lines, Borras and Lafayette (1994) concluded that multiple sources of input (authentic video in addition to subtitles) can increase performance on oral production tasks. However, when in comes to investigation of multimedia listening materials most of the work investigated learner attitudes towards multimedia as well as compared multimedia with traditional media.

To investigate learner attitudes towards multimedia as an independent learning tool, Brett (1996) surveyed 107 undergraduate EFL learners, enrolled at a university in the UK, who used a multimedia program for developing listening skills *English for Business* for the first time during a 45-minute period. More than 80% of the students believed that multimedia could improve their listening skills and that it would be beneficial for their language learning. Multimedia was also preferred to audio, video, and books as students reported at the end of the multimedia session. Brett's (2000) study about the integration of multimedia into a self-study curriculum yielded somewhat similar results. This time, 64 undergraduate learners of business English were instructed to use the software over a much longer period of 8 weeks for an hour a week and their attitudes towards the program were collected after the first and the last session. More than half of the participants reported improved listening skills although, after the last session, the general attitudes towards multimedia were less positive than at the beginning.

In addition to learners' attitudes, Brett (1997) also looked into learner performance while engaging with multimedia listening software. He used comprehension and language recall tests in addition to student questionnaires to investigate the effectiveness of a program from the same *English for Business* series as above. Forty three advanced EFL undergraduate students of business were exposed to six different listening texts delivered on different media: CDROM, audio, and video tape. Texts delivered on cassettes were accompanied by

tests on paper, while for multimedia texts tests appeared on the screen and had immediate feedback. Students scored better on 4 out of 6 comprehension and recall tests delivered on multimedia in comparison to audio and video delivery. The questionnaire results showed that students found the combination of multimedia learning features (audio, video, and text) very helpful and that they preferred multimedia over audio and video for listening, learning, and understanding. Overall, the work of Brett (1995, 1996, 1997, 2000) showed positive attitudes of learners towards multimedia listening materials although more research is needed into learner short and long term language gains. Finally, more studies investigating different possibilities that multimedia listening materials offer are necessary. The importance of incorporating video into such materials is discussed next.

## Video

Video represents an important part of multimedia materials and as storage capacities of CDs, DVDs, and hard drives increase together with network bandwidth it will be hard to imagine digital language learning materials without video in future. In its analog format, video has been a part of the language classroom, aid in teaching listening, and the subject of investigation for many years (Herron, Secules, Morris, & Curtis, 1995; Rubin, 1990; Secules, Herron, & Tomasello, 1992). Secules et al. (1992), for example, investigated the performance of fifty two college students of French whose curriculum was built around eighteen half-hour French video tapes that they watched in class over one semester. When those students were given a listening comprehension test at the end of the semester, they scored higher than students who followed a traditional non-video curriculum. These results are consistent with those of Rubin (1990) and Herron et al. (1995) who also found support that video can facilitate comprehension. Video is believed to help listening comprehension with its sufficient context that can assist with the understanding of the message, visual support, exposure, and motivation that it brings to the classroom (Rubin, 1995). Finally, the fact that video is the medium learners are predisposed to (Meskill, 1996; Rubin, 1995) adds to its advantages for language instruction.

At present, digital video as a part of multimedia has given learners more control than ever before. Some of the advantages include learner control over the viewing and the

reviewing process (Guillory, 1999), and learner control of the selection of other media in addition to video (Brett, 1995). This control can help the learner modify the input allowing for beneficial focus on form and added redundancy (Chapelle, 1998, p. 2001) which can potentially promote language acquisition.

## Subtitles

Close captioning or same language subtitles first appeared in television broadcasts in the USA in 1980 primarily as a service to hearing impaired population who could, from that time, follow TV programs through the printed version of the spoken text (Garza, 1991). However, the work on subtitles had started in the early 1960's during the debate whether TV programs in other languages should be dubbed or subtitled (Vanderplank, 1990). Unlike the USA, where foreign programs were always subtitled, this was an important issue in Europe, where some counties like France and Germany dubbed their programs at that time. This argument sparked interest into the use of subtitles in language teaching and ever since their influence on second and foreign language learning has been a subject of investigation (Vanderplank, 1988). In particular, researchers looked into the influence of intralingual subtitles (target language input-target language subtitles) on listening, reading and speaking as well as their effectiveness for students at different proficiency levels.

In this area, the work of Vanderplank (1988, 1990) on subtitling of TV programs represented a base for further investigation of the effectiveness of video and digital video subtitles. Vanderplank (1988) studied the effect of English subtitles in nine hours of BBC program on 15 European intermediate to advanced learners of English at a university in Scotland. The participants viewed the subtitled program recorded on video tapes for an hour each week and afterwards commented that they felt subtitles were useful and beneficial for their language development. In the post-viewing activities, a high level of retention and recall was reported which Vanderplank attributed to the development of participants' "chunking ability in both reading and listening" (p. 277). In his 1990 study, Vanderplank continued the same investigation this time with 18-25 volunteer students (15 of which attended regularly) who watched captioned BBC program for a longer period of three months. These students were thought to be on the same proficiency level as those in 1988 study although no language

tests were administered. The results supported the previous findings about learner gains but only for participants who paid conscious attention to the language used in programs by taking down notes and trying to retain information in other ways (p. 226).

At the same time Garza (1991) also investigated the use of subtitles in video recordings of TV programs but using speakers of different native languages--English and Russian. The participants were advanced learners of ESL (70 participants) and Russian as a foreign language (40 participants) who watched ten video segments in English and Russian, respectively. The students were divided into caption and no-caption groups and answered comprehension questions after each video segment. Comprehension tests, in addition to recall interviews, revealed that subtitles significantly increased comprehension and recall of content. All caption groups had 75.2% more correct answers than no-caption groups as well as 61.16% fewer incorrect answers.

While subtitles exhibited promising effects for listening comprehension in the work of Vanderplank (1988, 1990) and Garza (1991), Borras and Lafayette (1994) researched the influence of subtitles on students' speaking performance. The participants were 44 advanced learners of French who were randomly divided into 4 treatments that included the exposure to subtitled and non-subtitled digital video which was a part of a multimedia unit. The results revealed significantly higher overall performance scores of participants in the subtitles groups and "highly positive attitudes towards speaking practice with multimedia courseware" (p. 68). Borras and Lafayette concluded that learners who are given control of subtitles in an authentic video exhibit better comprehension and better production of the language. Findings by Vanderplank (1988, 1990), Garza (1991), and Borras and Lafayette (1994) showed that subtitles may enhance listening, reading and speaking competencies although the researchers believed that intermediate and advanced students may benefit more from subtitles than beginner students.

To investigate the influence of different types of subtitles on lower level students, Guillory (1999) used digital video clips to check the comprehension of 202 beginner students of French as a foreign language who were randomly assigned into three groups: no captions, full captions, and keyword caption groups. The study showed that the full caption group outperformed the keyword caption group although both caption conditions proved beneficial

for students' comprehension. In questionnaires, the participants commented on difficulties in understanding authentic video materials in spite of captions and expressed their preference for keyword to full text captions. It seems that participants also had processing difficulties since 26% of full caption students and 20% of keyword caption students reported difficulties in attending to audio and visual message at the same time. This finding adds a new perspective on subtitles as language learning resources and calls for more research with different proficiency levels, an issue my work is going to investigate. In more recent studies, subtitles have been investigated in connection to digital video and have been added to multimedia software as a part of help options which will be discussed in the next section.

## Help options in multimedia listening

According to Pujola (2002), help facilities are "resources of the program which assist the learner in performing a task" (p. 241). A number of studies have dealt with help options that are offered to learners while engaging in multimedia listening tasks (Hegelheimer & Tower, 2004; Hsu, 1994; Liou, 1997, 2000; Pujola, 2002). The studies used multimedia products available on the market or custom-made software (Pujola, 2002) to investigate the use of different types of help, patterns of user behavior, and effectiveness of help. Since these studies were very important for the design of materials used in this thesis and general set up of the study, they will be looked at in detail while the special focus will be given to textual help options-transcript and subtitles.

Hsu's study (1994) was one among the first core studies examining ESL students' use of help options (audio repetition, textual repetition, and dictionary) while listening to a story delivered on the computer. Textual repetition was in the form of the transcript which was displayed on the screen when students requested help after being unable to comprehend the spoken input. Frequencies of help function use showed that the transcript was the most often used tool followed by the aural repetition and the dictionary. The transcript was also the most effective tool as participants reported in questionnaires. Additionally, Hsu described participants' patterns of interaction with help functions to find out that the higher and lower group did not follow the same pattern. The lower group followed the audio repetition-transcript-dictionary pattern while the higher group followed the transcript-dictionary pattern

when requesting help during the activity. Hsu, however, did not attempt to explain the reasons behind this type of participants' behavior. Her study did show that the amount of requests for textual help positively correlated with the participants' listening comprehension scores. A summary of main findings in this study and others reviewed in the section can be found in Table 2.1.

The frequency and effectiveness of help use was also looked at by Liou (1997). She examined how 20 college students at a Taiwanese university interacted with eight on-line help functions of a self-paced multimedia video disc. The textual help options included an English and Chinese script in addition to the video control functions (pause and rewind) and other options presented in Table 2.1. Participants were previously divided into an effective and an ineffective group based on their listening abilities. Frequency of help access showed that the ineffective group requested twice as much help as the effective group and used the replay of aural input more than the English and Chinese script. The effective group, on the other hand, used the English script most, followed by the replay and the Chinese script. As can be seen from Hsu's and Liou's results, the transcript was among the most frequently used forms of help. Contrary to Hsu's results, in Liou's study, frequency of help use did not correlate with listening scores. In another article about the same study Liou (2000) reported on the usefulness of these help functions as perceived by participants. While 80% of participants found the English script useful, the Chinese script was useful for only 55% of the participants.

Pujola (2002) examined the use of subtitles (in addition to the transcript) which were a part of one help option because they share the same purpose (p. 251). In Pujola's multimedia program, there were six other help facilities apart from the subtitles and the transcript (the complete list can be found in Table 2.1). Based on participants' decoding level, Pujola divided 22 beginner EFL students into four groups (higher, average, lower, and poorer decoders) and observed their use of textual help. He found that participants in each group behaved in "varied, idiosyncratic ways" (p. 253) so it was difficult to draw conclusions that would apply to all participants in one group especially since some participants in lower groups never used textual help. Generally, the higher decoders used the reply and rewind functions more than the transcript and/or subtitles. The use of textual help increased as the

decoding level decreased, so some poorer decoders relied extensively on it "sometimes as they were doing a reading task instead of listening" (p. 253). Pujola believed that the explanation can partly be found in the participants' perception of the purpose behind textual help. While the higher groups used textual help as a backup for listening, lower groups tended to use it as a necessary part of the listening process. In terms of overall help use, there was no correlation between the use of help and participants' linguistic level. This shows that the proficiency level as well as comprehension scores may not be the best predictors of help function use.

Another study that looked at the use of textual help in multimedia software was that by Hegelheimer and Tower (2004). The participants were 94 beginner EFL students at a university in the Middle East who used a CALL program for self-study over the period of 8 weeks. Since using the software was a mandatory part of the course, teachers required participants to use some help options regularly while working with the program at their own pace. For that reason, the Repeat button, which offered repeated aural input, was introduced to participants, unlike the ABC button which enabled the participants to repeat the aural input simultaneously with the transcript. Perhaps, that is why, overall, the Repeat button was used more than the ABC button. The results showed that the simultaneous display of audiovisual and textual input was negatively correlated with comprehension scores. However, the ABC button was the most frequently used help option for the 30 lowest-performing participants. Thirty highest-performing participants used the Repeat button most frequently. As in Pujola's study (2002), the participants showed great variation when using help options. For example, half of the participants did not use the ABC button at all. More research is necessary to explain why this is the case. Liou (1997) hinted at this issue when wondering if it were the students who did not use the environment to their advantage or CALL developers who implement more help than necessary (p. 94). Whatever might be the case, Hegelheimer and Tower (2004) believe that teachers can help by demonstrating help option use while creating tasks that encourage the use of help (p. 201).

Table 2.1 Summary of studies dealing with help options in multimedia listening materials

| Author year | Title | Participants | Materials | Help options used | Frequency of help option use | Patterns described | Usefulness (self-reported) |
|---|---|---|---|---|---|---|---|
| Hsu, 1994 | Computer assisted language learning (CALL): the effect of ESL students' use of interactional modifications on listening comprehension. | 15 beginner ESL students divided into higher and lower group | Interactive Active English CD-ROM program | Three input modifications displayed from the beginning<br>1. aural repetition<br>2. text reinforcement<br>3. dictionary | 1. text reinforcement<br>2. aural repetition<br>3. dictionary | Lower group: aural repetition-text reinforcement-(dictionary)<br><br>Higher group: text reinforcement-(dictionary) | Most beneficial modifications are:<br>1. text reinforcement<br>2. audio repetition<br>3. dictionary |
| Liou, 1997<br><br>Liou, 2000 | Research on on-line help as learner strategies for multimedia CALL evaluation.<br><br>Assessing learner strategies using computers: new insights and limitations. | 20 EFL college students divided into effective and ineffective group | Self-paced interactive video unit (IVD) | Eight help functions displayed on request<br>1. Chinese script<br>2. English script<br>3. gist<br>4. background information<br>5. idiom search (dictionary)<br>6. word search (dictionary)<br>7. repetition (current sentence)<br>8. repetition (previous sentence)<br>9. video control (pause and rewind) | 1. video controller rewind<br>2. English script<br>3. Chinese script<br><br>Ineffective group:<br>1. rewind<br>2. English script<br>3. Chinese script<br><br>Effective group:<br>1. English script<br>2. rewind<br>3. Chinese script | | pause 85%[a]<br>backward 85%<br>English script 80%<br>glossary 80%<br>repetition of previous sentence 80%<br>repetition 70%<br>Chinese script 55% |

Note. [a] Out of 100%

Table 2.1 (continued)

| Author year | Title | Participants | Materials | Help options used | Frequency of help option use | Patterns described | Usefulness (self-reported) |
|---|---|---|---|---|---|---|---|
| Pujola, 2002 | CALLing for help: researching language learning strategies using help facilities in a web-based multimedia program. | 22 Spanish adult EFL students divided into four groups of decoders:<br>Higher<br>Average<br>Lower<br>Poorer | Web-based multimedia program (imPRESSions) | Seven help facilities<br>1. dictionary<br>2. cultural notes<br>3. transcript<br>4. subtitles<br>5. video controls<br>6. feedback<br>7. exerts module | | | |
| Hegelheimer & Tower, 2004 | Using CALL in the classroom: Analyzing student interactions in an authentic classroom. | 91 beginner EFL students | CD program New Dynamic English | Three options available on request<br>1. ABC button (repetition of aural input accompanied by displayed text)<br>2. repeat button (aural repetition)<br>3. glossary | 1. Repeat<br>2. ABC<br>3. Glossary<br><br>High performance group<br>1. Repeat<br>2. ABC<br><br>Low performance group<br>1. ABC<br>2. Repeat | | |

## Navigational patterns in CALL multimedia environments

Description and analysis of sequences of learner interaction with a CALL program are believed to be central to the work in CALL (Chapelle, 2003). These interactional descriptions are necessary to improve the design and use of multimedia environments such as placement and content of help options (Chapelle, 2003, p. 104).

As can be seen from the previous section, navigational patterns of participants in multimedia listening programs were investigated only by Hsu (1994). As noted, patterns of interaction varied between groups so lower proficiency learners tended to request sentence repetition before requesting sentence transcription. More research is urgently needed to investigate patterns in more controlled environments (one of the objectives of this study) as well as open environments (Desmarais, Duquette, Renie, & Laurier, 1998).

Desmarais et al. (1998) described and accounted for patterns of interaction with a multimedia videodisc. The researchers used verbal protocols and computer tracking to describe participants' navigational schema which they consequently classified into linear and chaotic. In a linear schema, the participants completed the activities and moved through them in a linear order. In a chaotic schema, the participants switched from one activity to another and tended to not complete them. There were several possible explanations for the chaotic pattern of behavior; perhaps participants could not establish learning goals and plan the order of activities. Furthermore, they may have perceived the task as too difficult, boring, or not fitting their needs. The chaotic schema may also be connected with participants' exploratory behavior. The researchers found out that many students exhibited chaotic behavior during the beginning sessions while they were exploring the environment and later followed a more linear pattern as they became accustomed to the program. This may be a useful finding for CALL program design since users should be prompted to explore the program either by their teachers (as Hegelheimer & Tower (2004) pointed out) or through tutorials if they are using the software for independent learning. In another article about the same study, Desmarais, Laurier, and Renie (1998) researched whether the proficiency level had an influence on navigational patterns. They found that intermediate learners established a more linear pattern while the beginners engaged in activities that were sometimes too difficult. This difficulty

prompted learners to interrupt the activities which resulted in a more chaotic pattern. This is another example that navigational patterns may be influenced by proficiency level in addition to many other variables such as learning styles and strategies.

# CHAPTER 3. METHODOLOGY

This chapter gives an overview of materials and methods used in the study. It opens with the description of participants and criteria for their division into two proficiency groups. Then, materials used in the study are presented and two pilot studies described. In the next section, each of the two questionnaires (pre and post-listening) and two tests (initial and recall) are discussed. Then, the main parts (video, subtitles and transcript) of the CALL data collection activity--the Astronomy unit are discussed together with its design features. Procedures used in the data collection and the methods of analysis by each research question conclude this chapter.

## Participants

Participants in this study were 18 ESL college students enrolled in the English 101 Listening, Academic Listening for non-native speakers of English, at Iowa State University in the Spring of 2004. This class is a required course for both undergraduate and graduate students who had been admitted to Iowa State but did not pass the listening part of the English placement test, an in-house listening, reading and writing test for all students whose native language is other than English.

The participants were from 7 different countries: China (8), Korea (4), Japan (2), India (1), Malaysia (1), Peru (1), and Vietnam (1). There was approximately the same number of men and women (ten participants were male and eight female). Most of the students were graduate students (11), while 6 were undergraduates (one participant failed to provide this information). The participants were all from different departments apart from two graduate students in Physics and Astronomy and one with an undecided graduate major.

In order to determine their overall language proficiency, participants were asked to report their most recent TOEFL score. Their scores ranged from 500 to 615 on the paper-based version of the TOEFL test, with the mean score of 559 (SD 30). Based on their TOEFL scores, their performance on the listening portion of the English placement test, and the instructor's judgment, the participants were classified as being on the intermediate level of general listening proficiency.

All participants in the study were enrolled in the same section of English 101 Listening, which had 20 students. However, one student was absent during the data collection and one had been exposed to the lecture used in the CALL listening activity so the data for 18 students were analyzed in this study.

For the purpose of the study, the participants were divided into two groups-higher (10 students) and lower (8 students) based on their listening proficiency. This division was done taking into account the participants' self reported TOEFL scores, their present grade on the course, their self-evaluation of listening proficiency, instructor's opinion of their listening ability, and scores on an initial listening test. For participants whose scores did not clearly categorize them either into higher or lower group, the instructor's opinion was used as the criteria for the final decision. For more details on how two groups were made please see Appendix 8.

## Materials

The materials used in this study consisted of two questionnaires, two tests, and the CALL activity especially designed for this research. Also, selected participants were interviewed after the recall test so the unstructured retrospective interview was also used. The whole study and all the materials were integrated into the existing English 101 syllabus and made a part of the course. Each of the materials will be discussed in the following sections in this order:

Initial listening comprehension test-The Horticulture lecture

Pre and post listening questionnaires

Delayed recall test

Multimedia CALL listening activity-The Astronomy unit

However, since both of the questionnaires, the recall test, and the CALL activity were trialed and improved during two pilot studies, the overview of pilot studies will be given before each of the materials is described.

## Pilot studies

Two pilot studies were conducted before the main data collection to further develop and evaluate data collection tools as well as to test the main material the CALL activity. The first pilot study with nineteen students from English 101, which was conducted a semester before the main data collection, trialed the comprehension questions for the Astronomy unit as well as the recall test questions. The students were taking the same course as the participants in the study and showed a good learner fit with an average TOEFL score of 591 (SD 31) and self-assessment of listening proficiency as low intermediate and intermediate. After the students viewed the video of the Astronomy lecture segment by segment and answered 19 multiple choice and open-ended questions, their answers were statistically analyzed. Item difficulty and discrimination were calculated and distractors analyzed. As the result, some of the test items were changed, some abandoned, and some answers used as distractors. Additionally, 5 open-ended questions that the participants completed afterwards were, with minor modifications, later used in the recall test.

The second pilot study conducted a month before the main data collection trialed the pre and post activity questionnaires as well as the design of the CALL activity. Five ESL students who evaluated their listening proficiency as intermediate pilot tested the Astronomy unit under the same conditions as in the main data collection. It was important to check the procedures for video capturing as well as handling and storing of large files (400-500 MB) that the video capturing program recorded. The pilot showed that the students had enough time to complete the activity during one 50-minute class period. Their feedback on the activity was used to implement minor design changes and polish the questionnaires. More details about both pilot studies can be found in Appendix 9.

## Initial listening comprehension test

The initial listening comprehension test administered before the CALL activity helped divide participants into two proficiency groups. The test consisted of an academic lecture on Horticulture accompanied by 10 multiple choice questions. This fifteen-minute lecture on Horticulture (transcript can be found in Appendix 1) is a part of the same series of lectures as the lecture on Astronomy intended for use in the ESL program at Iowa State

University. Both lectures feature a faculty member lecturing on the topic of his/her expertise in front of a group of students. This lecture was chosen because of its similar format and length to the Astronomy lecture. The comprehension questions (Appendix 2) used in the test were for the most part adapted from those used by Kon (2002). These questions were chosen because they were of the same type (multiple choice with 4 options) as in the CALL activity and because they were already tested and analyzed by Kon. When the results of the test were looked at, it turned out that the activity was not able to distinguish between the higher and the lower group with 72% of the students in the middle group. This is the reason why these scores were not used as the determining criteria for division into groups; instead, the final selection process was informed by the instructor's opinion.

**Questionnaires**

The two questionnaires used in the study were the pre-listening questionnaire (Appendix 3) given to participants at the beginning of the study and the post-listening questionnaire (Appendix 4) given immediately after they had completed the CALL activity. Both questionnaires were trialed during the second pilot study (more details in the section on Pilot studies and in Appendix 9).

The aim of the pre-listening questionnaire was to examine participants' familiarity and previous use of two help options from the study (subtitles and transcript) because this may have had an influence on the choice of help in the activity. Also, it was designed to survey which help option students would prefer in a hypothetical multimedia listening activity.

At the beginning of the questionnaire, subtitles (closed captioning) and the transcript were described. For all questions, students ticked the answer that applied to them. Question 3 asked about the familiarity with subtitles and transcript either in their native language or in English. The purpose of this question was to determine whether students had been exposed to L1 or L2 (English) subtitles and whether they had used transcripts as help in listening exercises prior to the questionnaire. Also, the questionnaire surveyed students' use of English subtitles for watching TV and DVD (questions 4, 5 and 6). Finally, question 2 described a

hypothetical situation in which the subtitles and the transcript would be offered in a computer-based listening activity.

The purpose of the post-listening questionnaire was to examine the participants' use and perceived effectiveness of the transcript and subtitles in the multimedia activity. It consisted of selected response and open-answer questions, as well as ratings using a Likert scale.

Question 7 asked the participants which help option they would prefer if they could have more activities like the Astronomy unit in the future. This question and question 2 in the pre-listening questionnaire examined whether the participants' opinion changed after they had actually interacted with help options. A seven-point Likert scale was used to survey the difficulty and helpfulness of the unit, as well as the helpfulness of each help option.

**Recall test**

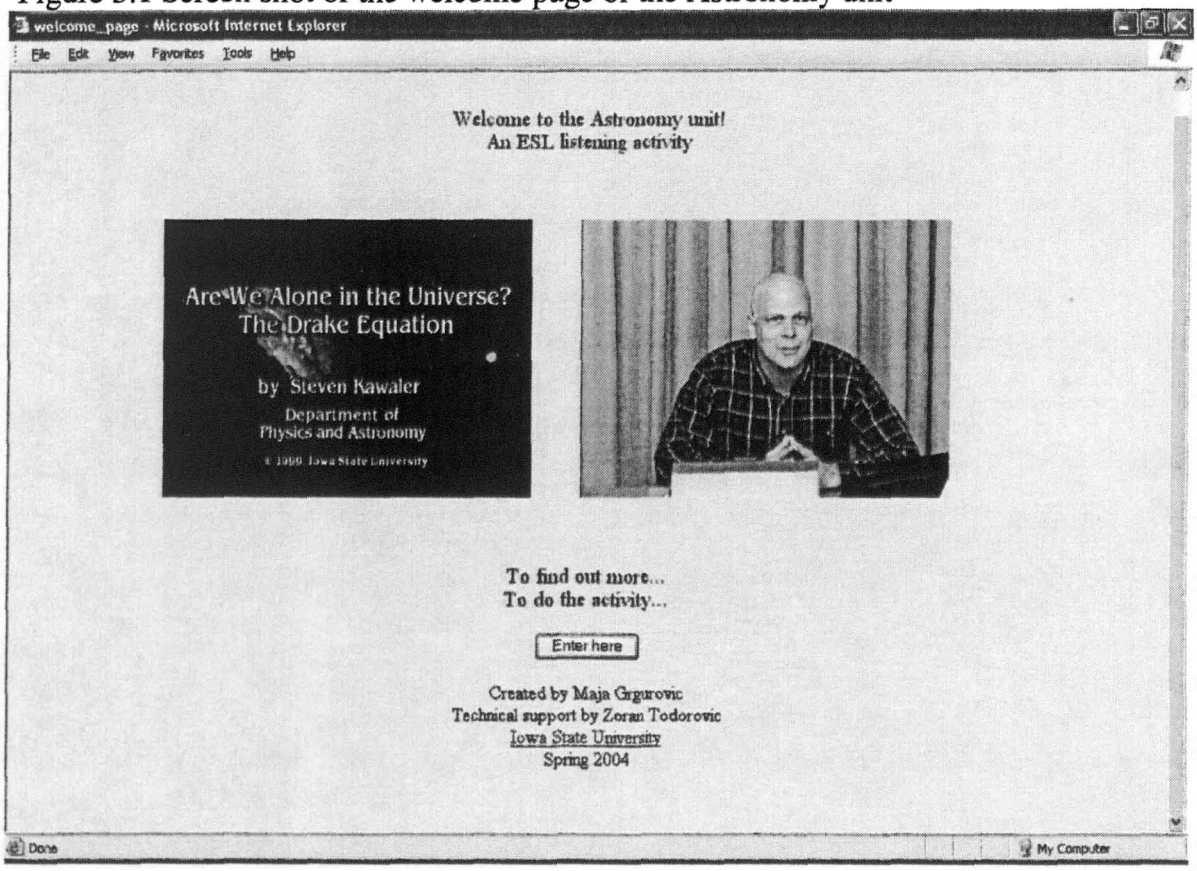
The final material that students were exposed to in the study was the delayed recall test (Appendix 7) which was given to the participants a week after they had completed the activity. The recall test was trialed during the first pilot study and aimed at checking the amount of information from the activity that the participants could recall. This was necessary to assess the learning outcome after a longer period and gather more quantitative data about the effectiveness of the activity. The test was conducted in class and it asked students to answer six open-ended questions. The questions covered the main ideas of the activity like its main topic (questions 1, 2) and the conclusion (questions 6). Also, it tested the main supporting details (questions 3, 4, 5).

**Multimedia CALL listening activity-the Astronomy unit**

The main material used in this study was the multimedia CALL listening activity entitled the Astronomy unit. This CDROM-based activity used multiple media--video, audio, text, and images to deliver an ESL academic listening activity. Through audio and video, ESL learners followed a lecture on Astronomy given by a professor in front of a group of students. As help in comprehending the video, text appeared in the form of two help options-English subtitles (Figure 3.3) and the lecture transcript (Figure 3.4). In case of the subtitles, the text appeared simultaneously with the video and audio input, while in case of

the transcript, the text was displayed either at the same time or separately from the video and audio depending on the learner's choice. The text also appeared in comprehension questions and in the tutorial (Figure 10.2 and 10.1 in Appendix 10) where the instructions how to use the unit were presented together with images of two help options. Two core components of the activity were the video and two textual help options which will be, together with unit navigation, discussed in detail in this section.

Figure 3.1 Screen shot of the welcome page of the Astronomy unit




A core component of the Astronomy unit was the video of a lecture delivered by Professor Steven Kawaler, a faculty member in the Department of Physics and Astronomy at Iowa State University. The video of the lecture "Are we alone in the Universe? The Drake equation" was eleven minutes long and for easier comprehension by ESL learners it was divided into ten segments of approximately 30 to 90 seconds (Appendix 5). This particular length was chosen because the literature suggests that for intermediate ESL students the optimal length of a segment should be between 30 and 120 seconds (Rubin, Quinn & Enos,

1988; Thompson & Rubin, 1993) since L2 listeners "may have limited memory capacity" (Rubin, 1995) especially in this case when they were not taking notes while viewing. The segments were divided based on thematical/logical units and cuts were made when the lecturer paused for several seconds and/or denoted a change of topic using a rhetorical device. Table 3.1 shows the division of the video into ten segments as well as the length of each segment while Figure 3.2 shows a screen shot of a video page.

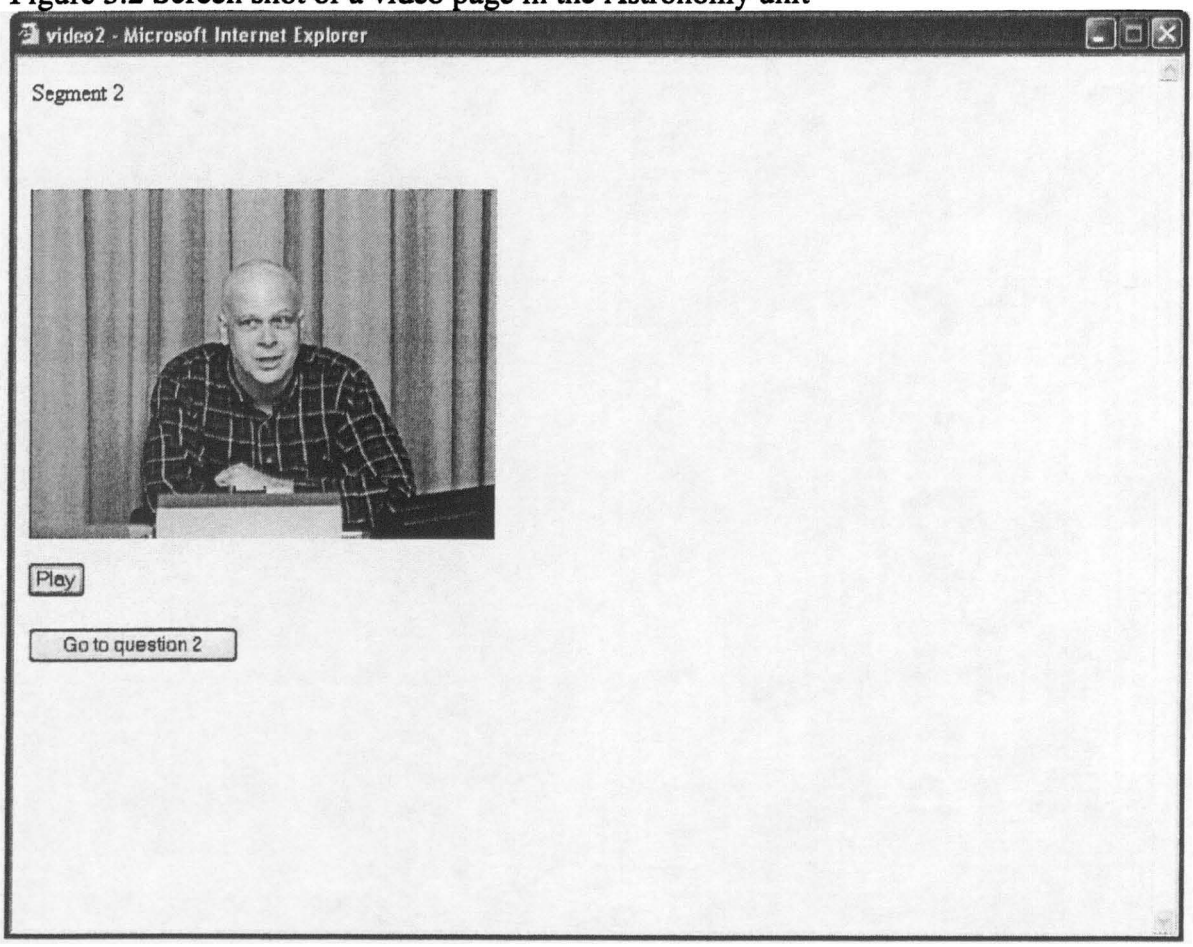
Table 3.1 Division of the video into ten segments

| Segment | Length (in seconds) | Start time (in minutes and seconds) | Stop time (in minutes and seconds) |
|---|---|---|---|
| Intro | 11 | 00:00 | |
| 1 | 86 | 00:11 | 01:37 |
| 2 | 58 | 01:39 | 02:37 |
| 3 | 63 | 02:38 | 03:40 |
| 4 | 73 | 03:41 | 04:54 |
| 5 | 65 | 04:55 | 06:00 |
| 6 | 31 | 06:01 | 06:32 |
| 7 | 62 | 06:34 | 07:36 |
| 8 | 90 | 07:37 | 09:07 |
| 9 | 55 | 09:08 | 10:03 |
| 10 | 74 | 10:04 | 11:18 |
| Mean | 65.7 | | |

When students clicked the play button, the whole video segment was played through and students could not replay it. This constraint was introduced on purpose to make the delivery of input as close to the one in a lecture hall where students don't hear the same lecture twice. This way, the viewing of the lecture for the first time was made as authentic as possible given that the lecture was used in ESL instruction.

Accompanying each video segment, there was a multiple choice comprehension question (sample question page is in Figure 10.2 in Appendix 10). Additionally, when the last video segment was over, there were four post-listening questions (the text of all the questions is in Appendix 6). These fourteen questions had been trialed in the first pilot study and later validated by three ESL instructors from the Department of English, one of them being an English 101 Listening instructor. In addition to content validation, the course instructor checked that the difficulty of the questions was appropriate for the participants in the study.

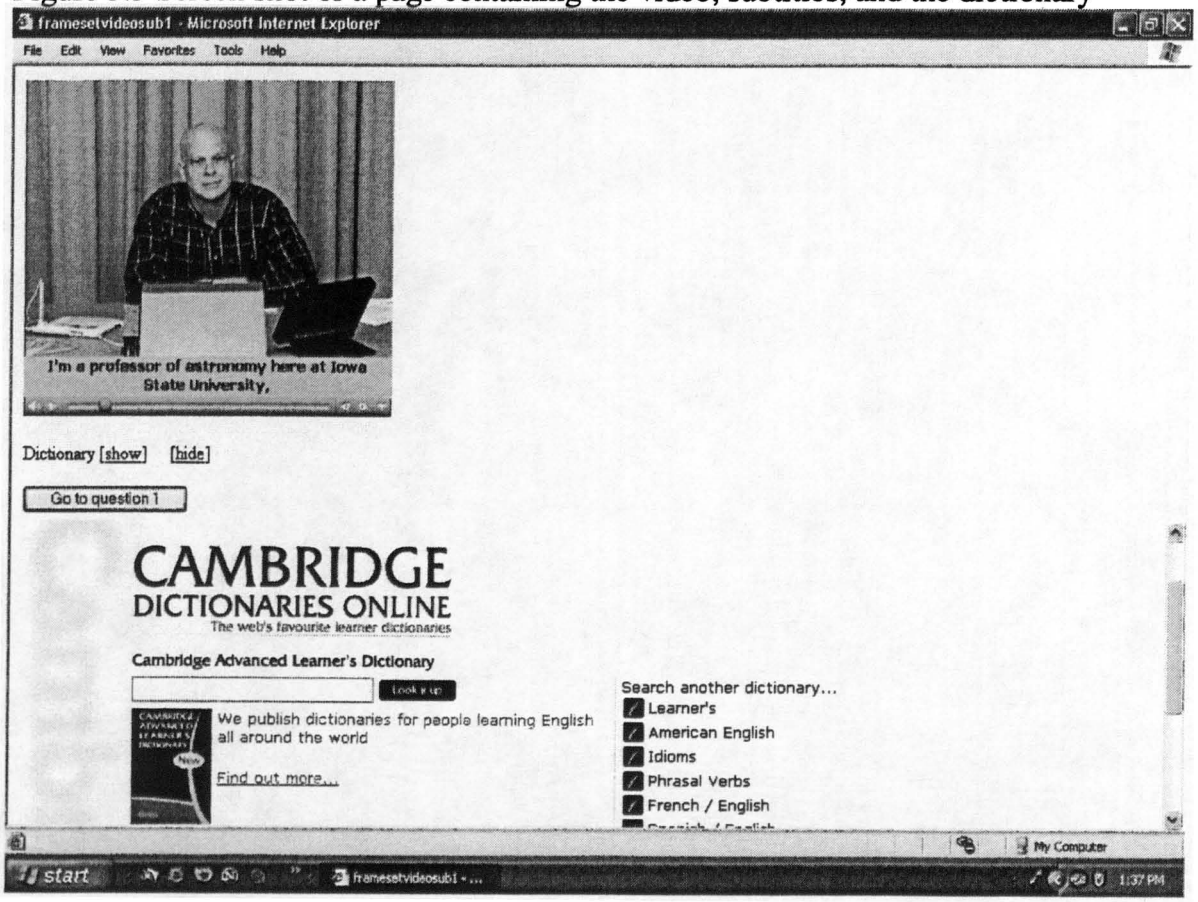
Figure 3.2 Screen shot of a video page in the Astronomy unit




Apart from the video, another core component of the unit was the set of two help options--the subtitles and the transcript. These options were given to students only after they had viewed the video segment and answered the comprehension questions incorrectly. This meant that help was offered only in cases of comprehension breakdowns when students needed to use help in order to move on. Special care was given to this issue during the design of the activity; the option to skip help and answer the question again from memory was not given to prompt the participants to make use of help. Each of the help options included the video segment and, for the second viewing, the participants could choose either to a) open the page with the subtitled video (Figure 3.3) or b) open the page with the lecture transcript and the video (Figure 3.4). In both cases, they had an access to the online English-English dictionary which was offered to participants as additional help in case they needed to look up unknown words. Although the dictionary is considered a type of help option (Hegelheimer &

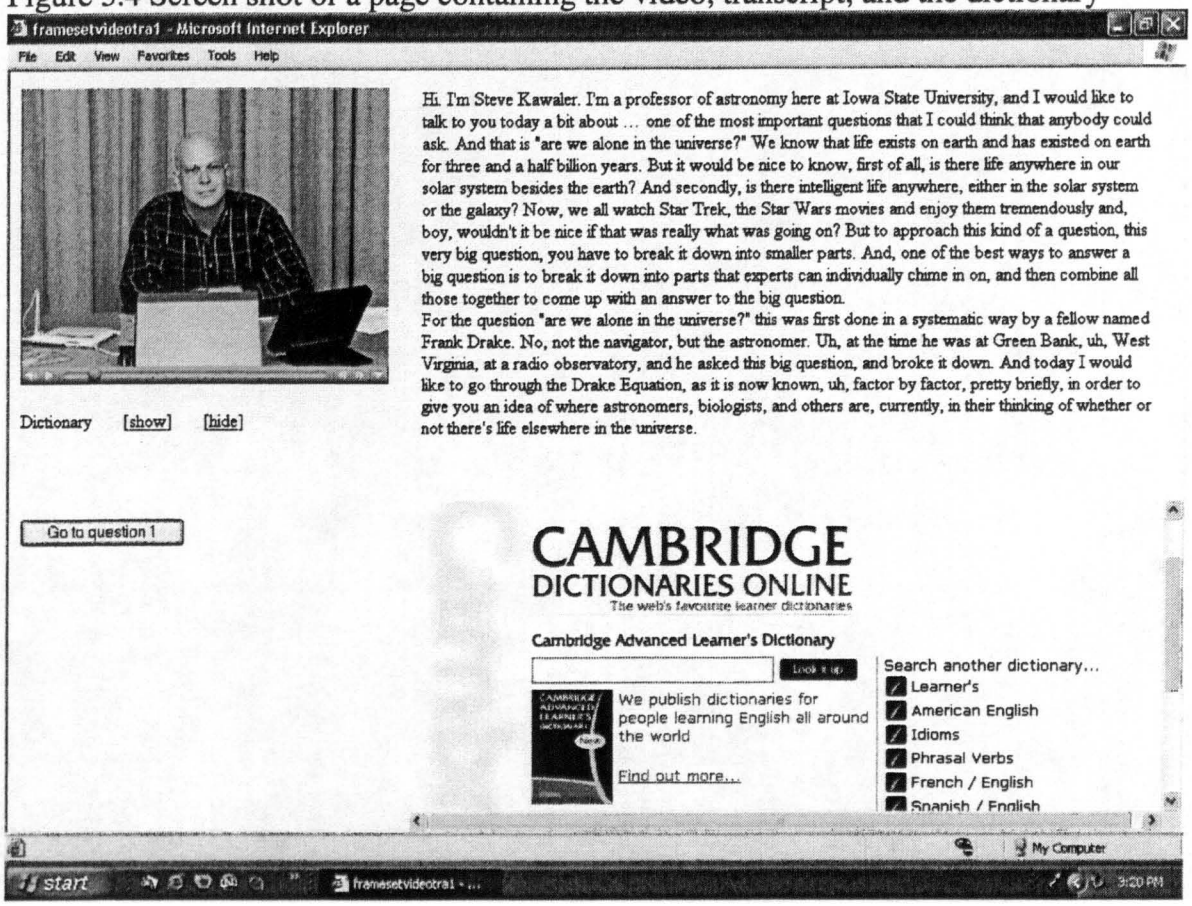
Tower, 2004; Hsu, 1994; Liou, 1997; Pujola, 2002), it is not the focus of this study and is not referred to as help option for the purpose of this research.

Figure 3.3 Screen shot of a page containing the video, subtitles, and the dictionary




The subtitles were synchronized with the video and appeared at the bottom of each video segment. The participants used the play and pause buttons to control the video and the subtitles. They could also rewind and fast forward the video using the video controller. The transcript for each segment appeared next to the video which was controlled in the same manner. A link to the Cambridge Advanced Learner's Dictionary appeared below the video and the dictionary opened in the bottom frame. The learners could either keep the dictionary open or hidden.

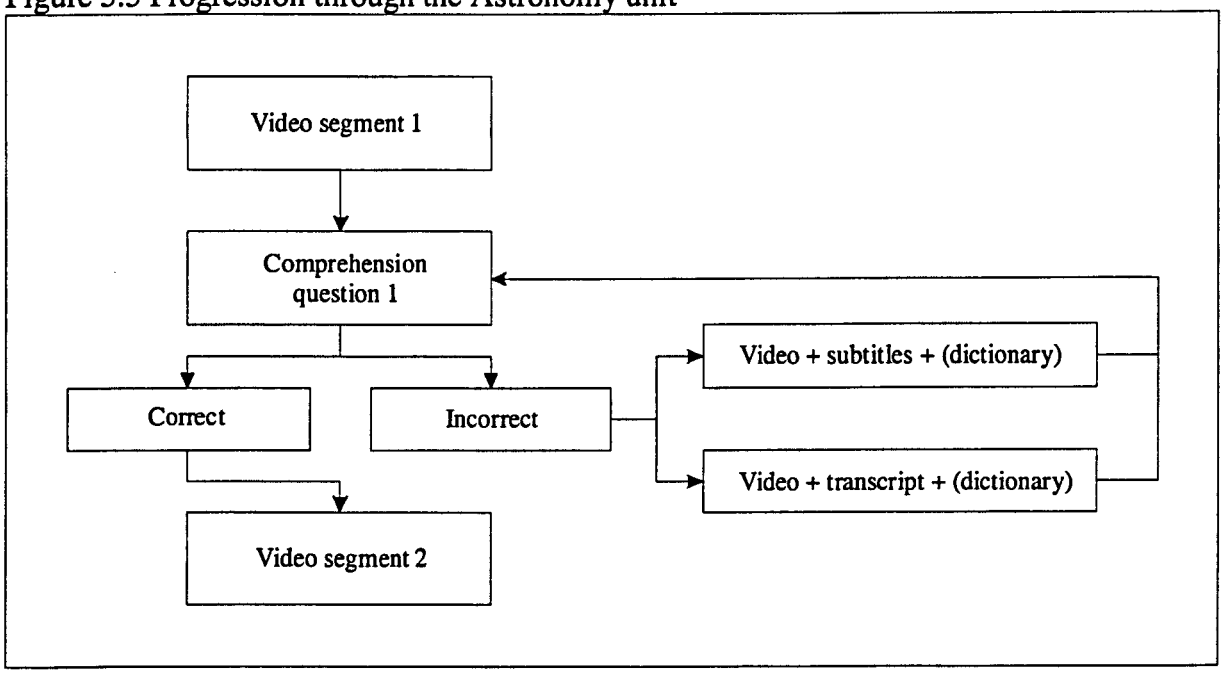
Figure 3.4 Screen shot of a page containing the video, transcript, and the dictionary




At the beginning of the unit, a tutorial explained how to go through the unit (Figure 10.1 in Appendix 10) and gave a sample video segment with help options that users could interact with. The participants were instructed to explore the tutorial before they started the activity to minimize requests for a help option made by guessing at what the option looked like.

In order to illustrate the participants' progression through the unit, Figure 3.5 shows movements from the first to the second segment.

Figure 3.5 Progression through the Astronomy unit




After video segment 1, which could be viewed only once, the participants were given a comprehension question that tested their understanding of the segment. This segmentation allowed them to check and monitor comprehension of the video as they went through the segments. If the answer was correct, they proceeded to segment 2. However, if the answer was incorrect, they were offered to view the same segment again either with the lecture transcript or the subtitles. The video had to be played to see the subtitles but the transcript was displayed even if the video was not played. When on the transcript page, the learners could choose whether to play the video or not. If necessary, they could also consult the dictionary. The participants could interact with help options for as long as they wanted to. Then, when they were ready, a slightly modified version of the first comprehension question was presented. The question stayed the same but the order of answers was changed to prevent guessing. If the correct answer was chosen, they went on to the next segment, if not, they were given the same help options until they got the correct answer. After they had answered all ten questions, four post-listening questions were displayed (Figure 10.3 in Appendix 10). When the students chose the answers and clicked the submit button, the program calculated their score. The activity was over when their score for the post-listening questions was displayed.

**Design issues in the Astronomy unit**

The Astronomy unit was created having in mind English 101 Listening students as primary users so the design choices are justified by the fact that the activity was done in an intact ESL class during one class period. The main design choice made regarded the issues of user navigation and control of the environment. Namely, the progression through the unit is linear and users can not return to the previous segment. Additionally, the use of help is compulsory when a comprehension breakdown occurs. The participants don't have control over first viewing of the video so it could be said that the activity is generally more program controlled than other multimedia materials used in similar studies (Hegelheimer & Tower, 2004; Liou, 1997; Pujola, 2002).

To complete the activity, the students needed to move through the unit in a linear fashion (from first to the last segment) and could not skip any segments. In order to ensure this, the right mouse click was disabled and the toolbar with the back button removed. Furthermore, each new page opened in the same window. Also, going through help options was a *compulsory* step in moving forward in cases of comprehension breakdowns because that is thought to increase the use of help. After the participants received feedback that their answer was incorrect, there were links to the subtitles and the transcript pages but no option to skip help. All of these were done to narrow the realm of investigation, try to control for variables, and prompt the use help.

When it comes to program control over exposure to video segments, once the participants clicked the play button to view a segment for the first time, the button disappeared and the video could not be played again. Also, once they were at the comprehension question, they could not go back and replay the video. As already mentioned, this was done to keep the exposure to the lecture the same for all students, to make listening to the lecture similar to the lecture hall listening, and to prevent some students from replaying the video, instead, if necessary, using video through help. When the video was used in a help option with the subtitles and transcript, it had control functions (play, pause) and the video controller. In this case, the learners had more control over the viewing and reviewing process. Thus, it could be said that the control in the Astronomy unit shifted from the complete control of the program at the beginning of the unit to the user control over the

choice between two help options and finally to user control of the interaction with help options.

Although the move towards the authenticity of use was made in the delivery of the input during the first viewing of the video segments, a reasonable claim can be made that that Astronomy unit may not be the type of material learners would encounter outside their ESL classes. However, the Astronomy unit was not created only to give students more exposure to the academic materials but also to facilitate their use of help and give them opportunities to monitor and check their listening comprehension. Additionally, participants in the study were taking 101 Listening because they had problems comprehending real-life lectures, so a task of this kind that has a number of target language use characteristics can be seen as an intermediate step towards the actual target language use task. In sum, I believe that the optimal design choices were made representing a balance between a language learning potential of the activity, target language use tasks, and the experimental nature of the study.

**Software and hardware requirements**

The Astronomy unit was created using Macromedia-Dreamweaver MX, a program for designing web pages. It was decided that seven web pages would be associated with each video segment and that each page would open in the same window. The subtitles and transcript pages had two and four frames, respectively. Course Builder, a program extension within Dreamweaver, was used to create multiple choice questions and provide immediate feedback to students. More advanced JavaScript programming was required for the issues of control such as opening and closing of windows that have no toolbars and disabling going back.

The video was edited in Quick Time Pro 6.0 and cut into ten segments. Then, text files were made using the lecture transcript for each segment. Subtitles were created from text files by converting them into movie files which were later added to the video and synchronized with the audio. Having subtitles integrated in the movie made them much easier to operate in comparison to cases when it was up to the user to achieve synchronization (Pujola, 2002). Next, the video was embedded into web pages which were then linked to each other. Although it was originally thought that the activity would be

exclusively on the Internet, this changed to prevent possible problems with low bandwidth of the network and limited server capacity to answer requests for data from ten users at a time. This was especially an issue since the video files ranged from 3 to12MB totaling 156MB of video for the whole activity. Consequently, the activity was put on hard drives of local machines and accessed locally. Although all the files were on local computers, Internet connection was present for accessing the online dictionary.

The activity was conducted in a computer lab with 10 Pentium III PC computers (996 MHZ with 256MB RAM, 20GB hard drive and a sound card) running Windows XP Professional (Service Pack 2). The computers were equipped with a headset and the screen resolution was set to 1024 x 768 for best viewing. In Internet Explorer 6.0, which was used for viewing the activity, the text size was set to medium and the top navigation bar removed.

The screen capturing program Camtasia Studio was used to record exactly what students did as they went through the activity. The Camtasia Recorder was set to capture the whole screen as well as mouse clicks which appeared highlighted. Also, hardware acceleration on each machine was disabled so that Camtasia Recorder could capture video. Although Camtasia can record external audio, this option was not used because students listened using headphones. This decision had to be made because the data collection was conducted in one computer lab with ten students listening at the same time. If they were listening using speakers, they would disturb each other so a tradeoff was made for sound not being recorded by Camtasia.

## Procedures

This study was conducted as a part of the English 101 Listening syllabus and took part over 3 weeks towards the end of Spring semester 2004. The class met twice a week (Monday and Wednesday) for 50 minutes in a classroom or a computer lab depending on the activities. I met with the participants during their regular class except for the interviews, which were scheduled outside the class time. The participants were given an overview of the study during the first session and told it was a required part of their course but that the data would not be used for research purposes if they did not want to give their consent.

Fortunately, all participants agreed to participate in the study and signed informed consent forms. The procedures used in the study session by session are given in Table 3.2.

Table 3.2 Overview of the procedures used in the study

| Sessions | Week/Day | Activity/place | Activity/place | Duration |
|---|---|---|---|---|
| 1. | Week1-Wednesday | Pre-listening questionnaire, division into two groups (classroom) | | 20 minutes |
| 2. | Week 2-Monday | Group one-CALL activity and post-listening questionnaire (computer lab) | Group two-diagnostic test (computer lab) | 50 minutes |
| 3. | Week 2-Wednesday | Group two-CALL activity and post-listening questionnaire (computer lab) | Group one-diagnostic test (computer lab) | 50 minutes |
| 4. | Week 3-Monday | Delayed recall test (classroom) | | 20 minutes |
| 5. | Week 4-Wednesday and Thursday | Interviews with selected participants (computer lab) | | 30 minutes per participant |

Apart from the overview of the study, in the first session the participants filled in the pre-listening questionnaire and were divided into two groups of ten according to the alphabetical class list. Since the computer lab used in sessions 2 and 3 had ten computers equipped with the screen recording software, only ten students could do the Astronomy unit at the same time. As a result, the next class period, the first group of students met with me to work on the Astronomy unit and the second group met with the instructor for the listening test. The next session, the groups switched activities.

Before students started the Astronomy unit, they were given verbal instructions how to use the activity. They were briefly told what the activity consisted of and how to move through it. Also, they were directed to explore the tutorial, read the detailed instructions, and

check how the two help functions worked before they moved to the first video segment. The verbal and written instructions were given to familiarize students with the layout and help functions and the tutorial was used to give students practice so that they could use the environment to their advantage. The importance of teacher instructions and demonstration of CALL materials was suggested to be beneficial for the participants' use of help (Hegelheimer & Tower, 2004). When the students were given headphones and after the Camtasia recording program started, they began work on the Astronomy unit.

Pujola (2002) believed that screen video recording was the fundamental method in data collection because it "provides accurate information about what learners *actually* do" (p. 244). Pujola (2002) identified screen recordings together with observations of learners working on the task as "two complimentary methods which provide rich and valuable data" (p. 258). In this study, the screen recordings were indeed the fundamental method of the data collection because they were the only way to know what exactly happened in the activity. Due to the data collection setting, it was impossible for a single researcher to observe every individual but I did observe the whole group and took down notes on student interaction with the activity, which later helped me select participants for retrospective interviews. Also, notes were important when viewing the Camtasia videos because I sometimes had explanations for atypical student behavior (for example when students experienced technical difficulties). This session ended with students filling the post-listening questionnaire.

Concurrently, the other group of students met with the instructor in another computer lab and was shown the Horticulture lecture on the video projector. They were instructed to answer comprehension questions on paper as they listened. The score on this test was added to other participants' information which was either self-reported or gathered from the instructor so that the division into two proficiency groups could be made.

Immediately after the Astronomy activity session ended, the student data were saved on local machines. On average, the Camtasia video files were between 400 and 500 MB so they were difficult to handle and store. For easier handling and data analysis, one CDROM was burned for each participant. A week later, in session 4, the participants did a delayed recall test in class and retrospective interviews were announced.

After looking at post-listening questionnaires, reviewing data collection notes and several videos, 6 participants were invited for retrospective interviews. Those were students whose answers and/or behavior attracted my attention either because they were shifting between two help options or not using help at all. Only three out of six participants agreed to meet me for individual interviews where I showed them their pre and post listening questionnaires and parts of their Camtasia video. I was interested in finding out the details about their choice of help and therefore reviewed videos with them to get their response (especially if they shifted between two help options). When all the data were collected, Camtasia videos of all participants were transcribed and methods of analysis for each research question determined precisely.

## Analysis

To answer the four research questions in this study both quantitative and qualitative data were used. Before the data analysis could be conducted, it was necessary to transcribe all Camtasia videos of participants interacting with the CALL activity because the transcriptions could show what exactly happened in the activity. Also, the answers to questionnaires and the recall test were complied. Finally, descriptive and inferential analyses were performed on the data for all eighteen participants. In this segment, methods of analysis used for each research question will be presented.

**Research question 1: How frequently and for how long do participants use two help options (subtitles and the transcript) offered in the multimedia activity--the Astronomy unit? How do these two help options compare in terms of frequency and time of use?**

To answer the first research question, the number of times each participant opened each help page was calculated since it was possible to open the help option page several times depending on how many times the answer was incorrect. During the calculations, it was noted that some students opened the help option page but would not interact with the help option. For these participants, opening the help page was the only way to continue the activity since there was no other way to get to the comprehension question. At the same time,

these participants spent a very short amount of time on the help page (for example, just enough for a page to load which usually took between 2 and 5 seconds depending on the size of the video file on that page). Consequently, help page openings did not give precise information about the help function use so criteria for *useful interaction* with help options were established after all the transcriptions had been looked at. Table 3.3 contains criteria for useful interaction for each of the help options.

Table 3.3 Criteria for determining useful interaction and non-interaction with help options

| Help page | *Useful interaction* | Non-interaction |
|---|---|---|
| Subtitles | *Subtitled video played* | No subtitled video played |
| Transcript | *Seven or more seconds spent on the page* | Less than seven seconds spent on the page |

In case of the subtitles, the time the participants played the subtitled video was counted as useful interaction time no matter how many seconds he/she spent on that page. In case of the transcript, it was much more difficult to judge useful interaction because there were no data collection tools to determine whether the participant was reading the transcript or not, when the video was not played. Nevertheless, it was decided that seven seconds would be a cut off time, which meant that if seven seconds or more were spent on the transcript page (playing the video, reading the transcript, or doing both) that time counted as useful interaction with help. The time on useful interaction with help was the only time included in the analysis in the second and third research question. As in case of time, the help openings that involved useful interaction with help (called *instances*) were looked at separately to determine the exact help use.

In addition to help page openings and instances of useful interaction, descriptive statistics for time on each help page were obtained. Furthermore, paired t-tests were performed to compare the two help options on all of the above mentioned variables.

**Research question 2: Are there patterns of participants' behavior and how could those patterns be explained? If there are patterns, how do groups of participants following those patterns compare in terms of comprehension, and time and frequency of help use?**

To answer the second research question, transcriptions of participant interaction with the activity were reviewed. In particular, the total number of help page openings, the total number of interactions with help (*instances*), and time in useful interaction with each help option were examined to group participants according to common navigational behavior. Based on these, four different patterns of interaction with help options during the activity were established:

1) Subtitles pattern
2) Transcript pattern
3) Non-interaction pattern
4) Mixed pattern

According to four patterns of interactions, four participant groups were made (more details on characteristics of each group in the Results section) and their behavior discussed. Means for six variables (incorrect answers, time on help page, help page openings, instances of interaction, post-listening questions, and the recall test questions) were also calculated for each group. After that, the means were compared using a single factor ANOVA to see whether there were statistical differences among groups, followed by a post hoc analysis (Tukey's Honestly Significant Difference, HSD) to determine exactly which group(s) were different.

**Research question 3: Do participants at different proficiency levels vary in their use of help options?**

Before the analysis for this research question could be performed, it was necessary to divide participants into groups based on their proficiency levels. The rationale for the

division into groups was given in Appendix 8. Once a higher group with ten students and a lower with eight were made, they were compared across ten variables (incorrect answers, time on subtitles page, time on transcript page, total time on help page, help page openings, instances of interaction, subtitles instances of interaction, transcript instances of interaction, post-listening questions, and the recall test questions). Then, a two-tailed t-Test was administered with the p value set at $p \leq 0.05$.

**Research question 4: What are participants' attitudes towards two help options before and after the activity?**

The fourth research question examined the attitudes of participants towards the subtitles and the transcript before and after the Astronomy unit. The students filled two questionnaires and their responses were compared to see how many students changed their preferences. Also, the reasons for the change of preferences were examined.

# CHAPTER 4. RESULTS AND DISCUSSION

This chapter presents the results of the study and supplies findings by four research questions. First, the use of two help options after comprehension breakdowns in a CALL listening activity is presented by looking at the time and frequency of interaction. Two help options are further compared on time and frequency of use. Next, four navigational patterns of interaction with help options are reported and accounted for. Then, four groups of participants following those patterns are compared on performance and behavior in the activity. In the next section, behavior and performance of participants at two different proficiency levels are discussed and compared. Participants' attitudes towards help options are given in the last part of the chapter.

## Help option use

The first research question in the study investigated how frequently and for how long participants opened pages with two help options--subtitles and the transcript. To determine the frequency and time of help option use for each participant, descriptive statistics were used. This research question also examined the relationship between two help options in terms of frequency and time by comparing the use of the subtitles and the transcript. For inferential statistics of this kind it is necessary to use larger numbers if the results are to be generalized to other contexts but the set-up of the study and the use of an intact class of 18 participants did not allow for that. As a result, these statistical results could detect significant trends, which could be explored with larger samples in future investigations.

### Frequency of help option use

After the transcriptions of participants' interaction were examined, the number of help page openings in total and by help function was calculated. Also, the instances (useful interaction) with help options were counted and are presented in Table 4.1.

Table 4.1 Number of help page openings and number of instances of interaction with help

| Participant N=18 | Subtitles page openings | Transcript page openings | Help page openings | Instances of interaction | Incorrect answers |
|---|---|---|---|---|---|
| 1 | 12 | 2 | 14 | 1 | 7 |
| 2 | 6 | 0 | 6 | 5 | 5 |
| 3 | 13 | 0 | 13 | 1 | 7 |
| 4 | 4 | 0 | 4 | 4 | 4 |
| 5 | 3 | 1 | 4 | 4 | 3 |
| 6 | 0 | 2 | 2 | 2 | 2 |
| 7 | 9 | 13 | 22 | 5 | 7 |
| 8 | 15 | 0 | 15 | 0 | 9 |
| 9 | 0 | 9 | 9 | 7 | 5 |
| 10 | 17 | 2 | 19 | 15 | 7 |
| 11 | 5 | 0 | 5 | 5 | 4 |
| 12 | 7 | 0 | 7 | 5 | 5 |
| 15 | 6 | 0 | 6 | 4 | 5 |
| 16 | 5 | 0 | 5 | 3 | 3 |
| 17 | 4 | 4 | 8 | 4 | 5 |
| 18 | 1 | 6 | 7 | 6 | 5 |
| 19 | 10 | 0 | 10 | 1 | 5 |
| 20 | 7 | 0 | 7 | 2 | 3 |
| Totals | 124 | 39 | 163 | 74 | 91 |
| Totals in % | 76 | 24 | 100 | 45 | |
| Mean | 6.88 | 2.17 | 9.06 | 4.11 | 5.06 |
| SD | 4.95 | 3.68 | 5.47 | 3.34 | 1.79 |

As can be seen in Table 4.1 the subtitles page was opened 6.88 times and the transcript page 2.17 times on average during the course of the activity. Consequently, subtitles page openings account for 76% of all the openings--three times more than the transcript page openings (24%).

When looking at individual help openings, participants exhibited great variation from no help openings of transcript, for example, to 17 openings of subtitles. This variation is reflected in the huge standard deviation (SD for subtitles 4.95, SD for transcript 3.68). It is important to keep in mind that the number of openings depended on the number of incorrect answers on each comprehension question offered for the first time (Incorrect answers column in Table 4.1). Moreover, the participants could open help more than once for one

comprehension question which explains why the mean for the help page openings is higher than the mean of incorrect answers after first viewing of each segment.

When looking at instances of interaction (*useful* interactions with help options after the page has been opened), Table 4.1 shows that 45% of help openings resulted in help use. The remaining 55% of the openings were classified as non-interaction because the participants opened help (since that was the only way to proceed with the activity) but they did not need or want to use it. This is an important issue for research into design and implementation of help options in CALL listening materials since help may not be always used although it is readily available. This may be due to a number of factors such as interface design and time on task. Thus, it is possible to add to claims that there may be more help than needed in multimedia materials (Liou, 1997; Pusack & Otto, 1997). Also, it can be concluded that the frequency of help page openings is not a good indicator of the actual help use.

To examine more closely the cases when an opened help page resulted in help use, the distribution of instances by two help options is presented in Table 4.2. When looking at the number of subtitle instances in Table 4.2, it can be seen that participant 10 had by far the highest frequency of useful interaction with the subtitles (13) in comparison to the frequency of useful interaction of other participants which ranged from 0 to 5. In this case, participant 10 can be considered an outlier but his data were kept in the calculation of the mean for this variable because it is important to note the highest as well as lowers frequency of interaction with subtitles. Additionally, the relationship between the means for the subtitles and the transcript instances of interaction would not change if the data for participant 10 were removed from the calculation.

Table 4.2 shows that the participants interacted with subtitles more than with the transcript (means 2.66 and 1.44 respectively). Sixty five percent of all interactions are those with the subtitles which is almost twice as many as with the transcript (35%). However, when 124:48 margin between subtitles page openings (Totals for subtitles in Table 4.1) and subtitles instances (Totals for subtitles in Table 4.2) is compared with 39:26 margin for transcript page openings (Totals for transcript in Table 4.1) and transcript instances (Totals for transcript in Table 4.2), it can be concluded that the participants opened subtitles page

more but usefully interacted with it less, while the transcript page was requested less but when it was opened it was used more.

Table 4.2 Number of instances of interaction by help options

| Participant N=18 | Subtitles instances | Transcript instances | Instances Total |
|---|---|---|---|
| 1 | 0 | 1 | 1 |
| 2 | 5 | 0 | 5 |
| 3 | 1 | 0 | 1 |
| 4 | 4 | 0 | 4 |
| 5 | 3 | 1 | 4 |
| 6 | 0 | 2 | 2 |
| 7 | 1 | 4 | 5 |
| 8 | 0 | 0 | 0 |
| 9 | 0 | 7 | 7 |
| 10 | 13 | 2 | 15 |
| 11 | 5 | 0 | 5 |
| 12 | 5 | 0 | 5 |
| 15 | 4 | 0 | 4 |
| 16 | 3 | 0 | 3 |
| 17 | 0 | 4 | 4 |
| 18 | 1 | 5 | 6 |
| 19 | 1 | 0 | 1 |
| 20 | 2 | 0 | 2 |
| Totals | 48 | 26 | 74 |
| Totals in % | 65 | 35 | 100 |
| Mean | 2.66 | 1.44 | 4.11 |
| SD | 3.20 | 2.15 | 3.34 |

One of the explanations for this can be purely technical in nature i.e. the link to the subtitles page was always displayed first on the page with the link to the transcript page under it. It can be speculated that the participants who wanted just to continue past help and go to the question would usually open the first link on the page that way causing a higher number of subtitles page openings. Other explanations could be that it was easier for the participants to play the video with subtitles than read the transcript and play the video at the same time. Moreover, some participants had used subtitles before and were accustomed to

them. Indeed, five participants who exclusively used the subtitles always (4) or sometimes (1) watched subtitled movies on DVD. For participants' comments on the use of subtitles, see the section on Patterns of interaction with help options below. Although, as already mentioned, the use of dictionary was not the subject of investigation in this study, it is interesting to note that not a single student opened the dictionary link nor interacted with the dictionary which was available in both help options. This finding was also reported in the literature since some other studies found low levels of dictionary use (Hegelheimer & Tower, 2004; Kon, 2002).

**Time on help options**

As already noted, the amount of time the participants spent on help options in the activity was also investigated in the first research question. When the transcriptions of participants' interactions with the activity were looked at, only the time on useful interaction with help options was counted as evidence of time on help. As already noted in the Analysis section, useful interaction time on the subtitles page was the time participants watched the subtitled video (on the subtitle page) and on the transcript page it was the amount of time (seven seconds or higher) spent reading the transcript, watching the video or doing both. The distinction between time on the help page and time on useful interaction with help options had to be made because the time on help page was not a good predictor of time the participants spent interacting with help on that page--just like the frequency of help request could not predict the actual help use. This finding has implications for data collection methods in studies investigating participants' interaction with CALL software. If participants' behavior is recorded using a tracking system that records only time spent on help, it may not be possible to capture the true interaction. Therefore, complimentary data collection methods are necessary (for example tracking system in addition to screen recordings, participants' observations, video recordings of data collection settings). This was one of the remarks in other help option studies (Pujola, 2002, p. 258).

In this study, only time on useful interaction (termed time) was investigated and the breakdown by help option and participants is given in Table 4.3.

Table 4.3 Time on each help option and total time on help

| Participant N=18 | Subtitles | Transcript | Total time on help |
|---|---|---|---|
| 1 | 00:00:00 | 00:00:07 | 00:00:07 |
| 2 | 00:04:49 | 00:00:00 | 00:04:49 |
| 3 | 00:01:39 | 00:00:00 | 00:01:39 |
| 4 | 00:05:40 | 00:00:00 | 00:05:40 |
| 5 | 00:03:40 | 00:01:55 | 00:05:35 |
| 6 | 00:00:00 | 00:01:16 | 00:01:16 |
| 7 | 00:00:57 | 00:01:32 | 00:02:29 |
| 8 | 00:00:00 | 00:00:00 | 00:00:00 |
| 9 | 00:00:00 | 00:07:11 | 00:07:11 |
| 10 | 00:08:48 | 00:01:45 | 00:10:33 |
| 11 | 00:06:05 | 00:00:00 | 00:06:05 |
| 12 | 00:04:11 | 00:00:00 | 00:04:11 |
| 15 | 00:02:17 | 00:00:00 | 00:02:17 |
| 16 | 00:01:54 | 00:00:00 | 00:01:54 |
| 17 | 00:00:00 | 00:03:13 | 00:03:13 |
| 18 | 00:00:45 | 00:04:18 | 00:05:03 |
| 19 | 00:01:50 | 00:00:00 | 00:01:50 |
| 20 | 00:02:08 | 00:00:00 | 00:02:08 |
| Totals | 0:44:43 | 0:21:17 | 1:06:00 |
| Totals in % | 67.75 | 32.25 | 100 |
| Mean | 0:02:29 | 0:01:11 | 0:03:40 |

*Note.* Time is displayed in hours, minutes and seconds.

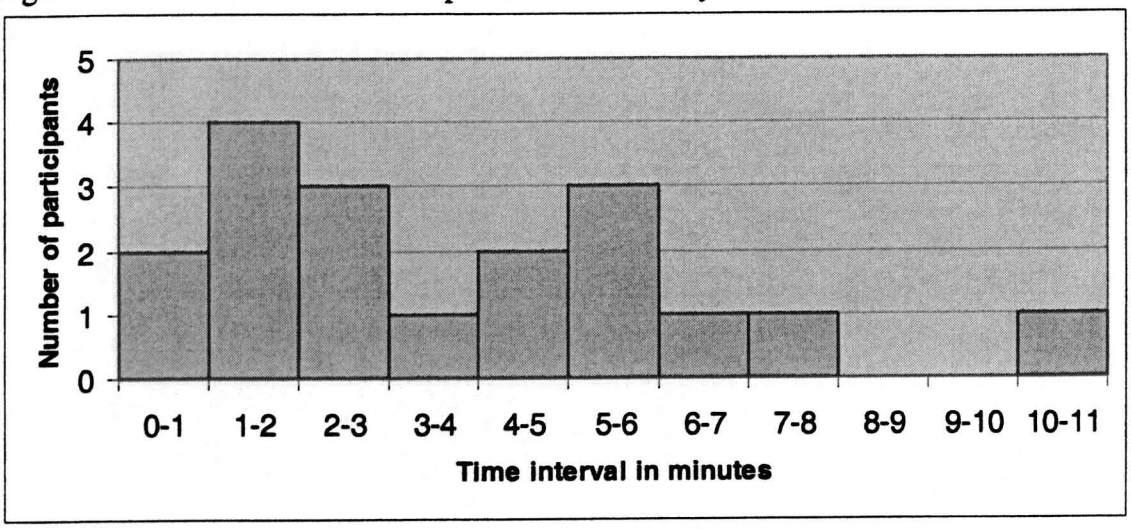
Table 4.3 shows that the participants spent twice as much time on the subtitles as on the transcript (total times are 44 min 43 sec and 21 min 17 sec respectively). It seems that participants made more use of subtitles judging both by time and frequency of interaction.

Table 4.3 also shows that participants within the group varied in the time on help. For example, participant 1 spent only seven seconds on help in the whole activity and participant 8 did not usefully interact with help at all. However, it was not because they performed well that these participants did not use help. On the contrary, their comprehension in the activity was rather weak. They incorrectly answered 7 and 9 questions respectively (see the number of incorrect answers in Table 4.1) although they opened help pages 14 and 15 times respectively (the number of help page openings in Table 4.1). The lack of comprehension of

content in many video segments in the activity, high number of help openings in relation to low number of interactions with help suggests that these two students needed help but did not make proper use of it. The small amount of time they spent interacting with help further supports this claim.

To further examine time on help options, distribution of time on help within one minute intervals is given in Figure 4.1.

Figure 4.1 Distribution of total help time in the activity in one minute intervals




The Figure 4.1 shows that the highest number of participants per interval is in 1-2 minute interval (4 participants). Additionally, it can be seen that the majority of students spent between 0-3 minutes (9 participants) and 4-6 minutes (5 participants) on help in total. The highest amount of time spent on help was 10 minutes while the participants engaged in the whole activity between 18 and 32 minutes (Table 11.1 in Appendix 11). This would suggest that the maximum help time is approximately one third of the time on the whole activity--the case of participant 10 who spent the longest time on help (10min 33sec) and on the activity overall (32min 38sec).

Time on each help option can be seen on Figure 4.2 (subtitles) and Figure 4.3 (transcript).

Figure 4.2 Distribution of time on subtitles in the activity in one minute intervals (N=18)

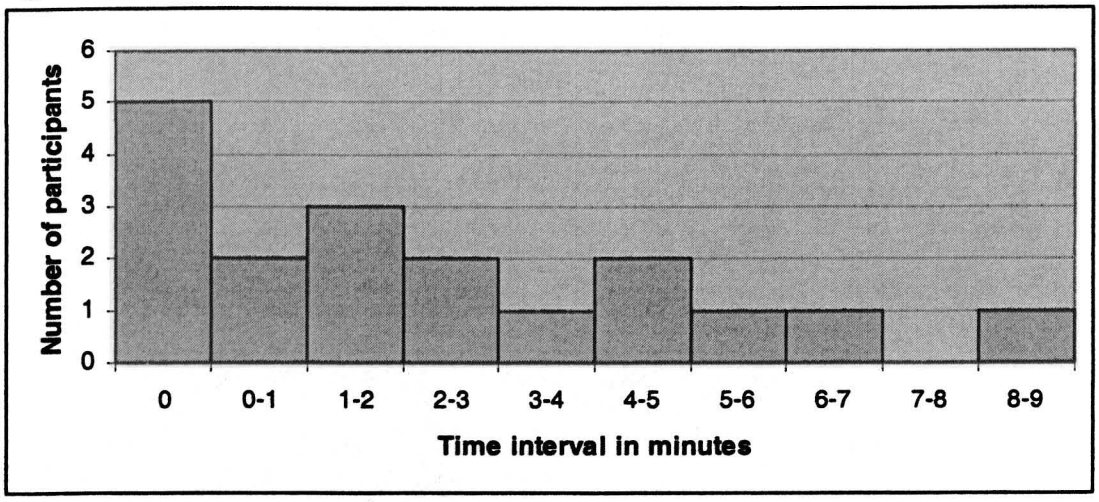
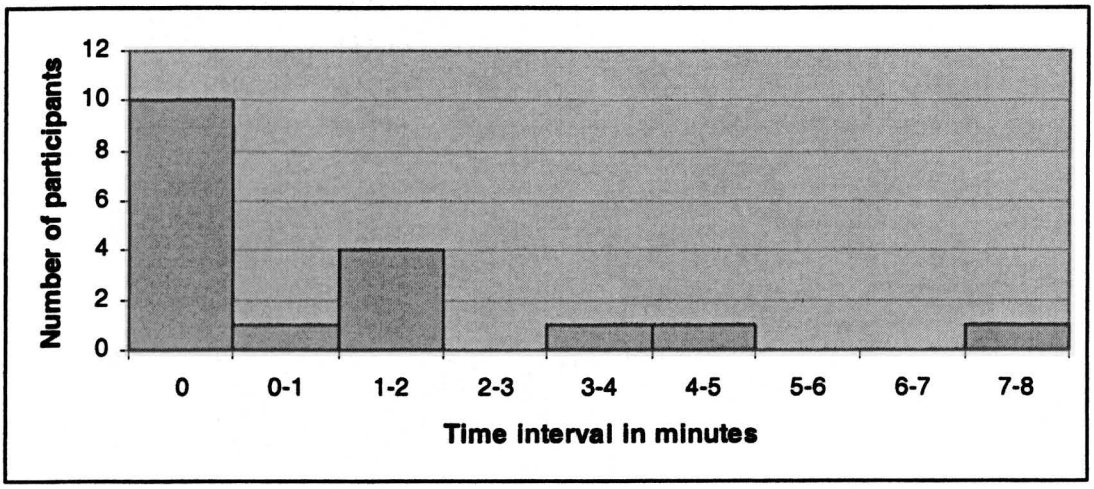



Figure 4.3 Distribution of time on transcript in the activity in one minute intervals (N=18)




It can be seen from both Figure 4.2 and Figure 4.3 that the highest number of students spent less than one minute both on the subtitles pages (7 participants) and the transcript pages (11 participants). Also, the participants used both help options between 1 and 2 minutes very frequently. The participants in the zero category were five students that did not use the subtitles at all (Figure 4.2) and ten that did not use the transcript at all (Figure 4.3).

In case of the subtitles, Figure 4.2 shows that the frequency decreased steadily as time increased. Transcriptions of participants' behavior showed that participants usually tended to watch the whole video segment without fast forwarding or replaying it. Also, an average video segment in the activity was 1minute and 5 seconds long (Table 3.1 in the Methodology chapter). These two could be the possible explanations for this type of time distribution.

In case of the transcript, there is a decline in frequency although not as gradual as in case of the subtitles. It seems that the transcript tended to be used up to two minutes by the majority of participants, with just 3 cases of use in 3-4, 4-5 and 7-8 minute intervals. This distribution could be explained by the fact that *not* a lot of participants used the video when reading the transcript (only 3 students 4 times in total). Furthermore, it was quicker to read the transcript than view the video in real time to find the correct answer. In sum, I believe that the differences between two help options in frequencies of use within one-minute intervals have to do with the nature of reading (quicker) and viewing (slower) in this type of activity. This belief was further supported by some participants in retrospective interviews. When shown his recording, participant 10 commented on using the transcript: "When I use the transcript it make me speed up". Similarly, when talking about the transcript participant 17 remarked "I can get information very quickly".

**Comparison of performance on two help options**

To compare the use of two help options, a paired two-tailed t-test (df =17, t =1.6083) was run with the p set at ≤ 0.05. Comparison of the use of the subtitles and the transcript was done on three variables (help page openings, instances of interaction, and time on help). The results are given in Table 4.4 and statistically significant values are marked with an *. It can be seen that the difference between two help page openings may not be due to a chance ($p<0.0089$) while no statistical significance was found for instances of interaction and time on help.

Table 4.4 Results of t-tests for the use of the subtitles and the transcript

| | Subtitles N=18 | | Transcript N=18 | | p |
|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | |
| Help page openings | 6.88 | 4.95 | 2.17 | 3.68 | 0.00894* |
| Instances of interaction | 2.66 | 3.20 | 1.44 | 2.15 | 0.24493 |
| Time on help | 02:29 | | 01:11 | | 0.14990 |

*Note.* Time is displayed in minutes and seconds.
*p* = probability level
* values significant at p≤ 0.05

To sum up answers to research question 1, the results showed that not all help page openings resulted in help use. Moreover, the participants used subtitles more frequently and for longer periods of time than the transcript. The majority of participants spent more than 2 minutes on help in the course of the whole activity. However, when each of the help options was looked at separately, the majority of participants spent less than 2 minutes on the subtitles and the transcript respectively. The results also indicated that participants greatly varied in the use of help from no use at all to spending up to one-third of total task time on help.

## Types of interaction exhibited by participants in the study

While transcriptions of participants' interaction with the activity were examined in order to answer the first research question, it was noted that there were two types of interaction exhibited by the participants: useful interaction and non-interaction. The distinction between these two types had to be made to get a clearer picture of what happened in the activity and obtain more precise calculations. In addition, the examination of transcriptions further revealed that there are two subtypes of useful interaction exhibited in the activity.

As already mentioned in the Analysis part of the Methodology chapter and in the results of the first research question, some participants in this study did not interact with help at all. Those participants did not play the subtitled video or the video on the transcript page and moved on to the question immediately after the help page loaded. On the other hand, the participants who usefully interacted with help options used help by playing the video and reading the transcript. However, not all the participants using help behaved in the same way. Therefore, two different subtypes of useful interaction were pinpointed:

A. *Meaningful interaction*-The participant opened the help option page and interacted with the help option. For example, the participant read the transcript and played the video that could potentially help him/her get the correct answer.
B. *Non-meaningful interaction*-The participant opened the help option page and interacted with the help option. However, from what the Camtasia videos showed, the strategies that the participant used could not help him/her comprehend better. For

example, the participant played the subtitled video but not the part that could help him/her answer the question. Also, the participant failed to notice the visual clue on the transparency used by the lecturer in the video that could have helped him answer the question.

This distinction between the two types of useful interaction was interesting to me as a researcher because I could, while viewing some Camtasia recordings, clearly see that some strategies that learners exhibited when using help did not aid comprehension. The investigation of unsuccessful learner strategies goes beyond the scope of this study but this type of interaction was worth noting because it may be one of the reasons why learners, although making use of help options in CALL materials, don't get most of help. In sum, this work is generally interested in useful interaction with help options and the results presented in this chapter try to quantify it so that useful interaction can be facilitated in a CALL activity.

## Patterns of interaction with help options

The second research question in this study looked at possible patterns of participants' interaction with help options in the course of the activity. Since the participants could use both the subtitles and the transcript, I investigated whether participants used the same help option all the time or shifted between options. Furthermore, the possible reasons for participants' behavior were also examined. Once four different patterns of interaction were determined, groups following those patterns were compared in terms of comprehension during and after the activity and time and frequency of help use.

### Four patterns of interaction

While the transcriptions of participants' interaction with the activity were coded, it was noticed that some students used only one help option, some used both, and some did not usefully interact with any. Based on the number of help page openings, instances of interaction and time on useful interaction, four patterns of behavior were established:

1. Subtitles pattern
2. Transcript pattern
3. Non-interaction pattern
4. Mixed interaction pattern

Depending on the patterns, the participants were divided into four different groups (the subtitles, the transcript, the non-interaction, and the mixed interaction group). Each of the groups is described in detail below.

The subtitles and the transcript groups

The participants who examined the subtitles pattern of interaction (Table 4.5) used only the subtitles as help during the activity while the transcript group participants (Table 4.6) used only the transcript. Consequently, the participants spent zero minutes and zero seconds on the transcript and the subtitles respectively. Also, the number of incorrectly answered questions is similar to the number of subtitle instances of interaction (the subtitle group) or transcript instances of interaction (the transcript group). This means that the students used help when they had problems understanding the lecture. There were seven students in the subtitles group (participants 2, 4, 11, 12, 15, 16, and 20) and three is the transcript group (participants 6, 9, and 17).

Table 4.5 Participants in the subtitles group

| Participant N=7 | Incorrect answers | Subtitles time | Transcript time | Total help time | Help opening | Instances | Division into groups[a] | Post-listening questions | Recall test |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 5 | 04:49 | 00:00 | 04:49 | 6 | 5 | H | 2 | 5 |
| 4 | 4 | 05:40 | 00:00 | 05:40 | 4 | 4 | L | 1 | 5 |
| 11 | 4 | 06:05 | 00:00 | 06:05 | 5 | 5 | H | 3 | 10 |
| 12 | 5 | 04:11 | 00:00 | 04:11 | 7 | 5 | H | 2 | 6 |
| 15 | 5 | 02:17 | 00:00 | 02:17 | 6 | 4 | L | 3 | 2 |
| 16 | 3 | 01:54 | 00:00 | 01:54 | 5 | 3 | H | 2 | 6 |
| 20 | 3 | 02:08 | 00:00 | 02:08 | 7 | 2 | H | 4 | 9 |
| Total | 29 | 27:04 | 00:00 | 27:04 | 40 | 28 | | 17 | 43 |
| Mean | 4.14 | 03:52 | 00:00 | 03:52 | 5.71 | 4 | | 2.43 | 6.14 |

*Note.* Time is displayed in minutes and seconds
[a]Division into groups: H-higher, L-lower

Table 4.6 Participants in the transcript group

| Participant N=3 | Incorrect answers | Subtitles time | Transcript time | Total help time | Help opening | Instances | Division into groups[a] | Post-listening questions | Recall test |
|---|---|---|---|---|---|---|---|---|---|
| 6 | 2 | 00:00 | 01:16 | 01:16 | 2 | 2 | H | 4 | 6 |
| 9 | 5 | 00:00 | 07:11 | 07:11 | 9 | 7 | L | 1 | 0 |
| 17 | 5 | 00:00 | 03:13 | 03:13 | 8 | 4 | H | 2 | 5 |
| Total | 12 | 00:00 | 11:40 | 11:40 | 19 | 13 | | 7 | 11 |
| Mean | 4 | 00:00 | 03:53 | 03:53 | 6.33 | 4.33 | | 2.33 | 3.66 |

*Note.* Time is displayed in minutes and seconds
[a]Division into groups: H-higher, L-lower

Retrospective interviews and questionnaires identified several possible reasons for this type of behavior which I classified into three categories--level of ease in using help, personal preferences, and previous experience with help. Participants' comments for each of the categories are given in Table 4.7.

Table 4.7 Participants' reasons for using only one help option in the activity

| Category | Participant's comment | Participant |
|---|---|---|
| Level of ease | *"Because it's the easiest way to follow all the words the speaker says"* | 9 |
| | *"I chose subtitles since it was easy for me to find what information I missed to answer a question..."* | 11 |
| Personal preference | *"I don't like transcript because I lost my concentration for listening from reading too much"* | 12 |
| | *"In spite of learning English from the middle school, it's hard to speak and listen English for most [Korean] students. But most students can read and write. Anyway what I want to tell you is that this is the problem for English education in my country. So for me to read sentence is to easy for understanding some story"* | 6 |
| Previous experience with help | *"I watch TV with subtitles"* | 20 |
| | *"I'm accustomed to reading subtitles while listening"* | 11 |

The non-interaction group

This group of participants exhibited very specific behavior in the activity. Data for all students in the non-interaction group are given in Table 4.8. These students opened help pages many times but did not interact with help options. They had 10 or more openings of help pages each but mostly only one instance of interaction. For this group it was not important which help option they opened because the number of meaningful interactions was less than 10% of help openings. This suggests that the participants did not use help although they opened help pages. Opening help was just a way to go on so they answered questions either from memory or by guessing. It is interesting to note that all of the participants in this group belonged to the lower proficiency group. This finding supports the claim by Desmarais, Laurier, and Renie (1998) that navigational patterns may be connected with participants' proficiency level since lower level learners in their study exhibited a more chaotic pattern. Also, it can be speculated that non-interaction students did not comprehend well (7--highest number of incorrect scores per group) but did not use help (0.75--lowest number of instances per group) although they needed it. One possible explanation for this kind of behavior could be that the activity was too difficult for these participants. On the post-listening questionnaire, one of the students remarked that the lecture was too fast for him and that he found the activity very difficult. Also, it is possible that the students were not sufficiently motivated to engage in the task. Task difficulty and lack of motivation were also pinpointed as reasons for this type of a navigational behavior in addition to the absence of learner goals (Desmarais, Laurier, & Renie, 1998).

Table 4.8 Participants in the non-interaction group

| Participant N=4 | Incorrect answer | Subtitles time | Transcript time | Total help time | Help opening | Instances | Division into groups[a] | Post-listening questions | Recall test |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 00:00 | 00:07 | 00:07 | 14 | 1 | L | 3 | NK[b] |
| 3 | 7 | 01:39 | 00:00 | 01:39 | 13 | 1 | L | 2 | NK |
| 8 | 9 | 00:00 | 00:00 | 00:00 | 15 | 0 | L | 1 | 0 |
| 19 | 5 | 01:50 | 00:00 | 01:50 | 10 | 1 | L | 1 | 0 |
| Total | 28 | 03:29 | 00:07 | 03:36 | 52 | 3 | | 7 | 0 |
| Mean | 7 | 00:52 | 00:02 | 00:54 | 13 | 0.75 | | 1.75 | 0 |

*Note.* Time is displayed in minutes and seconds

[a]Division into groups: H-higher, L-lower; [b]NK-Not known

The mixed interaction group

The mixed interaction group consisted of participants who used both the subtitles and the transcript at least once in the activity (Table 4.9). These students switched between two help options either within the same question or in the whole activity. When putting students into this group, it was not important which help option they used the most (in terms of time and instances) as long as they used them both. When compared to other groups, these four students had the highest number of instance of interaction (mean 7.5) in addition to longest total time on help (mean 5 min 55 sec) on average per participant. As a result, it may be possible to say that this group made good use of help.

Table 4.9 Participants in the mixed interaction group

| Participant N=4 | Incorrect answer | Subtitles time | Transcript time | Total help time | Help openings | Instances | S+T instances[a] | Division into groups[b] |
|---|---|---|---|---|---|---|---|---|
| 5 | 3 | 03:40 | 01:55 | 05:35 | 4 | 4 | 3+1 | H |
| 7 | 7 | 00:57 | 01:32 | 02:29 | 22 | 5 | 1+4 | L |
| 10 | 7 | 08:48 | 01:45 | 10:33 | 19 | 15 | 13+2 | H |
| 18 | 5 | 00:45 | 04:18 | 05:03 | 7 | 6 | 1+5 | H |
| Total | 22 | 14:10 | 09:30 | 23:40 | 52 | 30 | | |
| Mean | 5.5 | 03:33 | 02:23 | 05:55 | 13 | 7.5 | | |

Table 4.9 Continued

| Participant N=4 | Post-listening questions | Recall test |
|---|---|---|
| 5 | 2 | 10 |
| 7 | 1 | 0 |
| 10 | 2 | 4 |
| 18 | 3 | 1 |
| Total | 8 | 15 |
| Mean | 2 | 3.75 |

*Note.* Time is displayed in minutes and seconds
[a]S+Tinstances Breakdown of instances into subtitle and transcript instances
[b]Division into groups: H-higher, L-lower

From my point of view this group was very interesting to investigate in order to attempt to find a rationale behind the help option switching. This is the reason why two of the students from this group participated in retrospective interviews. It seems that experimenting was one of the reasons for switching since participant 18 remarked she wanted to have a

choice and see which help option worked better for her. The same reason is behind the comment of participant 10 "I just want to compare them". In addition, it appears that limited amount of time on task influenced participant 10 to speed up and switch to the transcript because he "was the last one and just wanted to finish".

**Comparison of four different pattern groups**

In order to get a better idea about the performance of participants from four different groups, their mean scores were compared across 6 variables (incorrect answers, total time on help, total number of help page openings, total number of instances of interaction, scores on post-listening questions, and the recall test) using a single factor ANOVA with the p value set at the .05 level. The results are presented in Table 4.10 and statistically significant differences between means are marked with an *. Since the groups used in this calculation of variance have small samples, the results need to be interpreted with caution when it comes to generalization to a larger population. However, these results serve as indicators of statistically significant relationships that warrant further investigation on larger samples.

Table 4.10 Results of ANOVA four pattern groups

| | Subtitles group N=7 | Transcript group N=3 | Non-interaction group N=4 | Mixed interaction group N=4 | p |
|---|---|---|---|---|---|
| | Mean | Mean | Mean | Mean | |
| Incorrect answers | 4.14 | 4 | 7 | 5.5 | 0.0318* |
| Total help time | 03:52 | 03:53 | 00:54 | 05:55 | 0.0540 |
| Help page openings | 5.71 | 6.33 | 13 | 13 | 0.0379* |
| Instances of interaction | 4 | 4.33 | 0.75 | 7.5 | 0.0234* |
| Correct answers on post-listening questions | 2.43 | 2.33 | 1.75 | 2 | 0.0740 |
| Delayed recall test | 6.14 | 3.67 | 0 | 3.75 | 0.1630 |

*Note.* Time is displayed in minutes and seconds.
*p* = probability level
*values significant at p≤ 0.05

Although the results of the ANOVA showed that there is statistical significance between means of four groups on some variables, this statistical procedure could not determine which specific group(s) may be different from others on those variables. In order to determine that, a post hoc comparisons procedure--Tukey's honestly significant difference (Tukey HSD) was used. This statistical procedure is more powerful than the Bonferroni method, often used for multiple comparisons in ANOVA, because it adjusts the p value and thus minimizes chances for making a type I error which is possible when performing multiple t-tests. Table 4.11 and 4.12 show the pairs of groups whose differences between means were statistically significant at the .05 level when two-tailed t-tests for all pairs of groups were run.

Table 4.11 Results of t-test for significant variable

| | Non-interaction group | | Subtitles group | | p |
|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | |
| Incorrect answers | 7 | 1.6 | 4.14 | 0.9 | 0.03336* |

*p* = probability level
*values significant at p≤ 0.05

Table 4.12 Results of t-tests for significant variables

| | Non-interaction group | | Mixed interaction group | | p |
|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | |
| Total help time | 00:54 | 00:59 | 05:15 | 3:22 | 0.0355* |
| Instances of interaction | 0.75 | 0.5 | 7.5 | 5.1 | 0.0136* |

*Note.* Time is displayed in minutes and seconds.
*p* = probability level
*values significant at p≤ 0.05

Incorrect answers

The first variable that the groups were compared on was the number of incorrect answers. There were 10 questions offered for the first time in the activity and each incorrect answer lowered the score by one point. Table 4.10 shows that the subtitle and the transcript

groups have almost the same number of incorrect answers on average (4.14 and 4 respectively) while the non-interaction group students performed the weakest in the activity with the highest number of incorrect answers (7). The mixed interaction group performed better than the non-interaction group with a mean score of 5.5. A statistical significance was found between four groups. When the Tukey's post hock comparison was conducted, it was found out that the means of the non-interaction and the subtitles groups were those different at the 0.05 level (Table 4.11).

Total time on help

When looking at the averages of total time participants in each group spent on useful interaction (Table 4.10), it can be seen that the participants in the mixed interaction group spent more time interacting with help than participants in any other group. This difference between the mixed interaction group and the subtitles/transcript group is more than two minutes while it is almost 5 minutes between the mixed and the non-interaction group. Non-interaction group spent less than a minute on average while the subtitles and the transcript groups spent the same amount of time (3 min and 52 or 53 sec). The difference between four groups is not statistically significant but is approaching significance. However, the Tukey's, multiple comparison revealed statistical significance between the non-interaction and the mixed interaction group (Table 4.12).

Help page openings

While the subtitles and the transcript group participants opened just the subtitles and the transcript help pages respectively, the non-interaction and the mixed interaction groups opened both help pages. Again, Table 4.10 shows that the subtitles and the transcript groups have very similar means while the non-interaction and the mixed interaction group have exactly the same means (13). The average difference between four means is statistically significant at the 0.05 level but since the p value was adjusted in the Tukey's comparison, no statistical significance was obtained in this procedure.

Instances of interaction

When looking at the means for the number of instances in Table 4.10, the results reveal that the subtitles and the transcript participants are again similar in their performance (4 and 4.33 respectively). The non-interaction group has the smallest number of instances (0.75) and the mixed interaction group has the highest number of useful interactions on average (7.5). Four means varied significantly at the .05 level and the significance was found between groups with the lowest (non-interaction) and the highest (mixed interaction) means (Table 4.12).

Scores on post-listening questions

Since the maximum score on the post-listening questions was four, Table 4.10 shows that the subtitles group students performed the best (2.43) closely followed by the transcript group students (2.33). The non-interaction group scored the lowest on the post-listening questions with a mean score of 1.75, while the mixed interaction group had an average score of 2. However, there was no statistical significance between groups on this variable.

Recall test scores

The recall test examined whether participants could recall ideas from the Astronomy unit a week after they were exposed to it. The mean score of 6.14 indicates that the subtitles group outperformed the transcript and the mixed interaction group the latter two having very similar means (3.67 and 3.75 respectively). As expected, lack of interaction resulted in poor recall of ideas because not a single student in the non-interaction group could recall any ideas from the lecture. However, as on the post-listening questions there was no statistical significance between the four groups.

Looking at the results of the Tukey's comparison method, it can be concluded that the non-interaction group significantly differed from the mixed interaction and subtitles groups on three variables. Indeed, the non-interaction group interacted minimally with help and performed the weakest of all groups. These results make this group of interest for CALL software designers and applied linguists who should try to create materials that would minimize this type of behavior and encourage useful interaction.

The findings of the second research question show that there are four patterns of participants' behavior and that participants differed in the type and amount of interaction with help options. The reasons for the choice of help options go from personal preferences and previous help use to motivation and experimenting with help.

Statistically significant differences between four groups were found for help openings and interactions with help as well as comprehension on the first 10 questions. The subtitles and the transcript groups appear very similar in their performance regarding time, help opening and useful interaction with help. While for comprehension in the activity, the subtitles and the transcript groups have similar scores, the subtitles group by large outperformed other groups on the recall test. The group that is the most different from the other three is the non-interaction group whose weak comprehension and low learning outcomes can be connected to the lack of help use and generally lack of interaction.

## Performance of participants at different proficiency levels

The third research question examined the performance of participants who were for the purpose of the study divided into two proficiency groups--higher (10 participants) and lower (8 participants). Data for the higher group is given in Table 4.13 and the lower group in Table 4.14.

When examining time on help, the higher group spent almost twice as much time on average interacting with help options. This finding is in contrast with Liou (1997) whose results showed that the ineffective group requested twice as much help as the effective group. In this study, the possible explanation for the overall low time on help for the lower group could be due to the behavior of two participants who did not use help almost at all (participant 1 and 8). Additionally, four out of eight participants in the lower proficiency group belonged to the non-interaction group which means that they interacted with help minimally if at all.

Table 4.13 Performance of the higher proficiency group on the activity

| Participant N=10 | Incorrect answers | Group[a] | Time on subtitles | Time on transcript | Total time | Help openings | Subtitles Help openings | Transcript Help openings | Instances | Sub title instances | Tran script instances | Post listening score | Recall test |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 5 | S | 04:49 | 00:00 | 04:49 | 6 | 6 | 0 | 5 | 5 | 0 | 2 | 5 |
| 5 | 3 | M | 03:40 | 01:55 | 05:35 | 4 | 3 | 1 | 4 | 3 | 1 | 2 | 10 |
| 6 | 2 | T | 00:00 | 01:16 | 01:16 | 2 | 0 | 2 | 2 | 0 | 2 | 4 | 6 |
| 10 | 7 | M | 08:48 | 01:45 | 10:33 | 19 | 17 | 2 | 15 | 13 | 2 | 2 | 4 |
| 11 | 4 | S | 06:05 | 00:00 | 06:05 | 5 | 5 | 0 | 5 | 5 | 0 | 3 | 10 |
| 12 | 5 | S | 04:11 | 00:00 | 04:11 | 7 | 7 | 0 | 5 | 5 | 0 | 2 | 6 |
| 16 | 3 | S | 01:54 | 00:00 | 01:54 | 5 | 5 | 0 | 3 | 3 | 0 | 2 | 6 |
| 17 | 5 | T | 00:00 | 03:13 | 03:13 | 8 | 4 | 4 | 4 | 0 | 4 | 2 | 5 |
| 18 | 5 | M | 00:45 | 04:18 | 05:03 | 7 | 1 | 6 | 6 | 1 | 5 | 3 | 1 |
| 20 | 3 | S | 02:08 | 00:00 | 02:08 | 7 | 7 | 0 | 2 | 2 | 0 | 4 | 9 |
| Total | 42 | | 32:20 | 12:27 | 44:47 | 70 | 55 | 15 | 51 | 37 | 14 | 26 | 62 |
| Mean | 4.20 | | 03:14 | 01:15 | 04:29 | 7.00 | 5.50 | 1.50 | 5.10 | 3.70 | 1.40 | 2.60 | 6.20 |
| SD | 1.48 | | | | | 4.57 | 4.67 | 2.07 | 3.73 | 3.80 | 1.84 | 0.84 | 2.82 |

*Note.* [a]Group (patterns of interaction)
S-subtitles
T-transcript
M-mixed interaction
Time is displayed in minutes and seconds

Table 4.14 Performance of the lower proficiency group on the activity

| Participant N=8 | Incorrect answers | Group[a] | Time on subtitles | Time on transcript | Total time | Help openings | Subtitles Help openings | Transcript Help openings | Instances | Sub title instances | Tran script instances | Post listening score | Recall test |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | NI | 00:00 | 00:07 | 00:07 | 14 | 12 | 2 | 1 | 0 | 1 | 3 | NK[b] |
| 3 | 7 | NI | 01:39 | 00:00 | 01:39 | 13 | 13 | 0 | 1 | 1 | 0 | 2 | NK[b] |
| 4 | 4 | S | 05:40 | 00:00 | 05:40 | 4 | 4 | 0 | 4 | 4 | 0 | 1 | 5 |
| 7 | 7 | M | 00:57 | 01:32 | 02:29 | 22 | 9 | 13 | 5 | 1 | 4 | 1 | 0 |
| 8 | 9 | NI | 00:00 | 00:00 | 00:00 | 15 | 15 | 0 | 0 | 0 | 0 | 1 | 0 |
| 9 | 5 | T | 00:00 | 07:11 | 07:11 | 9 | 0 | 9 | 7 | 0 | 7 | 1 | 0 |
| 15 | 5 | S | 02:17 | 00:00 | 02:17 | 6 | 6 | 0 | 4 | 4 | 0 | 3 | 2 |
| 19 | 5 | NI | 01:50 | 00:00 | 01:50 | 10 | 10 | 0 | 1 | 1 | 0 | 1 | 0 |
| Total | 49 | | 12:23 | 08:50 | 21:13 | 93 | 69 | 24 | 23 | 11 | 12 | 13 | 7 |
| Mean | 6.12 | | 01:33 | 01:06 | 02:39 | 11.63 | 8.63 | 3.00 | 2.88 | 1.37 | 1.5 | 1.62 | 1.17 |
| SD | 1.64 | | | | | 5.68 | 5.01 | 5.10 | 2.47 | 1.69 | 2.62 | 0.92 | 2.04 |

*Note.* [b]NK-not known
[a]Group (patterns of interaction)
S-subtitles
T-transcript
M-mixed interaction
NI-non-interaction
Time is displayed in minutes and seconds

While the higher group spent more time on the subtitles (3 min 14 sec) than the lower group and (1 min 33 sec), both groups had similar average time on the transcript (1 min 15 sec and 1 min 6 sec respectively). The lower group, on average, opened help pages more times than the higher group but the higher group had more useful instances of interaction- another proof that the higher group participants used help options more effectively. When usefully interacting with help, the higher group requested subtitles more than the lower group ( 3.70 and 1.37 respectively) while again the instances of interaction with the transcript do not seem to vary very much (1.4 and 1.5 respectively). Generally speaking, it seems that the higher group made more use of the subtitles according to time and frequency although the average difference between the means was not statistically significant. Interestingly, time and frequency of interaction with the transcript did not vary very much between groups.

The only variables where statistical significance between groups was found were the incorrect answers, the post-listening questions, and the recall test. The results of a two-tailed t-test with two sample equal variance are presented in Table 4.15. Means for each group show that, as expected, the higher group had better comprehension and recall. Obviously, the weaker students did not comprehend the content of the lecture well and did not answer many comprehension questions correctly. Immediately after the lecture their comprehension is even lower, while a week after the activity many of them could not recall any ideas from the activity. The explanations for this may be that the lecture was too difficult for students or that the help options were not employed effectively to compensate for comprehension breakdowns which consequently resulted in low learning outcomes. Finally, as already noted, some of the weaker students did interact with material at all since half of the group exhibited the non-interaction patterns of help use.

Table 4.15 Comprehension and recall data for two groups

| | Higher proficiency group N=10 | | Lower proficiency group N=8 | | p |
|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | |
| Incorrect answers | 4.2 | 1.48 | 6.12 | 1.64 | 0.018682* |
| Post listening score | 2.6 | 0.84 | 1.62 | 0.92 | 0.032143* |
| Recall test | 6.2 | 2.82 | 1.17 | 2.04 | 0.001977* |

*p* = probability level

* values significant at p≤ 0.05

To summarize the results for the third research question, it was found that participants at two different proficiency levels differed in their use of help options. The higher group spent more time usefully interacting with help than the lower group. Also, the participants in the higher group appear to have used the subtitles more frequently and for longer periods than participants in the lower group although there was no statistical significance between groups. However, the use of the transcript in terms of frequency as well as time seemed to be very similar for both of the groups. The difference between groups on comprehension and recall was found to be statistically significant at the 0.5 level.

## Attitudes towards help options

The fourth research question examined participants' attitudes towards the two help options. Questionnaires that the learners completed before and after the activity asked which help option they preferred. Table 4.16 shows the number of students who preferred a certain option before and after the Astronomy unit.

Table 4.16 Participants' preferences of help options

| N=18 | | Before the activity | | Total |
|---|---|---|---|---|
| | | Transcript | Subtitles | |
| After the activity | Transcript | 1 | 4 | 5 |
| | Subtitles | 4 | 9 | 13 |
| Total | | 5 | 13 | |

It can be seen from Table 4.16 that 5 students preferred the transcript and 13 students preferred the subtitles before the activity. After the activity, the ratio between the subtitles and the transcript did not change although only one student kept to the transcript and nine students to the subtitles. This means that the 8 students changed their preferences after they have encountered help--4 changed from transcript to subtitles and 4 from subtitles to transcript. A possible explanation for this shift in preferences could be that the students realized which help option worked better for them. A comment by participant 5 supports this: "I think I switch my answer because I realized that it was easy for me to follow the lecture and understanding the speaking using subtitles (like a close caption on TV)". Overall, the subtitles were preferred by the majority before and after the activity. I can speculate that the

participants chose the type of help they were most exposed to in everyday life. This was supported by answers to question 3 in the pre-listening questionnaire that asked about textual help (transcript or subtitles) students were familiar with either in English or their native language. Ten participants reported they were familiar with the subtitles, 1 with both the subtitles and the transcript, and 7 with none.

On the post-listening questionnaire, the participants also rated the helpfulness of help options. The average rating for the subtitles was 5.5 on the seven-point Likert scale (SD 1.51) which means that the participants found the subtitles helpful. The ratings for the transcript don't give precise results because the question asked students to rate the transcript that they used in conjunction with the video. However, many participants that used the transcript did not play the video so these ratings will not be reported. In the same questionnaire, the participants found the Astronomy unit helpful (mean rating 4.28; SD 1.71) and rated unit difficulty as 4.56 (SD 1.54)--somewhat difficult on the scale where 1 was easy and 7 very difficult. The Astronomy unit appeared to be helpful for improvement of listening skills because, as participants commented, of its organization and good video quality with clear sound.

# CHAPTER 5. CONCLUSION

With the advancement of technology at the beginning of the 21st century, multimedia language learning materials have entered ESL and EFL classrooms and language labs. With more and more language learning software becoming available daily, the questions posed for CALL researchers include how students use the software and how analyses of students' interaction with software can help both CALL software design and CALL pedagogies. To address these questions, this study examined the use of help functions in a multimedia listening activity by looking at student behavior, performance, and student impressions about the activity. Although research has dealt with help options that assist computer delivered listening, not a lot of research has explicitly addressed possible differences between two forms of textual help, subtitles and the transcript, that are offered in cases of comprehension breakdowns, the topic of investigation in this study.

## Major findings

The study identified four different patterns of participants' interaction with help options based on time and frequencies of help option use. The patterns identified were: subtitles, transcript, non-interaction, and mixed interaction pattern and participants were divided into four groups accordingly. The subtitles group participants used only the subtitles during the activity while the transcript group participants only used the transcript. The mixed interaction group switched between the transcript and the subtitles, while the non-interaction group interacted minimally with help options if at all. The analysis showed that the four groups varied in their performance on the activity, help page openings, and instances of useful interaction with help. While the subtitles and the transcript groups had very similar performance on comprehension questions during and after the activity as well as on time and frequencies of help use, the non-interaction group varied the most in behavior and performance from all other groups.

Two proficiency groups (higher and lower) also differed in their performance during and after the activity with the higher group having better comprehension. While for both groups time on the transcript was almost the same, the higher group spent more time on the

subtitles. It appears that the higher group took more advantage of help since the margin between the help page openings and the instances of useful interaction was lower for the higher proficiency group.

When comparing the use of the subtitles and the transcript for all participants, it was found that the participants interacted with the subtitles more frequently and for longer periods of time than with the transcript. Also, it was found that learners exhibited great variation in time on help which ranged from almost no help at all to approximately one-third of the total time on task. A large variation was also noted in frequency of help page openings and instances of useful interaction with help. This finding supports Pujola (2002) and Hegelheimer and Tower (2004) who also found variations in the use of textual help.

Overall, the results show that the participants spent less time interacting with options available in the program than was anticipated when the study was set up. As an illustration, 55% of help page openings did not result in help use which means that the participants opened the help page but did not use the subtitles or the transcript to their advantage. This result corroborates the findings by Liou (1997) and Hegelheimer and Tower (2004) about infrequent use of some help functions. The lack of participants' interaction with help was rather surprising given that the program was designed to facilitate and not discourage the use of help. This may not be a unique case since other sources indicated a similar problem (Desmarais, Laurel, & Renie, 1998). Obviously, "we can not assume students will behave in a particular way when they work within a learning environment" (Hsu, Chapelle & Thompson, 1993, p. 13). This issue warrants further investigation through a more extensive observation and evaluation of software use with lager and more varied user samples.

Some of the possible reasons for the failure of participants to make use of help options offered in the multimedia program could be found in task characteristics such as degree of control and time pressure. It appears that some of the learners did not like the fact that they could not skip help after their comprehension failed and that the only control of help they had was the choice between two help options. Moreover, time on task was limited by one class period of fifty minutes which may have forced some of the participants to finish quicker and did not allow them to use help to the extent they wanted. The time constraint was also noted as one of the explanations for the low level of interaction with the CALL material

in other studies (Hwu, 2003). Additionally, it could be speculated that external factors such as motivation and attitudes towards the task influenced students' behavior as well. For example, not all students may have found a lecture on Astronomy relevant for their field of study. Also, the participants had not encountered this type of activity in their class before, so the lack of experience may have influenced their behavior as well.

## Limitations

If this research was to be performed again, several technical and non-technical limitations should be addressed. The major technical limitation was the fact that Camtasia recording program could not record audio because the participants were using headphones to listen to the lecture. Headphones were needed since the data collection was conducted in a computer lab with 10 students at the time and if the students listened using speakers they would definitely distract each other. If students could work on the activity one at a time or in a different computer lab where using speakers would not be an issue, it would be possible to record external audio using Camtasia. Also, the files Camtasia generated were huge and went up to 660 MB so they were difficult to store and handle. In cases of a longer activity, an activity with more video, or a study with a larger number of participants, file size and handling would become an issue of major concern. Moreover, Camtasia files were not compatible with other media players so playing Camtasia video required Camtasia Studio program installed on the computer or importing the video into a different player which could sometimes decrease the quality of the recording. Finally, transcriptions of Camtasia files were a slow and time-consuming process so using a tracking system that would make logs of student behavior could speed up this part of the data collection in addition to giving more precise and easily analyzable data.

Some non-technical limitations included a small number of participants and the fact that not all students were available for retrospective interviews. The fact that the study was a part of the course but that students did not receive additional credit for participation may have influenced their motivation. Also, the students were all on the intermediate level of listening proficiency so the possible difference between two proficiency groups (higher and lower) may not have been that noticeable as in cases when beginner and advanced learners

are compared. Different results may have been obtained investigating more than one group of students who have been exposed to similar CALL materials as a part of the course curriculum. Lastly, fifty-minute class time limited the period students could spend on the activity so conducting the study in a setting different from the regular class could be explored.

## Implications

Implications of this study regard CALL software design and its use by learners and teacher. The major design implication addresses the balance of learner and program control that would allow the learner, in addition to the choice of the type of help, to choose whether or not to use help after comprehension breakdowns. It seems that studies that allowed more learner control (Borras & Lafayette, 1994) were more successful at keeping students on task (p. 69). Although this study showed that the participants used the subtitles more than the transcript, it seems that offering both help options would be the best solution. It appears to me that the question regarding help is not in the choice between the two types of textual help but in the choice between learners' using help and not using help. I believe that it is necessary to promote the use of help through demos and tutorials within the software. Also, I would argue for giving learners an option to skip help. Although in the research environment of this study skipping help was allowed to control for variables, I think it would be useful to include that option as well as give the choice of both the subtitles and the transcript to accommodate different learning styles and preferences.

On the side of teachers and students, teachers' assistance may be crucial in prompting learners to use help as previous research has suggested (Hegelheimer & Tower, 2004). For example, the teachers could prompt the use of help for both in class (use of lecture transcript, subtitled digital video) and out-of-class listening (subtitled TV program, movies on DVD). Furthermore, teachers can train students to use CALL materials and observe students working on tasks to encourage an effective use of learning strategies.

## Suggestions for further research

In my opinion, there are several possible paths to follow regarding future research on help options in multimedia listening materials. First, if a replication of the present study is to be done, it could include a different setting. The research could compare the use of the software in class, out of class for homework, and for independent study to see whether the amounts, frequencies, and patterns of learners' interaction with the material would be different. Results from different research settings could give more insight into the understanding of learner interaction with CALL materials and pinpoint other possible variables that could influence that interaction. Also, the results could add to a better software evaluation so that more customized programs for particular learners and settings could be made.

Second, data collection instruments could include a tracking system that would complement Camtasia screen recordings and retrospective interviews. The tracking system would automatically record mouse movements on the screen and send data about user behavior into a database. Then, the database could be used to generate reports and manipulate data in other ways. The tracking system would not only make the data collection more efficient, but solve the issues of a larger number of participants and data overflow. Furthermore, it would be interesting to explore another data collection method, think aloud protocols, which could shed more light on the choices learners make during the activity.

Third, future research could also look at a different and larger sample of participants. The use of the program by beginner and advanced students could add to the understanding of user behavior as influenced by the proficiency level. Another important variable, student computer familiarity, could be taken into account when designing future studies.

Finally, academic lectures from different fields could be included to accommodate students from various academic disciplines and make content more authentic. Also, it would be of interest to explore how the level of lecture difficulty influences the use of help.

These are just some of the suggestions for future investigation in this area. No matter which of these paths the future research takes, it will hopefully help CALL researchers and practitioners, CALL software designers, and CALL users get a better understanding of interaction and help use in multimedia listening materials.

# APPENDIX 1. HORTICULTURE LECTURE TRANSCRIPT

*Professor Nonnecke:* Good afternoon. What I'd like to do today is to define horticulture and as a goal throughout the next information, uh, you should be able to understand exactly where the term comes from and what it actually entails as far as the discipline. So with that, I'd like to start with the first overhead, looking at the concept of agriculture. [speaking with the graphic] And agriculture is the production of food, fiber, and shelter that provides benefit to humans and the world. And I think we can understand this in a very broad sense that that is what we would associate with the agricultural sciences. In addition to that, the plant sciences within agriculture, these include three areas. The first is agronomy, which would be the disciplines that are associated with the science of growing and producing grain crops. Forestry is the discipline associated with forest management and forest products such as wood. And then the last one is horticulture, which we are going to learn about today and define.

Horticulture has, um, has been important within the United States and within our country, and has, um, had professional significance throughout the entire, uh, life of our country. And I'd like to provide first a definition that includes what exactly horticulture provides to us as humans. So looking at the next overhead, you can see that horticulture is the science and art of producing nutritional plants that provide sustenance for the human body and many examples of these are listed here, such as fruits, vegetables, nuts, all of the different herbs that we use for culinary reasons as well as medicinal purposes, and then our beverages such as chocolate and coffee and tea. And the last aspect of the, uh, definition would include that it also provides beautiful plants for the human soul. So we have the human body and soul as part of horticulture and you can see these would include the flowering and ornamental plants, such as flowers, trees, shrubs, landscapes and lawns.

With that, I'd like to provide just a little bit of a background with regard to where the term horticulture came about. And the first writings that we have utilizing the term horticulture is from about 1600. And during the 1600s, we see the, uh, the use of the term horticulture and this is originally from Latin and if you look at the overhead, you can see that the Latin term for horticulture is hortus cultura and hortus means garden and cultura is Latin for cultivation. So if we were to put these words together, this would be the culture of garden plants. And you can see at the bottom of the overhead as we move down that the culture of garden plants or the cultivation of crops that are associated with the garden is how we would define horticulture currently. When we think about exactly, um, how this culture of garden crops comes about, um, knowing that the words came about from the 1600s, we need to think about what was happening in the, um, countries where Latin was the written language which would be primarily Europe in those times. So, as we look at what was going on within their lives at that time, it makes perfect sense why horticulture then as it's defined, would include all of the garden crops.

And if we would go to the first slide, what I'd like for you to see is the whole concept of, um, having a village and you can see that surrounding the village, we have, uh, farm lands where the grains of the agronomic crops would be raised. We have forests where the forestry products would be and then within the village or very close to the village where they would

have protection would be where the garden crops would be raised. If we look at the next slide, you can see an example of again another village. This is in a northern European country and you can see that the village is located very close to water source but then in the fields and the woods, um, the agronomic, the grazing of the animals and the forest products would be. In the next slide, you can see that, um, here's an example within a garden, um, where this happens to be a pear tree which is trailing along a house and so all of the crops that typically would have been grown within the village area or within the town would be a horticultural crop today. This would include then the flowers that you see in this slide in the, um, in the window boxes as well as fruit crops, vegetable crops, medicinal plants and so on. In the last slide, you'll notice that, um, this happens to be a monastery and within this monastery, this actually, uh, through the Dark Ages was how much of the horticultural, um, aspects were preserved, both the crops and the practices. So here in this particular, um, area of the monastery, you see a cemetery as well as gardens that are part of that convent and monastery.

If, um, if we think about then, um, ending with the slides, if we just think about exactly, uh, what would be an aspect that could define horticulture compared to the other plant sciences, there are several aspects that really separate horticulture from these particular areas such as forestry and agronomy. And, um, I have an overhead that lists these specifically, and the first one is that the factors that separate horticulture from forestry and agronomy, first of all, horticultural crops require very intensive management, labor and input costs. Um, much of the work that is done within a horticultural crop requires hand labor. Certainly we have, um, mechanization within horticulture, some of you may have had the opportunity to see a video or slides of mechanical harvesting of tomatoes, but in general, we associate the horticultural crops with having very high intensive management, labor, and input.

The second item on the list is that the individual plant in particular, um, is more important compared to large fielder plant that has higher value. In your readings, you'll see the example of a person backing up over a strawberry field compared to a wheat field. And when someone backs a car over one hundred wheat plants, it's of very little value and very little significance, yet our Horticulture Department is regularly asked to come up with value of ten strawberry plants or one shade tree because these have very high value in the landscape as well as for production purposes.

The next one is that we have a much higher return per unit area per unit time. And this relates specifically to the fact that again, the crops have such high value that if we think about a square foot of a glass house or a green house, just the value of what can be produced in that square foot compared to a square foot of a forest, it's very, very different.

The next one is that the plants and plant products are highly perishable and many of the practices that we use are because of this. And if you think about the concept of a cut flower, a cut rose on Valentine's Day and compare that to the value of a corn plant, in you, your just general life, you would know that the value is much different. And so these horticultural crops have great value and as part of that, part of that value comes from the fact that this rose plant is very perishable and does not have a long life after it's been removed from the plant. The last one that I just wanted to provide for you, um, is that plants are used for aesthetic purposes. While forests are very beautiful and agronomic fields are also part of our American culture and human culture, uh, we do have the sole purpose of horticultural crops being

produced and utilized only for aesthetic purposes. And this is different from the other plant sciences.

So just to summarize today, horticulture is the science and art of growing and producing plants for high nutritional purposes such as food crops and also for aesthetic purposes. There are three main disciplines within agricultural sciences that are plant-related; these are agronomy, forestry and horticulture. Horticulture differs from these other areas by having very high cost and input, by having very high returns and high value per plant, by actually having, um, a perishable product that we need to manipulate very carefully, and finally, by being used for aesthetic purposes. So with that, I'd like to open it up for any questions.

*Student 1:* Although the word horticulture was first used in the early 1600s, did horticulture exist before that?

*Professor Nonnecke:* Yes, um, that's an excellent question. Horticulture has been with us since, uh, the beginning of civilization, so 10000 years ago, horticultural crops were domesticated and grown for food purposes and hopefully, they also had some aesthetic value as well. Uh-hm.

*Student 2:* Would breakfast cereals like corn flakes or orange juice be considered part of horticulture?

*Professor Nonnecke:* The, uh, whole concept of horticultural, um, crops and production, um, are very important in students' lifestyles and the breakfast cereals such as corn flakes would not be a horticultural crop. That would be an agronomic crop because it is a grain crop. In comparison though, or in contrast, the orange juice is a horticultural crop. These would be tree fruits that would be grown within a garden normally, and so then they are classified as a horticultural crop.

*Student 3:* In the lecture, you mentioned the, uh, uh, uh that being part of the horticulture, it has to be like a, in a city, right, and also, like, in an intensive care management, then would you say golf courses part of horticulture?

*Professor Nonnecke:* Yes, when some of you complete your additional readings, you'll learn all the different disciplines within horticulture and it's very broad. We often say it goes from apples to zucchini, all the way through the alphabet, and turf grass management which would include golf courses would be part of horticulture. It's an aesthetic use of a plant, in a lawn care, in a sports turf such as football or soccer and also golf courses. Yes?

*Student 4:* What are some of the opportunities for students who graduate with a B.S. in Horticulture?

*Professor Nonnecke:* Students have possibilities ranging from management of a golf course, for example, managing the garden center, managing a nursery, managing a green house, through ownership, um, as well as also working within horticultural sales, horticultural therapy where they use the plants for, um, therapeutic value and through all of the crops that we have within our interior areas such as in malls and doctors' offices, so when you walk around on campus and when you see these other aspects in your life, it would just be part of horticulture. Ok. Thank you.

Retrieved from the Online Materials Development (MIAP) site-What is Horticulture? unit at http://www.tesl.iastate.edu/projects/onlineunits/miap/units/contents_page_fs.htm on February 13, 2005

# APPENDIX 2. HORTICULTURE LECTURE QUESTIONS

**101 Listening** Date_____
**Horticulture lecture**
Name_____

Listen to the lecture and circle the letter in front of the correct answer. Each question has only one correct answer. The questions are arranged in chronological order, that is, the answer to question 1 can be found before the answer to question 2.

1. The lecture covers each of these EXCEPT:
a) the meaning of horticulture
b) the idea of agriculture
c) the definition of plant sciences
d) the management of grain crops X

2. According to the lecturer, all of the following are used for culinary reasons EXCEPT:
a) grain X
b) vegetables
c) nuts
d) herbs

3. All of the following are examples of horticultural products EXCEPT:
a) wood X
b) flowering plants
c) drink products
d) trees

4. Horticulture can be defined as:
a) garden
b) cultivation
c) cultivation of garden crops X
d) culture of gardening

5. In the first two slides, the concept being explained is that of having:
a) villages X
b) farms
c) woods
d) fields

6. Horticulture survived the Dark Ages through planting in:
a) gardens
b) monasteries X
c) convents
d) cemeteries

7. It can be inferred that if you have very little space but a lot of manpower, you should concentrate on:
a) horticulture X
b) agriculture
c) agronomy
d) forestry

8. How long has horticulture been a part of humans' lives?
a) Since the 1600s
b) For 50,000 years
c) For 100,000 years
d) Since the beginning of civilization X

9. All of the following would be considered part of horticulture EXCEPT:
a) football fields
b) orange juice
c) corn flakes X
d) golf courses

10. Students who graduate with a B.S. in horticulture can work in:
a) nurseries X
b) buildings
c) malls
d) doctors' offices

Adapted from Kon (2002) by permission

Question 8 taken from the Online Materials Development (MIAP) site-What is Horticulture? unit at http://www.tesl.iastate.edu/projects/onlineunits/miap/units/contents_page_fs.htm on February 13, 2005

# APPENDIX 3. PRE-LISTENING QUESTIONNAIRE

English 101 Listening
Astronomy unit

Date_____

Pre-listening questionnaire

First name _____
Last name _____
E-mail address _____
Native language _____
Most recent TOEFL score _____
IO_____ PGC_____

Evaluate your listening proficiency (Tick one)
___High beginner
___Low intermediate
___Intermediate
___High intermediate
___Advanced

1. Which of the following can help you understand academic lectures better? Tick all that apply
___Writing down new words during the lecture and looking them up in the dictionary later
___Reading the lesson in the textbook before coming to class
___Following the lecturer by looking at the lecture summary on the handout
___Following the lecturer by looking at the lecture titles in the textbook

There are two most common textual options that can help listening comprehension

| | |
|---|---|
| **Subtitles (close captioning)** | The text of what the speaker is saying appears in real time. The text is synchronized with the speech. |
| **Transcript** | The text of everything what the speaker is saying is shown at once. The text is static. |

2. If you could have a computer based listening activity consisting of a video of an academic lecture and a choice of two textual options to help you when you don't understand something, which would those be? Tick only one
___Subtitles (close captioning)-the text appears at the bottom of the video. When the lecturer does not speak, the text does not show.
___Transcript-the text of the whole lecture appears on the page next to the video. The text stays on the screen all the time but you can use the scroll bar to move it up and down.
___None-I would not need any textual help.

3. Tick the column next to the textual option(s) you are familiar with either in English of in your native language. If you are not familiar with any, leave the column blank.

| | |
|---|---|
| Subtitles (close captioning) | |
| Transcript | |

4. Tick the option that applies to you
___I have seen others (friends for example) **watch TV** with English subtitles but I have never used them myself. (go to question 6)
___ My TV has English subtitles but I don't use them. (go to question 6)
___ My TV has English subtitles and I use them. (go to question 5)

Continues on next page

5. When do you use English subtitles on your **TV**?
___ All the time
___When watching movies
Other (please specify)_____
(go to question 6)

6. How often do you use subtitles (close captioning) when **watching video (VCR) and DVD**? Tick the appropriate column

| | Never | Rarely | Sometimes | Always |
|---|---|---|---|---|
| Video (VCR) | | | | |
| DVD | | | | |

7. The academic lecture that you will be viewing deals with a topic in astronomy. What is your background knowledge in this field? Tick the option that applies to you
___I know a lot about astronomy because it is my field of study.
___I know a lot about astronomy because I am interested in the subject.
___I have some background in astronomy.
___I don't know much about astronomy.
Other (please specify) _____

# APPENDIX 4. POST-LISTENING QUESTIONNAIRE

English 101 Listening
Astronomy unit
Name _____

Date_____
Group_____

Post-listening questionnaire

1. How many questions (out of 10) did you answer **incorrectly** when viewing the video for the first time? If you don't know exactly, put the approximate number.

_____

2. When viewing the video after you answered incorrectly, you had the following choices: Did you use them?

| Choice | Yes | No |
|---|---|---|
| Subtitles (close captioning) | | |
| Transcript | | |
| Dictionary | | |

3. Which of the help options did you choose **the most**? (Tick one)
___Subtitles (close captioning)
___Transcript

Why? _____
_____

4. After you answered incorrectly, did you view some segments with the subtitles and some with the transcript, that is, did you go back and forth between two help options? (Tick one)
___Yes (For example: in segment 1 you chose the transcript, in segment 2 the subtitles, in segment 3 the transcript again…)
___No (For example: you chose the transcript for almost all segments)

What did the choice of the help option depend on?

_____

5. Rate the following using the scale

The Astronomy unit is in my opinion
Very easy 1 2 3 4 5 6 7 Very difficult

The Astronomy unit is, in my opinion, helpful for improving listening comprehension
Not at all 1 2 3 4 5 6 7 a lot

Why?_____

Continues on next page

6. Now rate those that you have used

The subtitles (close captioning) in addition to video were helping me understand the Astronomy lecture:
Not at all 1 2 3 4 5 6 7 a lot

The transcript in addition to video was helping me understand the Astronomy lecture:
Not at all 1 2 3 4 5 6 7 a lot

The dictionary in addition to video was helping me understand the Astronomy lecture:
Not at all 1 2 3 4 5 6 7 a lot

7. If you could have more videos of academic lectures but only one help option with them, which option would that be? (Tick only one)
___Subtitles (close captioning)
___Transcript

8. Did you experience any technical difficulties? What kind?

_____

9. What can be done to improve the Astronomy unit?

_____

10. Other comments that you have for me

_____

_____

Thanks for your cooperation!

# APPENDIX 5. ASTRONOMY LECTURE TRANSCRIPT

Segment 1
Hi. I'm Steve Kawaler. I'm a professor of astronomy here at Iowa State University, and I would like to talk to you today a bit about ... one of the most important questions that I could think that anybody could ask. And that is "are we alone in the universe?" We know that life exists on earth and has existed on earth for three and a half billion years. But it would be nice to know, first of all, is there life anywhere in our solar system besides the earth? And secondly, is there intelligent life anywhere, either in the solar system or the galaxy? Now, we all watch Star Trek, the Star Wars movies and enjoy them tremendously and, boy, wouldn't it be nice if that was really what was going on? But to approach this kind of a question, this very big question, you have to break it down into smaller parts. And, one of the best ways to answer a big question is to break it down into parts that experts can individually chime in on, and then combine all those together to come up with an answer to the big question.
For the question "are we alone in the universe?" this was first done in a systematic way by a fellow named Frank Drake. No, not the navigator, but the astronomer. Uh, at the time he was at Green Bank, uh, West Virginia, at a radio observatory, and he asked this big question, and broke it down. And today I would like to go through the Drake Equation, as it is now known, uh, factor by factor, pretty briefly, in order to give you an idea of where astronomers, biologists, and others are, currently, in their thinking of whether or not there's life elsewhere in the universe.

Segment 2
Now, the Drake Equation has several factors, and I can point them out on this graphic here (Shows diagram). Uh, there are one, two, three, four, five, six, seven factors. The first three factors here summarize our astronomical understanding of aspects of this question. The next two factors are more or less biological, that is, having to do with the evolution of life. And the last two factors are sociological. That has to do with how civilization develops and behaves. Uh, by multiplying all these factors together we get to a number N, which is the number of intelligent and communicating civilizations in our galaxy. Uh, we know that N is at least 1, right? Because here I am talking to you electronically and this is a fairly technological way of, of communicating. This same sort of way of communicating can, can work across the galaxy as well. So, basically, we are looking for people like us. Well, people like us? Things like us, anyway.

Segment 3
So, let's take a look at all these various factors. The first factor is R sub S, that's the rate at which suitable stars form in our galaxy. Uh, the second astronomical factor is the fraction of stars, suitable stars that have planets around them. And the third factor is the number of planets, uh, on which life can develop. We then will move to the biological factors. The fraction of planets on which life can develop that does develop or that do develop life, and then the fraction of these planets, on which life develops, that develops intelligent life. Then, finally we move to the sociological factors, the fraction of those planets on which life and intelligent life develops that develop a civilization that is technologically, uh, able. And then

finally how long that civilization exists. Now, uh, I'm an astronomer so we'll talk a bit about the astronomical factors first. And, in fact we know the values of those astronomical factors better than we know any of the other factors. And the first of those factors is the rate at which stars form.

Segment 4
Now, stars ...for those of you who have limited astronomical knowledge, which may be a few of you – stars like our sun are very common in our galaxy, in our Milky Way Galaxy. Here's a picture of a galaxy that's not too dissimilar from our own, it's the Andromeda galaxy. You can actually see it in a nice fall night in the northern hemisphere as a pink smudge up in the sky. There are one hundred billion stars in this galaxy. That's 10 to the 11 stars. Many of them are very much like our sun. Now, that sounds like a very large number, but most of those stars are billions and billions of years old, and what we really want to know is how many of those stars are forming each year. And, so to get to that, astronomers have to know how old the galaxy is and the rate at which those stars form. And furthermore, how many of those stars are suitable. So, the first of the astronomical factors, R sub S, well, we know that the sun has been around for about five billion years, the star has to live that long, we can ask how many stars like our sun are in our galaxy today. And combining all these together we go from one hundred billion down to two, stars per year. That's two solar type stars, two stars like our sun formed in our galaxy each year. And we know that pretty well.

Segment 5
The next factor, the fraction of stars that have planetary systems, well, this factor was really unknown until about four or five years ago. Since then technological advances have allowed us to find planetary systems around other stars, and pretty much any star we look at, that's like our own sun, has a planet around it. This number is surprisingly large, close to one. Uh, we now know of seventeen solar systems around stars, uh, other stars. Of all the stars that have been looked at, these seventeen have been looked at the closest, and therefore they have revealed their planets to us. So, F sub P, the second factor, about a half, maybe a little more, maybe a little less.
The next astronomical, oh, before I talk about the next astronomical factor I just want to point out that a lot of money is being spent on astronomy to evaluate F sub P a little bit better. Uh, NASA and the NSF and various other federal agencies are putting a lot of money into astronomical research to look for planets around other stars. And they've been, as I said, very successful and hopefully they will be even more successful in the future.

Segment 6
Skipping down now to N sub P, the number of planets are, on which life can develop, that's a little bit more speculative because we need to know not only if the star has a planet but where that planet lies. Uh, it has to lie at the right distance for liquid water to appear, and that's a key ingredient. Uh, it has to lie in what we call the habitable zone. And, in our own solar system that number is two. Uh, if you look at random chance it should be about a half. So N sub P, the number of planets that are suitable, is somewhere between a half and two.

Segment 7
So much for what we know. Now, on to the less, uh certain things, the biological sciences. Not that the biological sciences are any less precise than the mathematical sciences, it's just that on the question of deve ... of the development of life on earth we only have one example: life on earth. And to extend that to other planets is a very difficult thing. So, the biological factors are a little bit less well known. The fraction of planets on which life can develop, that do develop life is probably a tenth to one. Probably not much smaller than that there's some tentative evidence for life on Mars. Uh, there's some... little better evidence for life on earth. As far as intelligent life goes, if you develop life on a planet, given enough time will evolution produce an intelligent species? Well, on the one example on which we know life did develop, intelligent life developed and there are several intelligent species on earth. Uh, so the fraction of planets that develop intelligent life could be one. Could be a lot less, could be a half. We don't know. And if you thought that was bad, let's get to the sociological factors.

Segment 8
Uh, assuming you develop an intelligent species, or two, or three on a planet, what are the chances of that, that species develop a technological civilization that's able to communicate across the galaxy? Any guesses? Uh, my guess is: pretty good. Because, given enough time and the development of intelligence, technology is a natural development in order to prolong the survival of such species. So, maybe that's one—I'm an optimist—one looks good to me. The final factor is L, how long does such a civilization communicate in a way that can be detected across the galaxy? Again, we only have one example, and that's life on earth. Uh, human beings have been communicating through radio and television for many, many years. Television broadcasts, however, are the ones that go as far as any other kind of broadcast that we make. They leave our earth, leave our atmosphere, leave our ionosphere and travel through the galaxy. The first television broadcast that qualifies was, of course, the 1936 Olympics in Munich, uh Munich or Berlin? I don't remember. Munich or Berlin? Nobody knows. The 1936 Olympics in Nazi Germany. Uh, that was about 65 years ago, so we continue to send out signals through the galaxy today. Our longevity is at least 65 years. That could be limited by nuclear war or just more efficient broadcasting. So, if everybody gets cable television, we will be quiet, uh, as far as the galaxy is concerned.

Segment 9
Combining all these factors together, uh, as you can see here, we see that N is equal to the longevity of the civilization in years divided by about 8. So, given my estimates, my somewhat optimistic est...estimates, uh, the number of technical civilizations in our galaxy is about 8, right now. That means, as we speak we are one of a family of about 10 possible communicating intelligent civilizations, in the Milky Way galaxy. Now, clearly these estimates can go up or down as we learn more. They will go up or down as people express their opinions. They will go up or down depending on who's doing it and a nice useful exercise for all of you would be to use your best guesses at these various factors and see what numbers you get. My number is 8, it might be considered optimistic; it might be considered pessimistic. Some say it's as high as a hundred. Others say it's one.

Segment 10
Now, that's our galaxy. Our galaxy is just one of a few more galaxies in the universe. And, as this next image shows, this was obtained by the Hubble space telescope of a supposedly blank piece of sky. And you can see that in this supposedly blank piece of sky, pretty much everywhere you look each smudge here is a galaxy, a distant galaxy in our universe. If you piece together all of these small pieces of the sky, over the entire sky, you can figure out how many galaxies there are in our universe. Remember, each galaxy has 10 to the 11 stars. The number you get? Same number, one hundred billion, 10 to the 11. As far as we know, there are at least a hundred billion galaxies in our universe. 1 galaxy for every star in our Milky Way galaxy. So, if it's one in our galaxy, chances are it's close to one in those 10 to the 11 galaxies as well. And, I don't see a way to escape the conclusion, based on the Drake Equation and based on the huge numbers that we have in astronomy, that we, that not only are we not alone in the universe but the universe is teeming with intelligent life. Thank you. If anyone has any questions, feel free.

Segment in the tutorial
We have now, on the earth, a radio telescope called the Arecibo radio telescope, which is built into the top of a mountain in Puerto Rico. It's three hundred meters across. This is a huge telescope. With that telescope and the receivers, we could pick up TV broadcasts like our own from across the galaxy. So, if someone has one of those things sixty-five light years away, the earth is detectable and they can watch the Olympics of 1936.

# APPENDIX 6. ASTRONOMY LECTURE COMPREHENSION QUESTIONS

Question 1
Professor Kawaler does not ask which of the questions below?
a) Is there intelligent life in the galaxy?
b) How was the Drake equation invented? X
c) Could science fiction movies be reality?
d) Are there intelligent species in our solar system?

Question 2
All of these are true about the number N EXCEPT:
a) it is the result we get when multiplying the seven factors
b) it cannot be less than one
c) it is the number of intelligent civilizations that can communicate
d) it represents the time a civilization needs to develop X

Question 3
Which factor, if any, tells about the development of a civilization that is technologically able?
a) Biological
b) Sociological X
c) Neither
d) Astronomical

Question 4
What does the instructor want the audience to really know about stars like our sun?
a) How dissimilar they are from our sun
b) How many billion there are
c) How many are forming each year X
d) How old they seem to be

Question 5
All of the below statements about the stars that have planetary systems around them have been mentioned in the segment EXCEPT:
a) future research by NASA and other federal agencies depends on available funding X
b) those stars have been discovered with the help of new technology
c) almost all stars similar to our sun have a planet around them
d) there are more than ten solar systems that we know of

Question 6
Planets suitable for life must be located in:
a) the random zone
b) the correct distance zone
c) the liquid water zone
d) the habitable zone X

Question 7
Why are the biological factors less known?
a) Because it is difficult to extend the development of life on earth to other planets X
b) Because more money is spent on astronomical than biological research
c) Because it takes a lot of time for intelligent species to develop
d) Because the biological science is less precise than the mathematical science

Question 8
What kind of broadcasts go out furthest in the universe?
a) Cable television
b) Wireless internet
c) Television X
d) Radio

Question 9
All of the following are estimates about the number of intelligent and technological civilizations mentioned in the lecture EXCEPT:
a) one hundred
b) twelve X
c) one
d) eight

Question 10
Why does the lecturer show the image obtained by the Hubble space telescope?
a) To give an example of the latest technology that astronomers are using
b) To summarize the Drake equation
c) To conclude that there are many galaxies in the universe X
d) To further discuss our solar system

**Post-listening questions**

Question 1
Why does the lecturer talk about the Drake equation?
a) Because the equation can tell us about the existence of life in the universe X
b) Because the equation helps approach the topic from multiple perspectives
c) Because the equation is important to understand astronomy, biology and sociology
d) Because the equation gives accurate numbers necessary in astronomic research

Question 2
Which factors in the Drake equation do we know most about?
a) Biological
b) Astronomical X
c) Cultural
d) Sociological

Question 3
The lecturer mentioned 1936 Olympics in Munich/Berlin for all the following reasons EXCEPT:
a) that was the first radio broadcast X
b) human beings have continued to send out signals in the galaxy since then
c) to give an example of a civilization that communicates
d) in connection to sociological factors

Question 4
It is difficult to estimate the number of technological civilizations in our galaxy because of all of the following EXCEPT:
a) different people have different opinions
b) we can only estimate some Drake equation factors
c) we still do not have enough data
d) the Drake equation has many factors X

# APPENDIX 7. DELAYED RECALL TEST

**English 101 Listening** Date_____

**Delayed recall test** Group_____

Name_____

____Undergraduate student

____Graduate student

Please answer the following questions as detailed as possible

1. What is the main topic of the lecture?

_____

_____

2. What does the lecturer discuss to present the main topic of the lecture?

_____

_____

3. How many factors are there?

_____

4. How can the factors be categorized?

_____

_____

5. Give more details about each factor

_____

_____

_____

_____

6. What does the lecturer conclude?

_____

_____

_____

# APPENDIX 8. DIVISION OF PARTICIPANTS INTO TWO PROFICIENCY GROUPS

Table 8.1 Five criteria for division of participants into higher and lower proficiency group

| Participant | Instructor's opinion | Participant self evaluation | Present course grade | TOEFL score | Initial pre-listening test | Group |
|---|---|---|---|---|---|---|
| 1 | l | 1 | 68 | 563 | 4 | L |
| 2 | h | 4 | 90 | 576 | 7 | H |
| 3 | l | 3 | 67 | 530 | 5 | L |
| 4 | l(m) | 1 | 89 | 615 | 9 | L |
| 5 | h | 4 | 93 | 563 | 9 | H |
| 6 | h | 5 | 92 | 583 | NK | H |
| 7 | l | 1 | 89 | 517 | 8 | L |
| 8 | l | 1 | 92 | 523 | 8 | L |
| 9 | l | 1 | 84 | 537 | 7 | L |
| 10 | h(m) | 2 | 91 | 580 | 6 | H |
| 11 | h(m) | 3 | 85 | NK | 8 | H |
| 12 | h | 3, 2 | 88 | 593 | 7 | H |
| 15 | l | 3, 2 | 68 | 547 | 6 | L |
| 16 | h | 3 | 94 | 560 | 8 | H |
| 17 | h(m) | 1 | 96 | 547 | 5 | H |
| 18 | h | 3 | 91 | 597 | 6 | H |
| 19 | l | 2 | 85 | 500 | 8 | L |
| 20 | h(m) | 1 | 98 | 570 | NK | H |
| Maximum | | | 100 | 677 | 10 | |

| *Note.* | Participant self evaluation: | Instructor's opinion: |
|---|---|---|
| H-higher group | 1- High Beginner | l low |
| L-lower group | 2- Low Intermediate | l(m) low medium |
| NK-not known | 3- Intermediate | h(m) high medium |
| | 4- High Intermediate | h high |
| | 5- Advanced | |

Table 8.1 shows the data for 18 participants in 5 criteria:

1) instructor's opinion of participants' listening ability,
2) participants' self-evaluation of listening proficiency,
3) participants' grade on the course at the time of data collection,
4) participants' self reported TOEFL scores,
5) scores on the initial listening comprehension test

The first criteria that was used for the division into two proficiency groups, the TOEFL score, is often used as a source for obtaining information about learners' language knowledge although it may not prove to be "the best method" (Kon, 2002, p. 73). Since the participants' scores were self-reported and the score for the listening comprehension section was not reported separately, other sources needed to be consulted as well. The course instructor was asked to rate students' listening comprehension as high or low (within that group) according to her impression of students' performance in class. Also, the participants were asked to evaluate their listening proficiency on the pre-listening questionnaire because self-assessment of second language abilities is thought to provide acceptable results in comparison to more objective assessment measures (Heilenman, 1990, p. 174). Furthermore, participants' grade on the course at the time of data collection was obtained from the instructor. This grade consisted of homework assignment scores expressed in percents. Finally, the results of the listening comprehension test containing an academic lecture similar to the one used in the study were also taken into account.

When all the criteria were looked at, it was clear that some students would fall into the higher (for example participants 5 and 6) and some into the lower group (participants 9 and 19) according to 4 criteria. Some students, however, had one high score/evaluation that did not match other low criteria. For example, the present grade on the course was very high for some students who were weak according to other criteria (for example participant 8). This was probably due to the fact that the grade was calculated based only on homework assignments so students may have received help from others and may have listened a number of times/looked at transcripts when answering comprehension questions. Also, the score that participants got on the diagnostic listening test did not differentiate well between the higher and lower group since 72% of the students fell into the middle category. Moreover, the students may have copied answers from each other due to the fact that the activity was done as a regular class exercise so test scores were not used as the determining criterion for division (although that was previously thought when the study was set up). Instead, it was decided that at least 3 criteria per student have to match the same group. In cases when that was not possible (participants 4, 17, and 20), the instructor's opinion was used as the determining criterion. Finally, the higher group of 10 and the lower of 8 students were compiled.

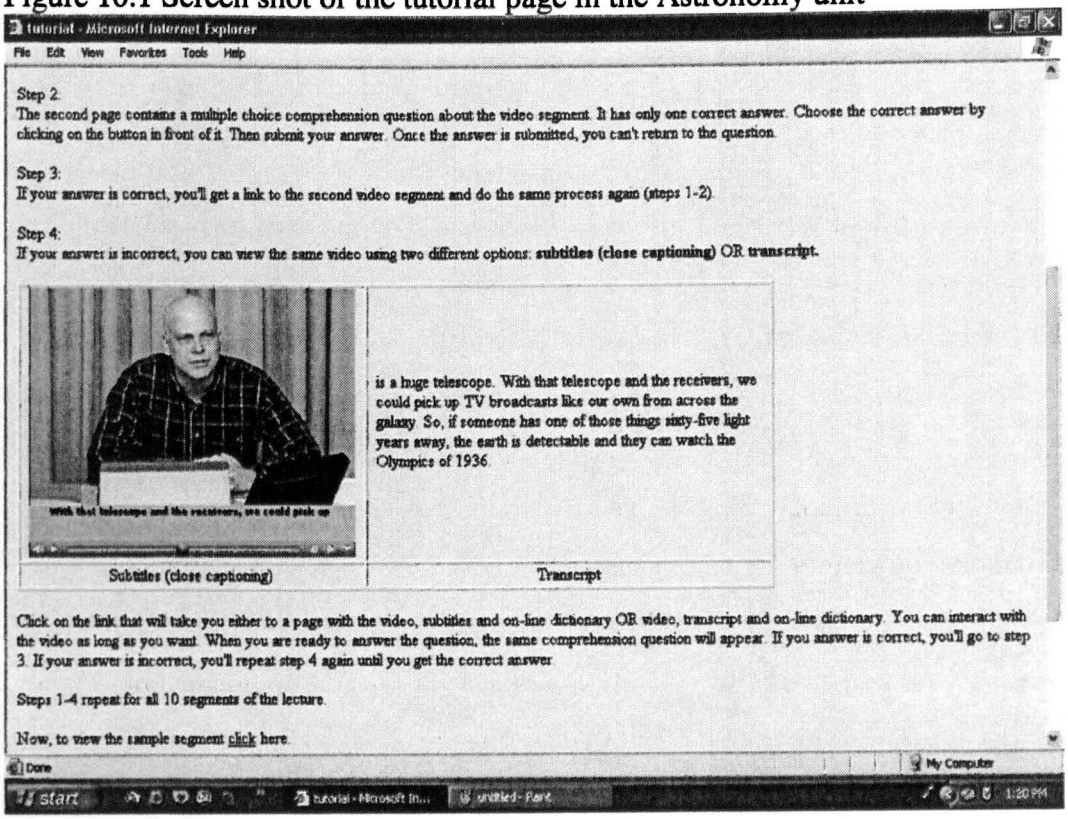
# APPENDIX 9. PILOT STUDIES

The first pilot test was done during one class period in December 2003 when 101 Listening students were shown the Astronomy lecture on the class video projector and instructed to answer 19 comprehension questions. The lecture was divided into 11 segments and the video was stopped after each segment so that the students could answer question(s) dealing with that segment. This viewing procedure was chosen because it resembled the way participants in the study would view the lecture. Out of nineteen questions, eleven questions were multiple choice and the other eight open-ended type. After the statistical analysis, the items that were too difficult or too easy were excluded and some distractors changed. The most common incorrect answers in the open-ended questions were used as distractors when those questions were turned into the multiple choice type questions. After the pilot, a revised version of the activity contained fourteen questions.

The second pilot study was done in Spring 2004 with five ESL students who volunteered to test the materials. Each student viewed the activity on the computer and answered comprehension questions while the screen capturing program-Camtasia recorded progression through the activity. The students also filled two questionnaires which asked them about the familiarity and use of subtitles and transcript and the design of the unit. They were asked to comment on the clarity of instructions and possible technical difficulties as well as to suggest improvements. Finally, they were asked to evaluate the usefulness and effectiveness of the activity. The students' comments showed that the instructions did not need to appear on every page in the Astronomy unit but only in segment 1 and in the tutorial. Also, some questions about the use of subtitles were changed because students had not found them clear enough. The students finished the activity in about 25-30 minutes which was the desired timing given the fact that students in the data collection needed another 15 minutes for the post-listening questionnaire.

# APPENDIX 10. SCREEN SHOTS OF THE ASTRONOMY UNIT PAGES

Figure 10.1 Screen shot of the tutorial page in the Astronomy unit




m

Figure 10.2 Screen shot of a comprehension question page in the Astronomy unit

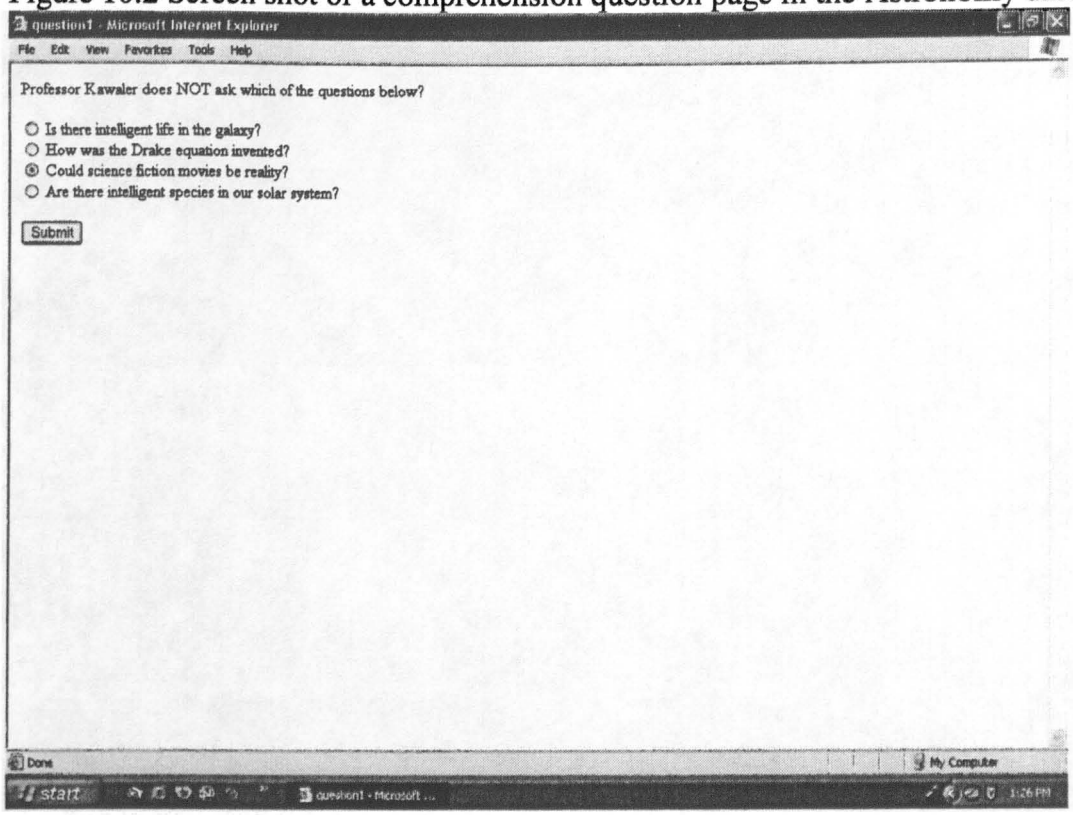
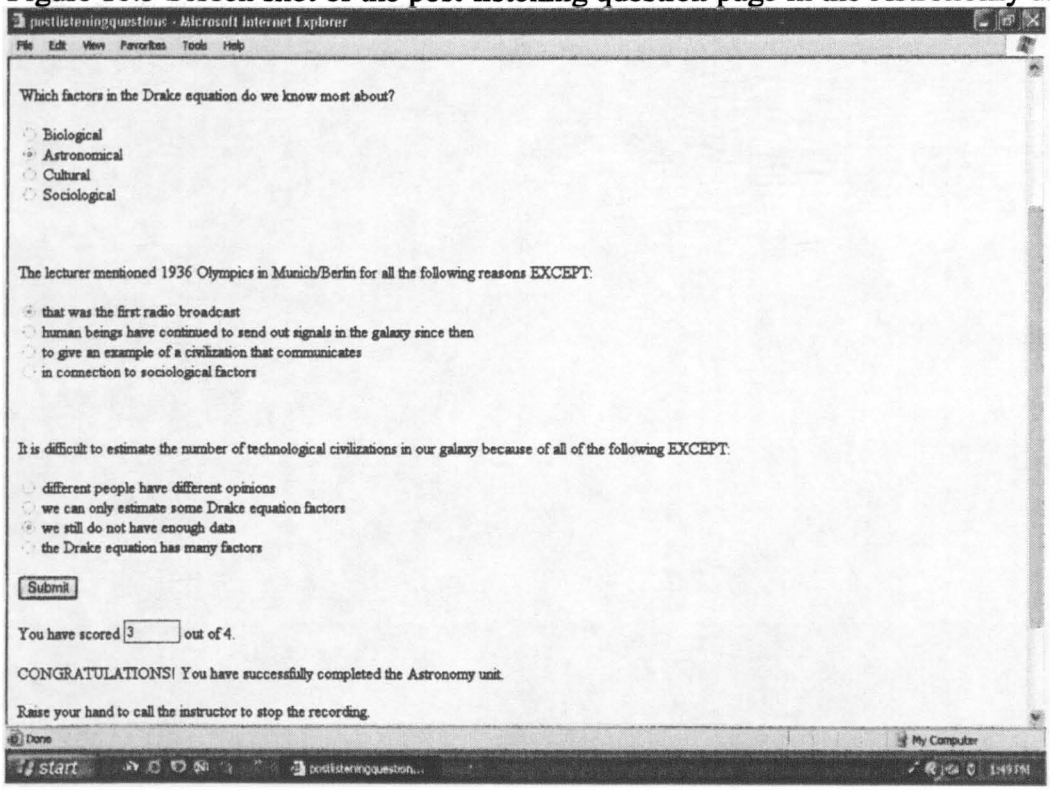



**Figure 10.3 Screen shot of the post-listening question page in the Astronomy unit**



# APPENDIX 11. RESULTS

**Time on task**

Table 11.1 Time on the activity (the Astronomy lecture and 10 comprehension questions)

| Participant | Time |
|---|---|
| 1 | 20:24 |
| 2 | 22:44 |
| 3 | 17:59 |
| 4 | 22:27 |
| 5 | 25:41 |
| 6 | 20:33 |
| 7 | 24:17 |
| 8 | 26:38 |
| 9 | 28:39 |
| 10 | 32:38 |
| 11 | 26:02 |
| 12 | 22:08 |
| 15 | 19:08 |
| 16 | 18:49 |
| 17 | 23:23 |
| 18 | 23:19 |
| 19 | 24:35 |
| 20 | 23:29 |
| Total | 07:02:53 |
| Mean | 23:30 |

*Note*. Time is displayed in hours, minutes and seconds.

# REFERENCES

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

Brett, P. (1995). Multimedia for listening comprehension: the design of a multimedia-based resource for developing listening skills. *System*, 23 (1), 77-85.

Brett, P. (1996). Using multimedia: an investigation of learners' attitudes. *Computer Assisted Language Learning*, 9 (2-3), 191-212.

Brett, P. (1997). A comparative study of the effects of the use of multimedia on listening comprehension. *System*, 25 (1), 39-53.

Brett, P. (2000). Developing cross-cultural competence in business through multimedia courseware. *RECALL*, 12 (2) 196-208.

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

Chapelle, C. A. (2001). *Computer applications in second language acquisition.* Cambridge: Cambridge University Press.

Chapelle, C. A. (2003). *English language learning and technology: Lectures on applied linguistics in the age of information and communication technology.* Amsterdam: John Benjamins.

Chun, D. M., & Plass, J. L. (1997). Research on text comprehension in multimedia environments. *Language Learning and Technology*, 1 (1), 1-35.

Desmarais, L., Duquette, L., Renie, D., & Laurier, M. (1998). Evaluating learning interactions in a multimedia environment. *Computers and the Humanities*, 22, 1-23.

Desmarais, L., Laurier, M., & Renie, D. (1998) The analysis of navigation patterns in CALL. *Computer Assisted Language Learning*, 11 (3), 309-315.

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

Gass, S. (1997). *Input, interaction and the second language learner.* Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Guillory, H.G. (1999). The effect of keyword captions to authentic French video on learner comprehension. *CALICO Journal*, 15(1-3), 89-108.

Hegelheimer, V., & Tower, D. (2004). Using CALL in the classroom: analyzing student interactions in an authentic classroom. *System*, 32, 185-205.

Heilenman, K. (1990). Self-assessment of second language ability: the role of response effects. *Language Testing,* 7(2), 174-201.

Herron, C., Secules, T., Morris, M., & Curtis, L. (1995). A comparative study of the effects of video versus text-centered instruction in the foreign language curriculum. *The French Review*, 68, 775-796.

Hsu, J. (1994). Computer assisted language learning (CALL): the effect of ESL students' use of interactional modifications on listening comprehension. Unpublished doctoral dissertation, Department of Curriculum and Instruction, College of Education, Iowa State University, Ames, IA.

Hsu, J., Chapelle, C. A., & Thompson, A. (1993). Exploratory environments: what are they and do students explore? *Journal of Educational Computing Research,* 9 (1), 1-15.

Hwu, F. (2003). Learners' behaviors in computer-based input activities elicited through tracking technologies. *Computer Assisted Language Learning*, 16 (1), 5-29.

Keng-Soon, S. (1999). Theory and research: learning styles, motivation, and the CALL classroom. In J. Egbert, & E. Hanson-Smith, (Eds.), *CALL environments-research, practice and critical issues* (pp. 289-301). Alexandria, VA: TESOL.

Kon, C. K. (2002). The influence on outcomes of ESL students' performance strategies on a CALL listening comprehension activity. Unpublished MA thesis, Department of English, Iowa State University, Ames, IA.

Krashen, S. (1982). *Principles and practice in second language acquisition.* Oxford: Pergamon.

Liou, H.C. (1997). Research on on-line help as learner strategies for multimedia CALL evaluation. *CALICO Journal,* 14 (2-4), 81-96.

Liou, H.C. (2000). Assessing learner strategies using computers: new insights and limitations. *Computer Assisted Language Learning*, 13 (1), 65-78.

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

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

Pica, T. (1994). Research on negotiation: What does it reveal about second-language learning conditions, processes, and outcomes? *Language Learning*, 44(3), 493-527.

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.

Pujolà, J.-T. (2002). CALLing for help: Researching language learning strategies using help facilities in a web-based multimedia program. *ReCALL*, 14 (2), 235-262.

Pusack, J. P., & Otto, S. K., (1997). Taking control of multimedia. In M. D. Bush, & R. M. Terry, (Eds.), *Technology-enhanced language learning.* (pp. 1-46). Lincolnwood, IL: National Textbook Company.

Rubin, J. (1990). How learner strategies can inform language teaching. In V. Bickley (Ed.), *Language use, language teaching, and curriculum* (pp. 270-284). Hong Kong: Hong Kong Institute of Language in Education, Education Department.

Rubin, J. (1995). The contribution of video to the development of competence in listening. In D. J. Mendelsohn, & J. Rubin, (Eds.), *A guide for teaching of second language listening.* (pp. 151-165). San Diego: Dominie Press.

Rubin, J., Quinn, J., & Enos, J. (1988). Improving foreign language listening comprehension. *Report on Project 017AH70028,* (Sept.) Washington, DC: International Research and Studies Program, U.S. Department of Education.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11 (2), 129-158.

Secules, T., Herron, C., & Tomasello, M. (1992). The effect of video context on foreign language learning. *The Modern Language Journal,* 76 (4), 480-490.

Thompson, I., & Rubin, J. (1993). Improving listening comprehension in Russian. *Report on Project 017A00032.* International Research and Studies Program, U.S. Department of Education.

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

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

# ACKNOWLEDGEMENTS

First and foremost I would like to thank my husband Zoran without whose love, support, and help, the Astronomy unit and this thesis would never have been possible. I'm truly blessed for having him by my side as my husband and friend. Also, I would like to thank my family, my mother Slobodanka and my brother Ivan, whose love and encouragement meant a lot while in pursuit of this professional dream of mine.

The person whose help proved invaluable for this study is my major professor Dr. Volker Hegelheimer and I'm most grateful for his assistance, advice, and guidance. He was the one who sparked my interest in computer-assisted language learning and I feel privileged I could work and learn from him. I also want to acknowledge statistician Hadley Wickham whose expertise helped me make the most of my data.

I'm in debt to my committee members, Dr. Carol Chapelle and Dr. Denise Schmidt, whose comments and suggestions were very helpful and were most appreciated. Also, many thanks to Ketty Reppert and Felicity Douglas for their help with polishing up the Astronomy unit comprehension questions. I'd also like to thank Cynthia Myers for her feedback on materials, help during the data collection, and letting me use her class. Finally, I thank the students who participated in the study.