

To the Graduate Council:

I am submitting herewith a dissertation written by Victoria B. Savchenkova entitled "Evaluation Criteria For Language Teaching Software: A Delphi Study Of Professional Perspectives." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Education.

Dr. Patricia Davis-Wiley

Major Professor

We have read this dissertation
and recommend its acceptance:

Dr. Edward Counts

Dr. John Ray

Dr. John Romeiser

Acceptance for the Council:

Dr. Anne Mayhew

Vice Provost and Dean of
Graduate Studies

(Original signatures are on file with official student records.)

**EVALUATION CRITERIA FOR LANGUAGE TEACHING SOFTWARE:
A DELPHI STUDY OF PROFESSIONAL PERSPECTIVES**

A Dissertation
Presented for the
Doctor of Philosophy Degree

The University of Tennessee, Knoxville

Victoria B. Savchenkova
August 2003

Copyright © Victoria Savchenkova, 2003

All Rights Reserved

DEDICATION

My dissertation is dedicated to my parents, Zoya and Boris Savchenkov, whose love and understanding have always been my support and my inspiration.

ACKNOWLEDGEMENTS

This dissertation is the final step in my doctoral career at The University of Tennessee and I would like to extend my sincere appreciation to the people who played an important role in this process. First and foremost, I would like to express my sincere gratitude to my chairperson, Dr. Patricia Davis-Wiley, whose thoughtful help, guidance and advice carried me through the years of my doctoral program, culminating in its last and most important stage of research and preparation of this manuscript. Her support and assistance and, most importantly, her willingness to ‘walk the extra mile’ in helping me with the research and writing were invaluable to me. I also would like to express my most heartfelt gratitude to the members of my committee: Dr. Edward Counts, for everything he taught me about multimedia design and computer animation; Dr. John Ray, for his advice and assistance with data collection and analysis; and, Dr. John Romeiser, for his invaluable advice and expertise in the field. Each committee member provided help and support in different areas critical in this study, and I am indeed very grateful for all of their time and selfless effort. Also, I would like to thank the 19 panelists, the experts in this Delphi study, who generously gave their time and expertise in the field of instructional technology. I am extremely grateful for all of their effort and significant professional contributions which brought this study to life. My sincere thanks are also extended to my near and dear friends and family, whose care and support made a great difference during these years of study and research.

ABSTRACT

Rapid development of computer technology in recent years has presented broad opportunities for its use in various fields, including the field of language teaching. Various materials have been designed to assist language teachers and learners with their goals. At present, a number of tools, including a vast variety of language teaching software, is readily available on the market. However, limited research has been conducted on identifying evaluation criteria for language teaching software from multiple perspectives. This study has undertaken to identify and reach consensus on the evaluation criteria for language teaching software essential for the software design and selection. Specifically, this study sought to answer the following two research questions:

1. What criteria for evaluation of language teaching software are identified in currently published literature, including textbooks, software reviews, and Instructional Technology and Computer-Assisted Language Learning research?
2. Which criteria are deemed important by the representatives of the Board and Council of the International Association of Language Learning Technology?

A three-phase research design was employed to carry out the purpose of the study and to provide answers to the research questions. First, published sources for existent language teaching evaluation criteria were identified. Second, a joint list of criteria was compiled from a variety of sources. Then, the criteria were organized into five thematic groups and eight subgroups. In order to validate and reach consensus on the identified criteria, a modified Delphi survey was conducted. Two

rounds of the Delphi method were implemented and produced the final list of evaluation criteria for the selection and design of language teaching software.

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION.....	1
Introduction.....	1
Statement of the Problem.....	3
Purpose of the Study.....	4
Research Questions.....	4
Significance of the Study.....	4
Delimitation of the Study.....	5
Limitations of the Study.....	5
Assumptions.....	5
Definitions of Terms.....	6
Organization of the Study.....	8
II. REVIEW OF THE LITERATURE.....	10
Introduction.....	10
Historical Perspective.....	11
CALL in the 60s – Early 70s.....	12
Late 70s – 80s.....	14
90s – Present.....	16
CALL at Present.....	18
Classifications.....	18
Reports of Practical Applications.....	18
Evaluation.....	22
The Perspective of Second Language Acquisition.....	23
The Perspective of Pedagogy.....	27
The Perspective of Instructional Technology.....	29
The Perspective of Socio-Cultural Studies.....	31
Chapter Summary.....	39
III. METHODS AND PROCEDURES.....	40
Introduction.....	40
Research Questions.....	41
Research Design.....	42
Phase One: Identifying Sources of Evaluation Criteria.....	42
Phase Two: Organization of Data.....	43
Phase Three Conducting a Modified Delphi Study.....	44
The Delphi Method.....	44
Reasons for Selection of the Delphi Method for Present Study.....	46
Selection of the Subjects.....	47
The Instrument.....	48
Procedures.....	49
Pilot Study.....	50
Data Collection.....	51

Analysis of Data.....	53
Chapter Summary.....	56
IV. ANALYSIS OF THE DATA.....	58
Introduction.....	58
Results of Phases One and Two.....	59
Selection of Sources.....	59
Developing the List of Criteria.....	60
Results of the Research Study Phase Three, Delphi Study.....	61
Round One.....	62
Round Two.....	64
Discussion of the Findings.....	84
Chapter Summary.....	90
V. SUMMARIES, CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS.....	91
Introduction.....	91
Summary.....	91
Conclusions.....	99
Implications.....	99
Recommendations.....	100
Chapter Summary.....	102
REFERENCES.....	103
APPENDICES.....	110
APPENDIX A. Round One: Letter to the Panel of Experts.....	111
APPENDIX B. Round Two: Letter to the Panel of Experts.....	114
APPENDIX C. Round One Survey Instrument.....	116
APPENDIX D. Round Two Survey Instrument.....	130
VITA.....	144

LIST OF TABLES

Table	Page
Table 1. Number Of Criteria Added and Eliminated After Round One In Each Group And Subgroup.....	55
Table 2. Criteria Eliminated After Round One.....	65
Table 3. Results from the Delphi Round Two for the Competencies in Group One: Pedagogical Aspect, Subgroup One: Conditions for Quality Comprehensible Input.....	69
Table 4. Results from the Delphi Round Two for the Competencies in Group One: Pedagogical Aspect, Subgroup Two: Conditions for Quality Comprehensible Output.....	70
Table 5. Results from the Delphi Round Two for the Competencies in Group One: Pedagogical Aspect, Subgroup Three: Application Of The Principles Of Cognition.....	72
Table 6. Results from the Delphi Round Two for the Competencies in Group One: Pedagogical Aspect, Subgroup Four: Motivation.....	72
Table 7. Results from the Delphi Round Two for the Competencies in Group One: Pedagogical Aspect, Subgroup Five: Presentation Of The Content.....	74
Table 8. Results from the Delphi Round Two for the Competencies in Group Two: Applicability Aspect, Subgroup One: Learner Perspective.....	75
Table 9. Results from the Delphi Round Two for the Competencies in Group Two: Applicability Aspect, Subgroup Two: Teacher Management.....	76
Table 10. Results from the Delphi Round Two for the Competencies in Group Two: Applicability Aspect, Subgroup Three: General Perspective.....	77
Table 11. Results from the Delphi Round Two for the Competencies in Group Three: Use Of Media Aspect.....	79
Table 12. Results from the Delphi Round Two for the Competencies in Group Four: Socio-Cultural Aspect.....	80
Table 13. Results from the Delphi Round Two for the Competencies in Group Five: Technical Aspect.....	82

Table 14. Comparison of Mean Scores Between the Groups (the order of group listing is based on their mean ratings for Round Two - from highest to lowest).....	83
--	----

Table 15. The List of Evaluation Criteria for Language Teaching Software Considered Essential by the Panel of Experts of IALLT Board and Council.....	85
---	----

CHAPTER I

INTRODUCTION

Introduction

Over the years, the field of Computer-Assisted Language Learning (CALL) has made exceptional progress in the development of its technology and methodology. Starting in the early 60s, as part of a military program, with very limited resources available to a selected few, CALL has become a common practice of language teaching and learning. Moreover, from being considered just a tool for language instruction, CALL has evolved into an independent area of knowledge and a successful pedagogy of its own.

There are certain features that make the area of CALL different from other academic disciplines. The rapid changes in the technological world have brought numerous updates into CALL practice. From one perspective, this phenomenon can be considered as positive since it extends the possibilities of learning a language, and making it less tedious and more effective to do so. The whole process of CALL development is an illustration of the following statements. From the mainframe to personal computers, later to the Internet, virtual reality, and voice recognition, computers have provided new avenues of opportunities and challenges for both language teachers and their students.

At the same time, the same changes have caused some feelings of instability and confusion in the field. New applications have been advancing forward so fast, and new methods have been developed at an equal rate that it has become a challenge for language instructors to follow all the latest upgrades in the field. Swift

technological progress has affected theory and practice of Computer-Assisted Language Learning. As a result of this process, there has emerged a need for a new paradigm that would incorporate the theories from various fields of knowledge related to CALL. However, reaching these objectives has met certain challenges.

Ehsani and Knodt (1999) comment that,

With recent advances in multimedia technology, computer-aided language learning (CALL) has emerged as a tempting alternative to traditional modes of supplementing or replacing direct student-teacher interaction, such as the language laboratory or audio-tape-based self-study. The integration of sound, voice interaction, text, video, and animation has made it possible to create self-paced interactive learning environments that promise to enhance the classroom model of language learning significantly. A growing number of textbook publishers now offer educational software of some sort, and educators can choose among a large variety of different products. Yet, the practical impact of CALL in the field of foreign language education has been rather modest. (p. 46)

The reasons behind theoretical limitations may lie in the following tendencies. First, the instructional design of CALL presents a challenge in including all the perspectives into a new paradigm. Second, there is a disagreement about applicability of some SLA theories for instructional design of CALL activities. Finally, much of the existing research focuses on specific issues rather than providing a broad theoretical perspective for the field of Computer-Assisted Language Learning.

Therefore, realizing this need, an effort is being made towards establishing a theoretical foundation for CALL (Chapell, 1997; Levy, 1997). One of the first steps in establishing a sound theoretical framework for the discipline may be identifying common evaluation criteria for CALL materials. This objective is also well justified from a practical perspective. Addressing the increasing number of CALL materials on the market, for example, Cameron (1999) stated, "In many cases, however, what we have is an excellent teaching resource but not one which has been designed for learning" (p. 3). With the growing availability of resources, it has become extremely challenging to make the right choice of CALL materials, which could make language teaching/learning more effective than it currently is. The purpose of this study, then, is to generate a list of evaluation criteria for language teaching software by incorporating perspectives from several fields, thus, providing a professional perspective on the importance of each of the criteria.

Statement of the Problem

The need has been expressed in the literature for reaching a common approach in formulating a common paradigm of Computer-Assisted Language Learning in general, and finding a common ground for evaluating CALL materials in particular (Chapell, 1997; Levy, 1997). To date, a unanimous approach has not been adopted to the evaluation of various CALL tools, including language teaching software. Therefore, there is a clear need to reach consensus about evaluation criteria for CALL software, since it is one of the most common tools used by instructors and learners in the field.

Purpose of the Study

The primary purpose of this study was to contribute to the knowledge and practice of effective evaluation of Computer-Assisted Language Learning resources for language instructional purposes. In particular, this study endeavored to gather CALL software evaluation criteria that would incorporate multiple perspectives of CALL-related fields (i.e., Second Language Acquisition (SLA), Instructional Technology (IT), pedagogy), and, therefore, potentially assist language teachers and learners in their search for appropriate software to fit their language instructional goals.

Research Questions

This study will examine the following research questions:

1. What criteria for evaluation of language teaching software are identified in currently published literature, including textbooks, software reviews, and Instructional Technology and Computer-Assisted Language Learning research?
2. Which criteria are deemed important by the representatives of the Board and Council of the International Association of Language Learning Technology?

Significance of the Study

The present research will assist language instructors and students who are in search of effective CALL materials. As a result of this study, language teachers and students will have guidelines, which will help them make the right choice in selecting appropriate materials from an increasingly large market of educational software. Secondly, it is anticipated that this study will benefit instructional designers who develop educational software for language teaching and learning. The results of this

research may help them create CALL software, which is not only visually appealing and equipped with the most current applications, but more importantly, follows the pedagogical principles that will make it effective for teaching and learning. Finally, the results of the study will make a small contribution to an ongoing effort of CALL researchers in synthesizing the theories of multiple disciplines to establish a common, multidimensional theory.

Delimitation of the Study

The study has the following delimitation: the population for the Delphi study will be limited to the opinion of chosen representatives of the Board and Council of the International Association of Language Learning Technology (IALLT).

Limitations of the Study

This study has two limitations, which include the following:

1. The sample will be limited to the criteria identified by the editors and research.
2. The study will be limited due to the sampling of the opinion of selected academic experts belonging to the same professional society.

Assumptions

Several assumptions were made in this study. The fundamental assumption was made that selected participants, members of the Board and Council of International Association for Language Learning Technology, representing those who provide leadership and assistance in technology use for the teachers of second/world language in higher education, can be considered the experts in Computer-Assisted Language Learning in general, and language teaching software evaluation in

particular. It was also assumed that the experts would be able and willing to provide valuable insights into the process of the selection of the right criteria for evaluation of language teaching software. Moreover, it was assumed that they would interpret the survey questions correctly and provide their honest responses. Thus, the assumption was made that their answers would reflect their knowledge and beliefs and provide accurate information about the evaluation criteria for language teaching software. It was also one of the assumptions of the study that developing criteria for language teaching software evaluation ought to be one of the objectives of Computer-Assisted Language Learning in order to assist teachers, learners, and instructional designers in the field in their selection of the materials and methods, which would better accommodate their goals. Finally, the assumption was made that material evaluation plays a vital role in the process of teaching/learning a language with technology assistance.

Definitions of Terms

The definitions of the terms used in this study appear below. Some definitions are drawn from cited sources, while others are based on the researcher's knowledge of the field.

Computer-Assisted Language Learning (CALL) is,

the search for and study of applications of the computer in language teaching and learning.... The subject is interdisciplinary in nature, and it has evolved out of early efforts to find ways of using the computer for teaching or for instructional purposes across a wide variety of subject areas, with the weight

of knowledge and breadth of application in language learning ultimately resulting in a more specialized field of study. (Levy, 1997, p.1)

The Delphi Method is a research strategy based on the following procedure, a small monitor team designs a questionnaire, which is sent to a larger respondent group. After the questionnaire is returned the monitor team summarizes the results and, based upon the results, develops a new questionnaire for the respondent group. The respondent group is usually given at least one opportunity to reevaluate its original answers based upon examination of the group response. To a degree, this form of Delphi is a combination of a polling procedure and a conference procedure which attempts to shift a significant portion of the effort needed for individuals to communicate from the larger respondent group to the smaller monitor team. (Linstone & Turoff, 2002, p.3)

Educational Software is a software package designed for presentation of the material for educational purposes.

EFL (English as a Foreign Languages) is a term describing teaching/learning English in the countries where English is not the native or primary language.

Instructional Technology (IT) is “a complex integrated process, involving, people, procedures, devices, and organization, for analyzing problems and devising, implementing, evaluating, and managing solutions to those problems in situations in which learning is purposing and controlled” (Educational Technology by Association

for Educational Communications and Technology as cited in Heinich, Molenda, Russell, & Smaldino, 1999, p. 406).

The Internet is

a computer-based global information system. The Internet is composed of many interconnected computer networks. Each network may link tens, hundreds, or even thousands of computers, enabling them to share information with one another and to share computational resources such as powerful supercomputers and databases of information. The Internet has made it possible for people all over the world to effectively and inexpensively communicate with one another.” (Encyclopedia Encarta Web site, 2003)

L2 – the second language or non-primary language being acquired by a language learner.

Pedagogy - the art and science of teaching.

Second Language Acquisition (SLA) is "the study of how second languages are learned. It is the study of how learners create a new language system” (Selinker & Gass, 1994, p. 1).

Target Language is a non-native language that is being acquired by a learner in the process of language learning.

Organization of the Study

This study will be organized into five chapters, followed by a list of references, a bibliography, and appendices.

Chapter One, Introduction, presents the introduction of the study, including the statement of the problem, the purpose of the study, and the significance of the study.

Chapter Two, Review of the Literature, reviews the relevant literature related to this study. Included in this study is a review of the literature related to historical perspectives, present conditions, and evaluation of Computer-Assisted Language Learning materials with particular emphasis on language teaching software.

Chapter Three, Methods and Procedures, identifies the methodology and procedures that will be used in the creation of the instrument, selection of the participants, and administration of the survey study.

Chapter Five, Summary, Conclusions, and Recommendations, will provide a summary of the findings, conclusions, and their implications for educational practice. Recommendations for further research will also be offered.

CHAPTER II

REVIEW OF THE LITERATURE

Introduction

From the beginning of its existence, computer technology has been used for various instructional purposes, including language teaching. The latter use has resulted in a relatively new area of knowledge, named Computer-Assisted Language Learning or CALL. Swift technological growth and multiple new developments in the theory of language teaching/learning have not only resulted in a drastic change in means and methodology of CALL, but also have transformed the role of the computer in that process (Levy, 1997). At present, there is an assortment of technological means available to language teachers and learners. However, the evaluation of CALL materials still remains an issue, mostly due to the following two reasons: a wide range of disciplines involved in their design, and the constant transformation of computer technology. The goal of this chapter was to review the factors involved in the evaluation of CALL materials with an emphasis on language teaching software. The subject was considered from multiple perspectives. First, a historical perspective on CALL and its evaluation criteria will be presented. Second, current CALL activities will be described and classified. Third, current examples of effective CALL use in language classrooms will be addressed. Last, multiple approaches to evaluation of CALL technology will be discussed.

Historical Perspective

Since the introduction of computers and their initial use for language teaching and learning, multiple transformations have occurred in both CALL technology and methodology. As Geissinger (1997) states,

Educational software has evolved over the past 40 years from Keller's Personalized System of Instruction to cognitively-stimulating microworlds and magical interactive multimedia experiences. Our views of the ways in which people learn effectively have changed with the evolution of media. (Geissinger, 1997, ¶2)

Those changes, derived from various sources, have greatly affected each other and, formed new trends and tendencies in the theory and practice of Computer-Assisted Language Learning. Retrospective studies on the subject affirm that both computer technology and its first application for language instruction were designed as a part of a military program. The former was invented in order to create a type of communication that would survive nuclear attack, and the latter was initially utilized to train military personnel (Delcloque, 2000). The first computers were capable of performing a very limited number of simple operations. Similarly language-teaching methods of that time could be described as "linear stepwise instruction" (Gessinger, 1997, ¶2). Yet, over a short period of time, computer applications have greatly evolved from text-based to multimedia-based and from local area networks to networks of millions of computers (Paramkas, 1998). In this same way, language instruction has advanced from being a structural and repetitious learning experience to one that is interactive and socio-cultural in nature.

Each development in this process has brought a certain unique aspect to the development of CALL; therefore, knowledge of the reasoning behind each change may provide a valuable insight into understanding the present condition of Computer-Assisted Language Learning. The following sections describe the most commonly identified stages of CALL history recognized by significant technological and pedagogical innovations.

CALL in the 60s – Early 70s

At the time when computers were first applied to language instruction, the nature of the latter could be described as an: “empiricist theory [that] was predominant in language teaching, a theory described by Stern (1983, p. 169) as 'pedagogically audiolinguism, psychologically behaviorism, linguistically structuralism'” (Stern, as cited in Levy, 1997, p. 14).

While it was believed that language teaching requires linear repetitious structure, the first computers, supporting few limited simple operations, were capable of providing assistance for this process. Since language instruction was limited to repetition, memorization, and straightforward translation, those operations were easily programmable for the first machines. The first two projects, PLATO and TICCIT, illustrate these principles in action.

PLATO

Programmed Logic for Automatic Teaching Operations (PLATO) was designed with the purpose of taking tedious drill and practice exercises out of the language learning classroom to free some instruction time for more meaningful experience. PLATO was offering a random set of activities, fitting immediate

instructional needs rather than presenting a language course in itself (Delcloque, 2000; Levy, 1997).

The design offered several levels of difficulty. Once the level was set, the prompt would give the student a word, combination or a sentence for translation or correction. The student was to type in his/her response. If the answer were correct, the student would be advanced to the next item. In the case of an incorrect reply, the same task would appear again. The student also had the option of looking up the correct answer or the corresponding rule. In addition, the system offered communication between learners and the teachers, who were presently working on it. It also kept a log of the results for each student (Kern & Warschauer, 2000).

The program for PLATO was written in the computer language TUTOR, which was easy to learn and use even for people with no previous computer-related knowledge. Therefore, language teachers were able to design CALL activities according to their classroom needs. Among the advantages of PLATO was an opportunity for a learner to self-pace his/her instruction and for a teacher to change the program to the individual needs of his/her students. Among the disadvantages indicated in the literature were the limited nature of exercises, unsystematic type of instruction, scarce feedback, and incapability of TUTOR to support any kind of sophisticated structure (Delcloque, 2000; Levy, 1997).

TICCIT

Unlike PLATO, Time-Shared, Interactive, Computer Controlled Information Television (TICCIT) was designed as a step-by-step course of instructions with a logical connection between exercises. Based on component display theory, in which

instruction was considered a sequence of three steps (i.e., classification, presentation, and prescription), access to the instructional content in TICITT was controlled by *rule, example, practice, advice, objective, easy, and hard* keys. While providing an adjustable pace of learning, TICCIT did not give the users any control over the content beyond specified frame (Levy, 1997).

Late 70s – 80s

At this time, the history of Computer-Assisted Language Learning was marked by several important milestones in both Second Language Acquisition (SLA) and computer engineering. In SLA, the newly introduced communicative approach to teaching languages has shifted not only the traditional understanding of how language learning occurs, but also how the roles of the student, instructor, and instruction have evolved (Kern & Warschauer, 2000). It had resulted in two significant changes. First, the learning style, not the language of the student, was made the center of instruction. This means that,

We now understand that learners operate on many levels, some of which are subconsciously chosen, and that they (who are also ourselves) are influenced by both intrinsic and extrinsic factors. Educators no longer think that linear stepwise instructions can supply the learning needs of even the most naive learner or that the learning context can be rigidly controlled. It seems well understood that learners bring a variety of skills, knowledge and responses to every learning situation and that the teacher must supply a number of ways for them to approach new learning, engage with it and absorb it. (Geissinger, 1997, ¶2)

Second, communication and imitation of linguistic patterns, like the one, which occurs in initial L1 learning, were believed the most important way of instruction. Therefore, the role of a language instructor was to facilitate linguistic interaction to bring "the independence of language and communication" to a student (Levy, 1997, p. 25).

The introduction of the first personal computers, invention of software, and advancement of flexible programming languages like BASIC, were among the most significant achievements of the time. These accomplishments allowed language instructors to accommodate CALL activities to fit students' needs (Levy, 1997).

The newly introduced concept of hypertext and hyperlinks allowed the use of non-linear structure in CALL activities and supported the design of highly interactive teaching/learning materials (Levy, 1997). By definition, *hypertext* is a type of text that contains *hyperlinks* in it. A *hyperlink* is a word or phrase, specifically marked in the text by the means of font, color, or underlining, which, by clicking on it, takes the user to a separate screen, which explains or describes the word or concept marked in the text. The second screen also allows easy navigation back to the initial screen, when the work with the second screen is completed. Hypertext or non-linear text structure allows simultaneous presentation of multiple concepts within one context. For example, a reading exercise can have new vocabulary words as hyperlinks taking the reader to the screen with their definitions. Thus, the function of the computer has evolved from tutoring and testing to interacting with a learner.

More opportunities for design were offered to teachers with programs like *Storyboard* or *HyperCard*, which gave options for designing from scratch and/or

applying ready-to-use templates to fit the instructional content. Growing availability and accessibility of computer programs for language teaching and learning brought increasing attention to this form of instruction. Yet, it was also the time that marked the beginning of an ongoing discussion about the appropriateness of computer use in language teaching/learning (Levy, 1997).

90s – Present

This period in history was signified by the introduction of the Internet to the public at large and by the implementation of the socio-cultural approach to teaching languages (Kern & Warschauer, 2000). Despite the fact that the Internet was invented in the 80s, it was only in the 90s that it became easily accessible by the general public. At the same time in SLA research, a socio-cultural theory was adopted from the sociological teaching by Hymes (1974) and Halliday (1989).

The socio-cultural theory supports the critical importance of context rather than text; therefore, in accordance with it, language instruction is no longer viewed as occurring through the study of grammar or even staged communication, but through learning socio-cultural aspects of the L2. With the growing understanding of the significance of cultural aspects of the target language, the need for authentic materials has increased exponentially (American Council on the Teaching of Foreign Languages, 1996; Kern and Warschauer, 2000). While real life authentic interaction involves high cost and time investment, the Internet provides a unique opportunity to reach this goal without leaving the comfort of the classroom or home. Therefore, Computer Mediated Communication was adapted by language teachers as an effective method of teaching.

Since the 90s, Computer Mediated Communication has been used by CALL instructors as one of the most important ways of providing authentic instruction. Several formats were developed depending on the capabilities and needs of learning groups, including e-mail pals, global classrooms, role-playing in virtual context, mentoring, tutoring, and teaching support (Paramskas, 1998). While this type of learning has been found effective in many cases, several disadvantages have been identified by language teachers. Among the most commonly noted disadvantages are the low reliability of some Internet resources, technological problem, associated with the use of computers and networks, and high cost of technology (Parra, 2000; Vilmi, 1995).

Present day Computer-Assisted Language Learning has been built on the latest developments in computer engineering and Second Language Acquisition as well as other sciences, and comes in a variety of formats. Most activities available to language students combine a number of approaches, including structural grammar components, communication, and studies of socio-cultural backgrounds of the L2. Only one type of CALL materials, language teaching software, comes in a variety of structures, including drill and practice, tutorials, games, simulations, discovery programs, and problem solving (Heinich, Molenda, Russell, & Smaldino, 1999). The Internet has also built a whole new type of language learning interaction, including accessing, retrieving, and sharing information and tools like authentic texts, pictures, movies, and shared software (shareware), and communicating synchronously and asynchronously with target language native speakers, teachers, and other learners through chat room, e-mail, and listservs (Paramkas, 1998).

CALL at Present

Classifications

Throughout its evolution, Computer-Assisted Language Learning has drawn from the theory and practice of each of its historical stages. It has adopted various methods and tools introduced at different times through different pedagogical perspectives (Kern & Warschauer, 2000; Levy, 1997). The following description of its present formats is offered by researchers. It can appear as electronic mail, remote access to library and other databases, subscription to lists and other electronic fora, subscription to electronic journals, file transfer (Marcos, 1994), the World Wide Web, streaming audio and video, search engines, chat, audio and video conferencing, messaging, Web course management (LeLoup & Ponterio, 1998), customizing, template and authoring programs, and compact disk technology (Higgins, 1993).

Reports of Practical Applications

Growing technological capabilities and an increasing number of formats, which can be used in language teaching/learning, have resulted in an increased expansion of CALL. Chapelle (1997) believes that, “As we approach the year 2000, the technical realities for CALL have expanded far beyond prior expectations. To attempt the best pedagogical applications for this technology, research on CALL use is needed” (p.16). The following descriptions of recent CALL experiences are drawn from current published sources. They address successful use of CALL whether targeting a specific language skill (i.e., reading, vocabulary) or a combination of skills (i.e., virtual classroom).

Levine, Ferenz, and Reves (2000) describe the use of the CALL approach to teaching reading in their world language classes in Bar-Ilan University, Israel. In one class that was used as an experimental group, reading instruction was given with the use of authentic sources presented through the computer. In the other class, which was a control group in this experiment, reading was taught through traditional printed sources. Statistical analysis of the results of the test that took place with both groups at the end of an instructional period, revealed that computerized learning environment positively affected the development of critical reading skills to a much greater extent than traditional instruction. Therefore, researchers now believe that CALL stimulates student transition from *learned reading skills* to *authentic reading skills*, making it likely that the students will turn into independent readers (Levine, Ferenz, & Reves, 2000).

Williams and Williams (2000) also used Computer-Assisted Language Learning techniques for teaching English to their ESL students who have a novice level of English proficiency. Computers were used as a medium for the students, who were learning to read and to key in words simultaneously. From their classroom observation and testing, Williams and Williams (2000) conclude, “The observations done for this study support the assumption that integrating reading and computers will improve the ESL student reading skills” (p.99). They state that the L2 reading proficiency of their students “improved dramatically” (p. 99).

In addition to general computer resources, such as the Internet or word processing programs, computer programs, especially designed for teaching different language skills, are often used by language instructors. Groot (2000) describes an

experience with one such program, intended for vocabulary acquisition. He expresses a concern that the usual number of words acquired in a language classroom is not nearly enough to make a student a proficient and confident speaker of the L2. The author suggests that if completely trusted to incidental learning, the necessary acquisition of the second language will require a very long time. Therefore, in search of different methods, Groot (2000) offers a special computer program, *CAVOCA* (Computer Assisted Vocabulary Acquisition), that allows vocabulary learning as fast as 25 words in 55 minutes. This program was tested with EFL students at Utrecht University, and proved to be effective with that limited group. However, for the most effective vocabulary instruction, the author suggests using both an incidental and cognitive approach, which is used in CAVOCA (Groot, 2000).

While the previous cases address acquisition of certain skills (i.e., reading, vocabulary, typing in L2), the type of instruction called *virtual classroom* targets several of them at the same time. This term describes instructional tasks that are performed collectively by the students through electronic synchronous (i.e., chat) and asynchronous (i.e., e-mail) communication. This type of communication can be organized by students in the same class as well as between learners from different classes, cultures, and locations (i.e., language students and native speakers of the learners' target language). The type of communication occurring in such activities is referred to as Computer Mediated Communication (CMC).

An example of effective use of the virtual classroom and CMC is presented in Vilmi's (1995) research. She describes her experience of organizing virtual classroom with students in her class. A total of 130 students from multiple countries

participated in this project. The objective of her project was to increase “student global awareness” as well as improve their language skills (Vilmi, 1995, ¶18). A similar project was undertaken by Shulman (2000) and involved students from Brazil and the U.S. Like Vilmi (1995), Shulman (2000) believes that cultural knowledge is an important aspect of language learning and wants her students to study it from authentic sources.

LeLoup and Ponteirio (1999) describe another type of Internet use in Computer-Assisted Language Learning. In their study, the Internet is considered to be a medium for the creation and publication of language teaching materials. Juan Ramon de Arana, a teacher of Spanish, designed an instructional interactive website that he has successfully used with his students (LeLoup & Ponterio, 1999). However, due to the accessibility of Internet resources and versatile and broad nature, this website can be used by all Spanish teachers and their students. Therefore, teacher authoring on the World Wide Web cannot only serve the purpose of a particular classroom, but can also provide valuable resources for the teacher’s colleagues (LeLoup & Ponterio, 1999).

Finally, computer software as well as Internet applications, are also used by language teachers for testing. Dunkel (1999) notes that tests, due to their structured nature, are easily programmable, and if put on-line, are also easily accessible. She compares computer testing with a traditional paper-and-pencil version and lists the advantages of the former. The latter include self-pacing, justified challenge ("test takers are challenged but not discouraged"), immediate feedback, security, and multimedia presentation (i.e., text, graphics, photo, video, and audio) (Dunkel, 1999,

p. 79). Considering both the advantages and disadvantages of electronic testing, the researcher drew the conclusion that although computer testing may not provide an in-depth assessment of students' abilities, it makes it easy for the teacher to make a quick evaluation of a student's current language level and needs (Dunkel, 1999).

The examples described above illustrate various applications of Computer-Assisted Language Learning applications. They address a variety of formats, classroom types, age groups, language proficiency levels, and target skills. They support the idea that computers are a flexible medium in language instruction and can be utilized in a number of ways. Yet, at the same time, with such a large and diverse population of language teachers and students using Computer-Assisted Language Learning, a need for a common pedagogy and a new paradigm has been addressed by several researchers in the area of CALL (Chapelle, 1997, 1999; Levy, 1997). Among other advantages, a new paradigm may assist language instructors to evaluate CALL materials and select the ones that best fit up-to-date SLA research (Chappelle, 1998), their current technological proficiency (Kasper, 2000; LeLoup, 1997), and their students' level of L2 proficiency (Dunkel, 1999), interests, age, and culture (Kung, 2002; O'Dowd, 2001; Vilmi, 1995).

Evaluation

The need for a sound paradigm in the field of Computer-Assisted Language Learning has been frequently addressed in the research. As Chapelle (1997) states, "A glance through the computer-assisted language learning (CALL) literature of the 1990s reveals the profession's quest for principled means of designing and evaluating CALL" (p.19). At the same time, Wood (2001) admits that, "the number of software

products on the market grows exponentially. In 1996, a record-breaking 693 educational software titles were released [*The bright future of children's software, 1997*]” as cited by Wood (2001, p. 167). With the rapid spread of computer technology, increasing authoring possibilities, and a large spectrum of CALL materials on the market, the question of what makes a CALL activity suitable for a particular classroom still remains unanswered (Chapell, 1997).

Although much research has been conducted in this area, in most cases, it often addresses only a specific situation. Yet, while Computer-Assisted Language Learning is drawn from multiple disciplines, it is believed to require a multi-dimensional approach in evaluation. The purpose of this study is to examine the criteria for evaluation of CALL software that would embrace the perspectives of various disciplines related to Computer-Assisted Language Learning.

From the analysis of current CALL literature, the following categories related to the following disciplines were identified: Second Language Acquisition, Pedagogy, Instructional Technology, and Socio-Cultural Studies. While these categories may not include the entire scope of perspectives related to CALL, their choice is based upon the analysis of current published sources in Computer-Assisted Language Learning and Instructional Technology. It is possible that some of the aspects in these categories may overlap, however; each of them has its own number of distinctive characteristics that may contribute to the general picture of software evaluation.

The Perspective of Second Language Acquisition

There is a number of sources in current CALL literature that deal with SLA-related aspects of CALL materials. One of the biggest concerns of the researchers

working in the area is the presence and quality of linguistic input and output in the design of CALL instruction. This approach is built on the development of Krashen's theory of Comprehensible Input, which states that in order to be able to acquire a second language, a learner needs to be exposed to what Krashen (1985) calls *comprehensible input* in the target language. He believes that, "Acquisition requires meaningful interaction in the target language - natural communication - in which speakers are concerned not with the form of their utterances but with the messages they are conveying and understanding" (Krashen, 1985, p. 39). He further argues that the input should be one step beyond than a learner's current level of language competency to be effective – a principle usually demonstrated with the *i+1* formula.

Further research by Swain and Lapkin, cited by Nagata (1998), offers a position that the presence of comprehensible input alone develops comprehension, yet it is not effective in teaching language production. Therefore, the researchers suggest not only providing quality input, but also creating conditions for learners to create what they call *comprehensible output*.

Swain conjectures that producing the language, as opposed to simply comprehending the language may force the learner to move from semantic processing to syntactic processing, thereby facilitating more grammatical competence. Swain also refers to the phenomenon of individuals who can understand a language and yet can only produce limited utterances in it: a ninth grade immersion student said, 'I understand everything everyone says to me, and I can hear in my head how I should sound when I talk, but it never

comes out that way' (p.248). This indicates that comprehension does not necessarily transfer to production. (Nagata, 1998, p 25)

Chapelle (1997; 1998) and Nagata (1998) take this requirement even further. They both are concerned not only with the present, but also with the amount and quality of comprehensive output in CALL materials. Thus, Chapelle (1997) criticizes some of the CALL activities for either not providing or providing very limited output, such as mouse clicks or yes/no options. In the first example that she describes, a learner is offered a text with a number of idioms that are to be introduced. The new idioms are highlighted in text as hyperlinks that take the user to pages with vocabulary definitions and examples. It is apparent that in this type of activity, the learner's role is limited to mouse clicks, therefore, the operational knowledge of the new vocabulary is not likely to be gained from that experience. In a second example, a group of students reads the text, performs certain tasks, and discusses it on-line amongst the group members. This type of CALL activity supports not only input, but also potentially rich output; therefore, according to Chapelle (1997), the information is more likely to be acquired. Finally, in the third example, a learner *talks* to the computer by asking content-related questions making his/her final conclusions based on the results of that interaction. (The learner is playing investigator in a murder case and finding *who done it* by asking questions and building a hypothesis based on the answers.) Due to the authentic nature of both input and output and non-restricted output, this activity was also found to be effective by the author (Chapelle, 1997).

Examining the role of output in CALL, Nagata (1998) also conducted an experimental study where two groups of language students received the same

instruction, yet the control group's practice offered the exercises with multiple choice output; the experimental group had to construct their own output. Both groups were tested at the end of the course. Comparative statistical analysis of the means of the results from both groups showed a significant difference in favor of the experimental group, which may support the idea that the opportunity for comprehensible output makes a crucial investment into the effectiveness of L2 acquisition.

The importance of output in CALL materials is also emphasized by Bishop (2001) in her list of guidelines for teachers, in which she advises them to look for ...true educational interactivity. Immediate feedback for self-correction and practice or rewards for accuracy are essential for language learning. Programs that record a user's verbal responses and compare them to a native speaker's, or that use voice recognition, are very useful. (p. 41)

Therefore, in choosing, designing, and applying CALL activities, the nature of student interaction with the software or of a student with another student with the help of software, remains an important issue. Thus, Egbert (2001) states that while student work with software is important, students "are not just part of the same group, but they have reasons to talk to each other, to make decisions, negotiate meaning, and develop understanding together" (§1), while Stepp-Greany (2001) concludes that, ...students are not passive recipients of knowledge. Instead, they are active participants in the construction of new knowledge that is idiosyncratic, and derived from the learner's prior experience and need to create equilibrium (i.e., find meaning), when faced with a new situation that creates cognitive dissonance. (p.167)

Therefore, from the perspective of Second Language Acquisition, the conditions for significant and meaningful learner output leading to active learning are one of the most important criteria for effective CALL software.

The Perspective of Pedagogy

Among the pedagogical aspects in CALL material evaluation, the most common ones usually address either the types of cognitive perception or the issue of interest and motivation. The nature of learning has been described in the field of pedagogy for a long time. It is believed that in the planning and designing of any educational activity, it is important to consider the processes of cognition. In their findings, CALL researchers have emphasized the significance of multiple pedagogical factors (Chapelle, 1997, 1998).

While it is believed that learners differ in their abilities, learning styles, strengths and weaknesses, it is important to target various skills and include a variety of activities and types of instruction within the same course (Chun & Plass, 1997). Plass (1998) suggests employing a cognitive approach in software interface design, an approach that does *not* depend on the nature of computer application, but rather on a student's learning style and needs. He argues that evaluation criteria need to be developed based on domain specific learning processes and activities and on the cognitive processes that these activities involve (Plass, 1998).

In addressing the same issue, much attention has also been given to combining incidental and directed instruction and structural and integrative approaches to serve different purposes and to accommodate different types of learners. Thus, analyzing computerized phonetic instruction, Chun (1998) comments, "computers can both

provide training in a production and perception of speech and create environments to facilitate interaction” (p.64). She believes that phonetic instruction should not only include “practice of sounds and nuances” but also “elicit and encourage [the] practice of specific types of interactions” (p. 65).

Many researchers have commented on the importance of the relevance of teaching content related to student goals and interests. While interaction and language production are crucial for language learning, there are several factors related to interest and motivation that can support or *break* the interaction. Thus, Kung (2002) citing the studies by El-Wardi and Johns (1998), Warschauer (1995), and Blatt (2000), describes the four types of *broken interaction* in her e-mail pen pal project,

lack of response [El-Wardi & Johns, 1998; Warschauer, 1995]: the partners are not reliable and do not write regularly [Wegner, 1997 in Blatt, 2000], there are demands of other activities for students [El-Wardi & Johns, 1998].

lack of purpose [Warschauer, 1995]: interaction wanes after the initial getting-acquainted period because participants do not know what to say to each other.

low linguistic level [Blatt, 2000]: students write carelessly repeating the same mistakes and without making any progress in their learning of the second language.

logistical difficulties: lack of experience with software and/or typing [El-Wardi & Johns, 1998], difficulty accessing equipment [El-Wardi & Johns,

1998; Sela, 1997], frequent and severe technical problems [Wegner, 1997 in Blatt, 2000]. (Kung, 2002, ¶11-14)

Therefore, from the perspective of pedagogy, addressing various learners with various learning styles and motivational natures through relevance to learner interests are important criteria for the evaluation of language teaching software.

The Perspective of Instructional Technology

Multimedia computer applications have presented a unique opportunity to put into practice some important language learning principles. Chun and Plass (1997) address the dual coding theory and its implications for the use of multimedia in language teaching. This theory states that in the process of language acquisition, a learner operates with two systems of storage, *verbal* and *non-verbal*, which mutually support each other in that process. A student's interaction with verbal information in the L2 is associated with his/her previous verbal and non-verbal knowledge. In the case of L2 learners, the verbal information may be limited, yet the non-verbal information can help the L2 learner understand and process information in the target language (Chun & Plass, 1997).

In this case, computer multimedia provides a variety of formats to make language learning meaningful for a student. In addition to text, it offers pictures, audio, video, and animation. By also offering different types of interaction with a learner, it targets learners with different learning styles (Chun & Plass, 1997).

On the other hand, however, researchers state that if overused, this same type of multimedia (animation) can become distracting, and as a result, can decrease the effectiveness of language learning material. The difference between meaningful and

non-meaningful use of animation is described by Wood (2001) in her analysis of vocabulary teaching software for children.

Many of the animations reviewed were gratuitous by nature, attempting to captivate children but not to illuminate words meaning: for example, a Greek stature that comes to life and tells a joke or a flamboyant quiz show host whose speech is full of puns. Some animations, however, played a more central role in adding visual (and often auditory) information to word meanings. In *What's the Secret?*, for example, the animation is used to show laws of physics at work (including technical words) in propelling a roller coaster. (Wood, 2001, p. 179)

Therefore, it is suggested that in order for multimedia use to be effective, the following guidelines be adopted:

- Multimedia elements need to support and relate to the instruction (Chun & Plass, 1997; Wood, 2001).
- Multimedia elements need to be selected with the understanding and evaluation of a user's needs and abilities (Chun & Plass, 1997).
- Depending on the content of instruction and the user's skills, the right orchestration should be applied to multimedia elements (simultaneous, before, or after instruction) (Wood, 2001).

Another important instructional issue in software evaluation is student attitude towards learning. For example, Bitter and Pierson (1999) include *interest* in selection guidelines. They give the following advice to the instructors selecting educational

software, "Rate whether the way content is organized and presented will hold the interest of these students" (p. 115). Heinich, Molenda, Russell, and Smaldino (1999) also state that in educational software "information needs to be presented in an interesting manner to maintain student interest and involvement in the tasks" (p. 218). In their software *Appraisal Checklist*, 1 of the 11 major criteria is *Arouse Motivation/Maintain Interest* (Heinich, Molenda, Russell, & Smaldino, 1999). The same principle is formulated by Eskenazi (1999), when she includes "Learners should feel at ease in the language learning situation" (p.62) in the list of her CALL materials requirements. It clearly shows that lack of interest, lack of challenge, or, on the other hand, problems with understanding, can result in low motivation. Therefore, it is believed important to consider all these factors in choosing or designing a CALL activity.

Thus, from the perspective of Instructional Technology, consideration of rational multimedia use according to the learner's needs, abilities, and interests (Bitter & Pierson, 1999), and forming a positive attitude and motivation by facilitating success and addressing interests of a student (Heinich, Molenda, Russell, & Smaldino, 1999), are essential factors to be considered in educational software selection.

The Perspective of Socio-Cultural Studies

The roots of the socio-cultural perspective on language learning lie in sociological theories by Hymes (1974) and Halliday (Halliday & Hasan, 1989). These scholars believed that the socio-cultural context defines the content of language, and knowledge of the target language culture is as crucial to language

learning as knowledge of formal language structure. Thus, Hymes (1974) offered a SPEAKING model, which, he claims, defines the content of language production:

- Setting and Scene – time, place, and physical circumstances of speech.
- Participants – speakers and audience.
- Ends – purposes, goals, and outcomes of speech.
- Act sequence – form and order of speech.
- Key – tone, manner and spirit.
- Instrumentalities – forms and styles of speech.
- Norms – social rules that governs participant's actions and reactions.
- Genre – ways of presenting the same story differently (as an anecdote or as a moral example).

Like Hymes, Halliday (Halliday & Hassan, 1989) believed in the importance of the Socio-Cultural aspect of the language. While he defined three functions in any language: ideational (use of language to express content), interpersonal (use of language to maintain social relationships), and textual (use of language to create situationally acceptable discourse), the last two, in his viewpoint, were at least as important as the first one, and therefore, must be addressed in language teaching/learning.

Recently, these perspectives have been addressed in the Standards for Foreign Language Learning (American Council on the Teaching of Foreign Languages, 1996) in their *5C Model* (Figure 1), which presents a basic paradigm for world language education (i.e. Communication, Cultures, Communities, Connections, and Comparisons).

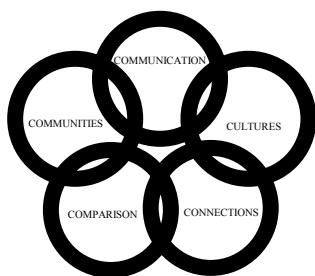


Figure 1. Five C's of Foreign Language Education

Note. From *Standards For Foreign Language Learning. Preparing For The 21st Century*, by American Council on the Teaching of Foreign Languages, 1996 Yonkers, NY: American Council on the Teaching of Foreign Languages, Inc.

The meaning of each competency is specifically described in the Standards and can be addressed through CALL materials (LeLoup, 1998)

Communication

The Communication Standard requires students to "engage in conversation, provide and obtain information, express feelings and emotions, and express opinions," "understand and interpret written and spoken language on a variety topics," "present concepts and ideas to an audience of listeners or readers on a variety of topics" (American Council on the Teaching of Foreign Languages, 1996, p.4). LeLoup (1998) has a clear vision of implementing these values into a language curriculum through the use of computer tools.

Can you think of any electronic communications technology that would be an asset here? How about email? IRS? Usenet groups? MOOs? FL learners could be in contact with native speakers, teachers, or even other learners of target language and fulfill all of the objectives of this standard easily. (LeLoup, 1998, ¶12)

Cultures

The socio-communicative language teaching theory, which is rooted in studies conducted by Hymes (1974) and Halliday (as cited in Kern & Warschauer, 2000), stresses the importance of cultural and social context of communication. Hymes (1974) believed that in communication, the context plays a more important role than the content itself. Similar to this model, the cultural competency in the Standards emphasizes the necessity of learning a target culture while learning a target language. As is stated, the students "cannot truly master the language until they have also mastered the cultural contexts in which the language occurs," thus, it is required that learners "demonstrate an understanding of the relationship between the practices and perspectives of the cultures studied" as well as "between their products and perspectives" (American Council on the Teaching of Foreign Languages, 1996, p. 5).

While it is hard to provide authentic products of culture and even explain its perspectives at times, Internet technology has become a valuable source for information and learning activities. As LeLoup (1998) notes,

How might students come in contact or become familiar with practices and perspectives of the target language culture, particularly if no native speakers are to be found in their vicinity? Might the WWW be useful here? Or FTP? For example, a local newspaper briefly covers a fiesta in Spain called the Tomatina with a photo and two-line explanation of people throwing tomatoes at each other. Why do they do that? Is this an ancient ritual? A recently introduced practice? A long-standing tradition? Language learners can use WWW to look up information on this happening that most likely will not be in

the encyclopedia. Again, technology can bring the world to our classroom, and distance difficulties melt away. (§14)

The other side of the same matter is an issue of cultural sensitivity, which should be specifically addressed in the design of each activity. If it is a Computer Mediated Communication type of activity, Bond (2002) notes that " Lack of face-to-face interpersonal communication (lack of feeling you really know your student), lack of non-verbal communication, plus different cultures and different native language, can result in miscommunication and misunderstandings" (§37).

Even in language teaching software cultural misconceptions can be present. One of the striking examples of socio-cultural insensitivity can be found in a language teaching software review by Sokolik (2000).

More troubling is the content. For the same item presented above, the choices are:

- A. The computer is broken.
- B. The man is welding.
- C. The man is wealthy.
- D. He is taking a break.

The feedback from item C says: "No, wealthy means RICH, but this is clearly a workman." I'm disturbed by the implications here. Although a distracter is trying to address possible confusion between pronunciation of 'welding' and 'wealthy,' the feedback implies that the understanding should be based on the implication that a 'workman' [sic] wouldn't be wealthy. (Sokolik, 2000, §9-11)

Another example from Bishop cited in Pofstak (1999), describes the same type of culturally unsound approach in educational computer resources:

I call those engines ‘translations on the fly.’ They are mechanical translations that can range from being funny to insulting. For example, the U.S. Department of Education Web page once translated ‘undersecretary of education’ to ‘inferior secretary of education’ in Spanish. (p. 56)

Therefore, researchers encourage teachers to closely examine the cultural content of instructional materials, reflecting on possible assumptions they may produce by their content. Thus, analyzing his experience with Computer Mediated Communication, O’Dowd (2001), citing Hollet (1997), gives the following advice to language instructors.

Therefore, I would see the second challenge for teachers is to help develop their students' ability to interact with sensitivity, insight and tolerance with people from other cultures - whether that be in an on-line , face-to-face or written environment. Hollet (1997) sums this up nicely:

...learners need to be conservative in what they send, so their messages can be easily understood across cultural boundaries. But at the same time they must be liberal in what they receive, being both able and willing to understand the discourse of other cultures [1997, p. 19]. (O’Dowd, 2001, ¶42)

While socio-cultural mistakes lead to misconceptions in language, true cultural knowledge can benefit a language learner immensely. Rich multimedia formats give a great opportunity for presenting cultural content.

Through multimedia in particular, language is no longer just a list of grammatical paradigms or lexical items. Rather, it is intimately associated with all kinds of verbal and paraverbal behaviors, an acoustic and visual context that is dissociable from the large societal context in which the words are uttered. Language, in a sense, has become culture. Similarly, culture is no longer just factual pieces of information that textbooks present in the form of culture capsules on foreign mores, but is produced and reproduced under our very eyes, on the screen, in what people say and how they say it. In multimedia, culture is inscribed in language use. (Kramsch & Andersen, 1999, p. 31)

Connection

The Connection standard refers to the importance of connecting the knowledge obtained in the world language classroom with other disciplines. The students are to "reinforce and further their knowledge of other disciplines through the foreign language" and at the same time "acquire information and recognize the distinctive viewpoints that are only available through the foreign language and its cultures" (American Council on the Teaching of Foreign Languages, 1996, p. 5).

Several disciplines in the school curriculum are connected with the studies of other countries. Among those are literature, geography, history, political science, science and math. While a thin layer of very basic knowledge on the subject can be obtained from the courses in the discipline, the deeper understanding of some knowledge might involve understanding of a second culture and even further require knowledge that is only available in another language.

For instance, a philosophy student might be interested in what Russian philosophers have to say about American philosophical schools. This information can only be found in original Russian texts in written form. The WWW might be of help here locating original target language texts on this topic. Another possibility would be to engage in a target language conversation with Russian philosophers, most likely on electronic Philosophy discussion lists. (LeLoup, 1998, ¶14)

Comparisons

The standard of Comparisons refers to the students' ability to understand "the nature of language through comparisons of the language studied and their own" as well as "cultures studied and their own" (American Council on the Teaching of Foreign Languages, 1996, p.5). In fact, this standard brings in a higher cognitive ability to generalize knowledge and to develop the concept of the culture and language and their roles in the lives of people. The ways to fulfill this standard involve exposure to multiple authentic cultural and linguistic materials. The standard also involves participation in a dialogue with the people who represent a particular target language and culture, all of which can become easily accessible through computer technology.

Community

The Community standard brings forward the goal of turning the students into life-long learners, which means that they "use their language both within and beyond the school setting" (American Council on the Teaching of Foreign Languages, 1996, p. 6). This particular goal might be viewed in direct connection with computer

media, for it was computer technology that broke down the barrier of time and distance and is gradually turning the world into one global community.

Chapter Summary

This chapter has presented a literature review related to several issues related to the evaluation of CALL materials, including language teaching software. A historical perspective demonstrates the evolution of CALL materials as well as the evolution of a CALL evaluation system. A description of several CALL present day experiences presents a broad variety of CALL use in language instructional classrooms. Yet, despite the multiplicity of its applications, the researchers address the lack of a common pedagogically valid evaluation system for CALL materials. The review of the published research on evaluation of CALL materials (with emphasis on language teaching software) reveals a number of issues related to pedagogy (e.g., SLA, IT, and socio-cultural studies) that need to be addressed in an approach to language teaching software evaluation. The next Chapter, Chapter Three, will describe the design of the study, including methods, instrument, and participants.

CHAPTER III

METHODS AND PROCEDURES

Introduction

Since the creation of Computer-Assisted Language Learning, a broad discussion has existed in the literature concerning the effectiveness of its methods. A number of experimental studies have been described, where the traditional ways of language teaching were compared with CALL. This discussion still continues (Chapell, 1997, 1999; Salaberry, 1999), yet, with computer technology development and growing availability of computers, increasing numbers of language instructors and students use the computer as a means of teaching and learning language (Wood, 2001). The question being raised at present concerns which computer materials available on the market are the most effective (Chapelle, 1997, 1998). The choice of the best means for teaching and learning necessitates defining the criteria to be used in the selection of the CALL materials.

At present, there was a great number of ways in which computers can be used for teaching/learning languages (i.e., from word processing to e-mails, chat, World Wide Web, and voice recognition tools). Software is only one of the formats of CALL materials, yet it is quite commonly used. This study examined professional perspectives on evaluation criteria for CALL software. It is therefore hoped that the results of this research may contribute to the design of selection guidelines for language teaching software. Realizing the multidisciplinary nature of Computer-Assisted Language Learning, research in the following disciplines was used as an

initial source for establishing this list of criteria: Second Language Acquisition, Instructional Technology, pedagogy, and socio-cultural studies.

A descriptive research design was applied to achieve the goals of this study. In the early stages of this research, lists of software evaluation criteria were developed based on current CALL published research, educational software evaluation presented in print, criteria published in on-line journals, and on the websites of professional CALL and IT associations, IT textbooks and articles. Based on the review and analysis of the studied materials, a set of criteria was developed by the principal investigator and incorporated into a Delphi survey, which was subsequently presented to a panel of experts for their review and consideration.

Research Questions

The development of the above mentioned criteria list for the Delphi study addressed the two previously formulated research questions:

1. What criteria for evaluation of language teaching software are identified in currently published literature, including textbooks, software reviews, and Instructional Technology and Computer-Assisted Language Learning research?
2. Which criteria are deemed important by the representatives of the Board and Council of the International Association of Language Learning Technology?

The following three-phase research design was adopted to answer the research questions.

Research Design

In order to answer the questions raised in the study, the following research design was developed and implemented by the researcher. It consisted of the following three phases.

- Phase One: A continuous list of evaluation criteria for language teaching software was gathered from a number of diverse published sources (i.e., research journal publications on Computer-Assisted Language Learning and Instructional Technology, Instructional Technology textbooks for teachers, educational software and language teaching software reviews).
- Phase Two: The list of criteria developed in Phase One was modified by eliminating repetitious criteria and dividing the remaining items into the following five major groups based on area or approach related to each criterion: Pedagogical, Applicability, Use of Media, Socio-Cultural, and Technical.
- Phase Three: Based on the criteria collected in Phase Two, a Delphi Study was conducted to further gather information on the importance of each of the criteria from experts in the field of Computer-Assisted Language Learning in order to provide teachers, students, and software designers with an adequate list of criteria for software selection.

Phase One: Identifying Sources of Evaluation Criteria

In this phase of the research, a thorough analysis was given to the sources related to Computer Assisted Languages Learning and Instructional Technology in

their relevance to software evaluation. The following sources were found to address the issue and, therefore, were used for criteria identification: published educational and language teaching software reviews, Instructional Technology textbooks for teachers providing a list of criteria for evaluation and selection of educational software, and research articles in CALL and IT areas discussing issues related to evaluation of educational and language teaching software.

Phase Two: Organization of Data

During this phase, language teaching software evaluation criteria were drawn from all the sources identified in Phase One in order to form a continuous list. The repetitious items were deleted from the list and the remaining criteria were analyzed and organized into five major group in reference to the aspect it was addressing:

Group One: Pedagogical Aspect; Group Two: Applicability Aspect; Group Three: Use Of Media Aspect; Group Four: Socio-Cultural Aspect; and Group Five: Technical Aspect.

Due to the large number of criteria included in Groups One and Two and a variety of issues embraced by those criteria, the former were further organized by the researcher into thematic subgroups within main groups. Thus, in Group One, the following subgroups were identified by the researcher:

1. Conditions for quality comprehensible input
2. Conditions for quality comprehensible output
3. Application of the principles of cognition
4. Motivation
5. Presentation of the content

In Group Two the following subgroups were implemented into the survey design:

1. Learner Perspective
2. Teacher Management
3. General Perspective

Finally, Groups Three, Four, and Five have remained in their original condition.

Phase Three: Conducting a Modified Delphi Study

In order to evaluate and validate the list of criteria described above, the researcher used a modified Delphi Study, a technique successfully used by many researchers for similar purposes. The Delphi method applied in this study is described below, followed by specifics of its implementation in the current research study.

The Delphi Method

The Delphi method is a way of facilitating anonymous communication between selected experts in a field, which otherwise would be impossible for geographical and other reasons (Coates, 1974). The purpose of the discussion is to reach a consensus about a certain problem or question by corroborating information. The discussion is facilitated by the researcher and usually consists of more than one round. After the panel of the experts for the study is selected, the researcher designs a questionnaire to be sent to the members of the group. Upon the receipt of their responses, another questionnaire, which includes all the opinions in anonymous form,

is created and sent to the group once again. This step may be repeated several more times for clarification (Coates, 1974).

The Delphi method was first created and used for a RAND project in 1950s. In its foundation, there was a statement that said that in the areas where the scientific laws of development had yet to be defined, the judgment of the experts would be valid for forming a prediction. This type of research received its name after an ancient Greek oracle, who, according to the legend, was able to speak to the gods and predict the future. The creators of the method, however, did not like this name since it implied something unscientific (Linstone, 1978).

Although criticized by some scholars, the Delphi method continues to be used and has been proven to be efficient in a large number of studies (Linstone, 1978). For example, among some successful studies, Linstone (1978) lists applications of the Delphi method for studies in business (i.e., TRW, Inc., Bell Canada) and government (National Materials Advisory Board, Michigan Sea Grant Program).

The following steps in the Delphi method are recommended by Linstone (1978).

1. Formation of a team to undertake and monitor a Delphi on a given subject
2. Selection of one or more panels to participate in the exercise. Customarily, the panelists are experts in the area to be investigated
3. Development of the first round Delphi questionnaire
4. Testing the questionnaire for proper wording (e.g., ambiguities, vagueness)
5. Transmission of the first questionnaires to the panelists

6. Analysis of the first round responses
7. Preparation of the second round questionnaires (and possible testing)
8. Transmission of the second round questionnaires to the panelists
9. Analysis of the second round responses (Steps 7 to 9 are reiterated as long as desired or necessary to achieve stability in the results.)
10. Preparation of a report by the analysis team to present the conclusions of the exercise. (p. 274-275)

Finally, from technological predictions, the Delphi method has evolved into being applicable for most areas of human knowledge. As Coats (1978) argues, it “is applicable in any situation to which quantitative values may be assigned, whether these are dates, weightings, or scalings. Consequently, it is applicable to most any area of inquiry” (p. 406).

Reasons for Selection of the Delphi Method for Present Study

The following factors led to the application of the Delphi method for the current research study:

- CALL is a relatively new discipline with no firmly established theoretical tenets. Therefore, the opinions of experts in the field need to be accepted for the evaluation of CALL materials.
- Quantitative data on the subject can be gathered and later analyzed by applying a Likert-type scale to the research design (Linstone, 1978)
- There is a number of reliable professional groups in the field, which enables the researcher to identify a population of experts most suitable for this study and current research objectives.

Therefore, the application of the Delphi method for the present study appears to be justified.

Selection of the Subjects

The panel of experts selected for the subjects of present study was formed from the members of the Board and Council of International Association for Language Learning Technology (IALLT). Established in 1965, IALLT is a professional organization whose members provide leadership in the development, integration, evaluation, and management of instructional technology for the teaching and learning of language, literature, and culture. Its strong sense of community promotes the sharing of expertise in a variety of educational contexts (The International Organization for Language Learning Technology Web site).

In the words of its membership director,

IALLT is a professional organization dedicated to promoting effective uses of instructional technology for language learning, teaching, and research.

Professionals in the field of language learning technology, including language media center directors and FL instructional technologists, require knowledge from a broad range of disciplines. IALLT facilitates the exchange of information in areas such as, instructional media design and development, curricular integration, needs assessment, learning center layout, hardware and software evaluation, funding and grants, personnel, copyright, cataloguing and lab management. IALLT is also a meeting ground for professionals with overlapping concerns: language teachers, librarians, lawyers, architects, and

others (The International Organization for Language Learning Technology Web site, 2003).

Thirty-seven experts chosen for the panel were identified by the researcher to participate in required at least two rounds of surveys.

The Instrument

The instrument for Round One (Appendix C) was developed by the principal investigator, based on the gathered criteria and its organization into five thematic groups and eight subgroups. The format for the survey was developed by reviewing the other examples of Delphi studies. It was constructed with the following structure: questions were presented in corresponding groups and subgroups. Each subgroup (in Groups One and Two) or group (in Groups Three, Four, and Five) of questions was followed by a space provided for expert comments. In accordance with the traditional Delphi design, each criterion on the list was followed by a five-point Likert scale for rating by the respondents. The following Likert scale was used in the present research:

1 = (NI) Not Important

2 = (SI) Somewhat Important

3 = (I) Important

4 = (VI) Very Important

5 = (E) Essential

The instrument was preceded by a clear set of instructions for the respondents. The criteria were offered to the subjects in thematically organized groups and subgroups. In order to provide clear and logical organization of the criteria, it was

decided to offer one subgroup (in case of Groups One and Two) and only one group per page of the survey. The criteria were numbered within a group/subgroup. Each criterion was followed by a described above five-point Likert-scale for evaluation. The panelists were asked to choose one of the values on scale. At the end of each group or subgroup, the respondents were given blocks of space (fields) to enter their own comments. The Delphi survey was designed in electronic format for access and completion on the World Wide Web.

Procedures

The following procedures were followed by principal investigator:

1. A printed version of the Delphi survey was designed by the principal investigator and approved by the researcher's doctoral committee members.
2. Committee members also approved the survey for on-line distribution.
3. The on-line version of the Round One survey was designed by the researcher with the help of SPSS Data Entry Builder 3.0 (part of the SPSS Suite program package). Each subgroup and group of the survey was presented on a separate web page with convenient navigation between the pages (NEXT PAGE and PREVIOUS PAGE buttons). A Likert scale was placed on the left of each criterion and was represented by a set of five radio buttons, each with a corresponding legend. Empty box fields were placed on the bottom of each web page for the respondents' comments. The final page of the survey contained a larger field for final comments as well as a SUBMIT ANSWERS button.

The online Round One survey was uploaded to the specifically allocated space on the official University of Tennessee Web server. The researcher was provided with the URL. Once the researcher tested the survey in various browsers and accounted for all types of data input, the online Round One survey was considered ready for use.

An introduction letter explaining the research study was e-mailed to the participants of the study with the link to the Delphi survey URL. Upon activation, the survey could be opened in a separate window and was intended for immediate or continuous completion and submission on line. It was determined that neither viewing, nor interaction with the electronic survey required additional downloads. Prior to contacting the subjects with the Delphi Survey, a pilot study was conducted.

Pilot Study

A preliminary form of the survey was submitted to the members of the researcher's doctoral committee, who suggested a few minor changes. After that, a pilot study was conducted with selected experts in the field whose professional status is similar to the one of the selected panel of experts. The small population selected for the pilot study consisted of professionals specializing in Computer-Assisted Language Learning in an academic institution. Like the members of IALLT Board and Council, one of them was a director of a University Language Resource Center, while the other two were World Language professors heavily involved into applications and development of CALL for their classes.

E-mail letters were sent to the pilot study participants, inviting them to partake in the research. As in the final copy of invitation letter to be sent out in Round One,

each e-mail included the link to the online survey. The participants were asked to complete and submit the survey as well as offer their critiques of the survey and recommendations for possible changes and improvements.

Three out of three responses were received from the respondents. In general, they were positive in their comments about clarity and readability of the Round One survey. Several suggestions and some minor corrections were given and respective changes were subsequently made to the instrument, which is believed to have been improved in its clarity and organization.

Data Collection

Prior to sending Round One survey and in order to improve the return of the survey, the researcher attempted to contact the panel members via phone. She was able to have phone contact with 24 panelists out of the total 37 and leave voicemail messages for the rest of the participants. In a brief message, the researcher gave a self-introduction and described the nature and objective of the study. The rationale for the panel selection, benefits of the study, and the importance of participation were given. Out of the above-mentioned 24 subjects, 16 promised their definite participation, 7 considered participation, provided they had time in their schedules, and 1 declined participating due to the fact that the subject of the research did not relate to her interests and field of expertise.

On January 29, 2003 e-mail letters of invitation (Appendix A) were set to the remaining 36 panel members. The letters reiterated the previous contact by phone, and restated the topic, goals, and nature of the research, as well as listed its general benefits to the area of study and some personal advantages to the panel of experts.

The e-mail letters contained the link to the URL of the survey website, which once opened in a separate window, was easily completed by mouse clicking on the radio buttons and typing short comments in the specially provided and clearly marked fields. The respondents were given a suggested deadline of one week to complete the questionnaire and submit their answers.

Three days after sending the initial e-mail letters, a reminder was sent to the panelists via e-mail. Ten days after the first invitation letters were sent, the results were extracted from the server. Round One of the survey was completed with 19 sets of valid responses. Four other panelists contacted the researcher via e-mail and declined participation due to reasons of health or demanding professional involvements. The data from Round One survey were gathered and analyzed in the SPSS 11.0 program package. Descriptive statistics were applied to the Round One results: means and standard deviations were calculated for each criterion as well as for all groups and subgroups.

Based on the results of the Round One survey several modifications were made in the Round Two survey. First, following recommendations, offered by panel members, several statements on the survey were re-formulated in a more precise manner. Secondly, all criteria with mean ratings below 3.95 (3 = Important, 4 = Very Important) in the Round One survey were eliminated from the Round Two survey. Added and eliminated criteria for each group and subgroup are presented in Table 1.

The necessary updates were made to the instrument and the Round Two Survey was uploaded to the website and tested by the researcher. On February 27, 2003, e-mail letters (Appendix B) with the link to the Round Two survey (Appendix

D) were sent to the panelists. The deadline of one week was given for the completion of the survey by the panel of experts. On March 1, 2003, the panel was contacted with a reminder. Ten days after the beginning of Round Two, the results of 19 sets of completed data were gathered and subsequently analyzed.

Analysis of Data

The data gathered in Round One and Round Two of the Delphi survey were statistically analyzed for the following purposes:

1. to statistically examine the degree of importance given to each criterion as well as to the thematic groups and subgroups by the panel of experts.
2. to statistically examine the presence or lack of consensus of expert opinion about each criterion as well as thematic groups and subgroups.
3. to measure the internal consistency rating of this study.

The results of the study were statistically analyzed by means of the SPSS 11.0 program package. First, the mean scores were calculated for each criterion on the list, for each subgroup of criteria within a bigger group, and, finally, for each of the five major groups. An analysis of mean scores was conducted in order to determine the degree of importance of each of the criteria as well as each thematic group and subgroup. Thus, a total of 79 criteria were presented to the panel of experts for their evaluation. The calculated means of the result scores ranged from 2.39 (the smallest) to 4.84 (the largest).

After the completion of the Round One survey, a decision was made by the researcher, in agreement with her dissertation chair, to have a cut-off point for the mean scores of 3.95 (3 = Important and 4 = Very Important in the Likert scale used

for the research). Therefore, only the criteria that had received a mean score equal to or higher than 3.95 were applied to the Round Two Delphi Survey. This procedure was applied to address Research Question 2, targeted only criteria considered important by the panel. (Table 1)

The mean scores were also calculated for each criterion in the Round Two survey as well as for each subgroup within the groups and for each big thematic group. The criteria mean scores in Round Two ranged from 2.95 (the smallest) to 4.89 (the largest). A detailed description and evaluation of the mean and standard deviation scores, as well their interpretation, are presented in the Chapter IV of this study.

The standard deviation scores were also calculated for each criterion on the survey. These scores were used to measure the level of agreement amongst the panelists on each specific item. Specifically, a larger standard deviation (equal or more than 1.00) indicates a lack of consensus, and a smaller standard deviation (of less than 1.00) indicates a consensus among the experts. In Round One, the standard deviations of the result scores ranged from 0.375 (the smallest) to 1.389 (the largest), while in Round Two they ranged from 0.315 (the smallest) to 1.335 (the largest). A detailed description and evaluation of the standard deviation results, as well their interpretation for this study, are presented in Chapter IV.

Statistical analyses of the results were also conducted for both rounds of the survey. Thus, the mean scores of groups, subgroups, as well as individual criteria in Round One were compared with the corresponding scores of the Round Two survey.

Table 1

Number Of Criteria Added and Eliminated After Round One In Each Group And Subgroup

Group/Subgroup	Number of Criteria			
	Round 1	Round 2	Deleted	Added
Group One: Pedagogical Aspect	33	15	18	0
1. Conditions for quality				
comprehensible input	8	3	5	0
2. Conditions for quality				
comprehensible output	9	5	4	0
3. Application of the principles				
of cognition	6	1	5	0
4. Motivation	5	4	1	0
5. Presentation of the content	5	2	3	0
Group Two: Applicability Aspect	21	12	9	0
1. Learner Perspective	8	5	3	0
2. Teacher Management	7	1	6	0
3. General Perspective	6	6	0	0
Group Three: Use Of Media Aspect	9	4	5	0
Group Four: Socio-Cultural Aspect	8	4	4	0
Group Five: Technical Aspect	8	6	4	2

The same procedure was applied to the standard deviations. Detailed results of this analysis are presented in Chapter IV.

Finally, appropriate statistical techniques, used to further examine the data and to measure reliability of the survey, were applied to the research. Thus, in order to measure internal consistency of ratings, reliability analyses were conducted for the final results of Round Two. Cronbach's coefficient Alpha was calculated for this purpose. The total scale reliability was high (0.8921) and N of cases = 15 (only the participants, who provided the responses to each question of the survey were counted in this calculation).

Chapter Summary

This chapter discusses methods and procedures used in this research. It describes how the criteria were identified and modified for the use of the Delphi method. First, the sources of criteria were identified by the researcher. Second, the criteria were drawn from the identified sources and included into a continuous list. Next, the criteria on the list were organized into thematic groups and subgroups and incorporated into the Round One Delphi survey instrument. The survey in electronic online format was tested in a pilot study, and after some minor changes, was presented to the selected panel of experts, members of the International Association of Language Learning Technology (IALLT), to be completed. The results of Round One were statistically analyzed, and after a number of changes to the instrument, the survey was once again presented to the panel of experts. The results of the two rounds of the Delphi study were analyzed with the application of proper statistical techniques addressing the degree of importance of each criterion, presence of

consensus among the experts, and measurement of the internal consistency rating of this study. The results of data collection and analyses of the findings are presented in Chapter IV of this research study.

CHAPTER IV

ANALYSIS OF THE DATA

Introduction

This chapter presents an analysis of the data collected in the research and discussion of the findings. The purpose of the current research was to draw language teaching software evaluation criteria from various CALL-related fields of study and construct a collective list of criteria. After presenting it for evaluation to a panel of experts, a final list from the selected items, which were deemed important by a panel of experts in the field, was developed. More specifically, the researcher attempted to answer to the following research questions:

1. What criteria for evaluation of language teaching software are identified in currently published literature, including textbooks, software reviews, and Instructional Technology and Computer-Assisted Language Learning research?
2. Which criteria were deemed important by the representatives of board and council of the International Association of Language Learning Technology?

A three-phase design was applied to the research study in order to carry out its purpose. First, the published sources related to the issue were identified by the researcher. Second, a cumulative list of criteria was developed. Third, a thematic organization was applied to the list of criteria, from which the repetitious criteria were removed from the master list and the remaining items were organized into thematic groups and subgroups according to the subject or theme they were addressing. Following this process, a final list of evaluation criteria was integrated into the Delphi

survey by applying a five-point Likert scale to each criterion and adding a comment field after each group and subgroup of the criteria. Finally, two rounds of the Delphi survey were consequently conducted with a pre-selected panel of experts (members of the Board and Council of the International Association of Language Learning Technology) for their evaluation, with Round Two survey reflecting the changes initiated by the results of the Round One survey.

The examination and statistical analyses of the data, as well as the discussion of the findings are presented in this chapter. In order to address the purpose of the study and answer the research questions, this chapter is organized into the following sections: Results of Phases One and Two, Results of Phase Three – Delphi Study, Discussion of the Findings, and Chapter Summary.

Results of Phases One and Two

Selection of Sources

In the beginning of the study, the researcher identified the potential sources for evaluation criteria of language teaching software. In order to include the approaches of various disciplines, the following areas were examined by the principal investigator:

1. published educational and language teaching software reviews (i.e., reviews published in *CALICO Software Review*; *CALL@Chorus*; *Web Tools Newsletter*; *International Review of Research in Open and Distant Learning*; *International Review of Research in Open and Distant Learning EvaluTech*, the Database of *The Southern Regional Education Board*, and the official website of *Association for Computing Machinery*);

2. Instructional Technology textbooks for teachers, providing a list of criteria for evaluation and selection of educational software (i.e., Heinich, R., Molenda, M., Russell, J., & Smaldino, S. (1999); *Instructional media and technologies for learning*. (6th ed.); and Bitter, G., & Pierson, M. (1999); *Using technology in the classroom* (4th ed.));
3. research articles in CALL and IT areas discussing issues related to evaluation of educational and language teaching software (i.e., research articles in *CALL-EJ Online*, *ReCALL Newsletter*, and *Language Learning & Technology* journals).

In the process of selecting published sources for the study, the researcher noticed the difference in approaches offered by different resources. While CALL-related resources mostly concentrated on software characteristics related to SLA or a particular level of difficulty involved in software execution and application, technology-related sources addressed various technical aspects, including compatibility, system and extra software requirements, and download time. Finally, Instructional Technology textbooks for teachers have also emphasized some pedagogical items, such as relativity to the curriculum and teaching/learning objectives, motivation, and sequence of material presentation.

Developing the List of Criteria

The following principle was applied to the design of the criteria list. First, the preexisting lists were stripped of the titles and combined in one continuous list. Then, the research articles on language teaching software from the previously mentioned sources were analyzed by the researcher and every quality mentioned by the authors

was identified, formulated as a criterion, and added to the previous list. Next, repetitious criteria were eliminated from the list and the remaining items were separated into various groups in accordance with the field of study identified for each criterion. Thus, a list of 79 criteria gathered in five different groups was created. Furthermore, the first two thematic groups, Pedagogical Aspect and Applicability Aspect, were broken down into several subgroups, which addressed more specific issues within each major thematic group.

The creation of the list of evaluation criteria addressed Research Question 1. Therefore, at this point in the research, major sources of educational software evaluation criteria were identified; the criteria were extracted from the sources, and, after sorting and classifying them, converted into a master list. From this list, the researcher incorporated collected criteria into the modified Delphi study, the results of which are described in the following sections.

Results of the Research Study Phase Three, Delphi Study

The third phase was designed to address Research Question 2. In this phase, a modified Delphi study was conducted in order to obtain the opinions of the panel of experts in the area of Computer-Assisted Language Learning in general, and language teaching software in particular. Two rounds of the Delphi study were conducted, which produced the data which was considered sufficient to produce a final list of evaluation criteria for language teaching software, deemed important by the experts of the field. The panel of experts consisted of 19 panelists in each of the two rounds. (The original panel consisted of 37 experts, 18 of whom chose not to participate in the study based on the reasons discussed in the previous chapter.) All the participants

were considered experts in the field, based on their leadership role (members of the Board and Council) in International Association of Language Learning Technology (IALLT) and their professional experience (directors of the language learning laboratories in various colleges and universities in the United States and Canada).

Round One

A total of 79 criteria were used to compose the instrument for Round One, consisting of five groups and eight subgroups (five and three subgroups in Groups One and Two, respectively), organized by the researcher in accordance with the themes addressed by the criteria. The return rate was 51.4% for Round One survey. In this round, the mean scores for evaluation of importance for each criterion ranged from 4.84 (the largest) to 2.39 (the smallest) on a five-point Likert type scale. The standard deviations of the scores in this round ranged from 0.375 (the smallest) to 1.389 (the largest).

Based on the results received in Round One, several changes were made to the initial survey before developing Round Two. First, in accordance with the comments made by the experts, one criterion was re-worded for clarification purposes. Secondly, two additional criteria suggested by the panelists were added to the list. Finally, the criteria with mean scores less than 3.95, a total of 38 criteria, were eliminated from the list. The score of 3.95 was selected since the data distributed themselves in a natural distribution such that 3.95 was a logical cut-off point.

Re-Wording

In Round One 3 out of 19 respondents mentioned the ambiguity in understanding the first criterion in the first subgroup of Group One in the original

survey. The initial statement was as follows: *Language teaching software should use mostly authentic, non-simplified language (some exceptions, such as slower rate of speech for the novice learners may apply)*. According to those panelists, the statement in parenthesis was “somewhat contradictory to the main statement” and, therefore, needed to be rephrased. Thus, in Round Two the same statement was rephrased in the following way for Round Two: *Language teaching software should mostly authentic, non-simplified language for those students with above novice level of proficiency in the target language*.

Adding extra criteria

In accordance with the suggestions of two panel members, the following two criteria were added to the survey: *Language teaching software should offer an affordable multi-user price* and *Language teaching software should be workable with other software (including language teaching and non-language teaching software) in a networked environment*. Those criteria were added to Group Five: Technical Aspect of the original survey.

Eliminating criteria

One of the goals of the current research, which was embedded into the selection of the research design, was reaching a consensus by the panel of experts in the field (Linstone, 1978). Another goal, addressed by the Research Question 2, was to identify the important criteria for evaluation of language teaching software. Therefore, the principal investigator undertook the following steps in order to fulfill the goals. First, although reaching consensus was of outmost importance for this study, it was suggested by the Research Question 2 that, in fact, it was most

significant to reach consensus about the criteria, which were viewed as being most important by the subjects. Therefore, based on these reasons, the principal investigator eliminated the less important criteria from the survey at this phase of the study.

Thus, the criteria with mean results under an arbitrary elimination point of 3.95 (3 = Very Important; 4 = Essential) were eliminated from the survey before constructing Round Two. Table 2 represents the criteria which were eliminated from the Round Two survey.

Round Two

The second round of the Delphi study was sent to each expert of the selected panel to rate the remaining criteria based on the collective input from Round One of the study. A total of 19 responses were received from Round Two for a return rate of 100%. After Round Two was conducted, a decision was made by the researcher, in concert with her committee members and statistician, that since there were no major changes in the results, the study would be limited to two rounds. In Round Two the mean scores ranged from 2.95 (the smallest) to 4.89 (the largest) on a five-point Likert-type scale. The standard deviation ranged from 0.315 (the smallest) to 1.335 (the largest). The large standard deviation (equal to or more than 1.00) indicated a lack of consensus, while the small standard deviation (less than 1.00) indicated the presence of consensus among the experts of the selected panel. The results received in Round Two addressed Research Question 2.

Table 2

Criteria Eliminated After Round One

Criteria by Group/Subgroup	Mean
Language learning software should...	
Group One: Pedagogical Aspect	
1. Conditions for quality comprehensible input	
include information on syntax of L2.	3.84
present the content from a variety of perspectives.	3.68
use correct grammar, spelling, and sentence structure.	3.68
offer communication patterns applicable in a variety of situations (i.e., in formal and informal communication, in communication with people of different age group, professional and social statuses, etc.)	3.32
use examples based on content.	3.21
2. Conditions for quality comprehensible output	
offer feedback as a separate summary of the performance for each part, level, or task.	3.89
give opportunities for multiple practice/use of newly acquired language skills.	3.84
provide open-ended activities.	3.68
offer activities requiring problem solving.	3.16
3. Application of the principles of cognition	
balance structural and integrative approaches.	3.84

Table 2

Continued

Criteria by Group/Subgroup	Mean
Language learning software should...	
relate new material to previous knowledge.	3.78
require in-depth processing (i.e., association, comprehension, and generalization).	3.67
encourage students to be strategic learners.	3.44
involve higher-order thinking skills.	3.39
4. Motivation	
develop independent learning.	3.47
5. Presentation of the content	
follow logical sequence in content arrangement (i.e., present material in a correct order of difficulty, keep continuity and connections between themes and subjects, etc.).	3.89
be adaptable in a variety of communicational situations with minimum modifications.	3.68
fit the target audience.	3.68
Group Two: Applicability Aspect	
1. Learner Perspective	
keep the same level of difficulty within an activity.	3.84
provide good and easy instruction.	2.84

Table 2

Continued

Criteria by Group/Subgroup	Mean
Language Learning Software Should	
allow student tracking.	3.21
allow comparison of student progress.	3.11
3. General Perspective	
Group Three: Use Of Media Aspect	
have content-related multimedia only.	3.89
use multimedia, which does not distract from content.	3.83
have well thought-out organization of multimedia.	3.83
present visually appealing multimedia.	3.35
offer a variety of multimedia.	3.06
Group Four: Socio-Cultural Aspect	
relate to the learners' life.	3.53
encourage intercultural communication.	3.58
provide sources of additional information on L2 countries and cultures.	3.37
avoid diminishing stereotyping (i.e., race, sex, gender).	3.16
Group Five: Technical Aspect	
load easily and without error.	3.53
present appropriate learner's features (i.e., thesaurus, spell check, URLs with additional information, note taking option).	3.84

Table 2

Continued

Criteria by Group/Subgroup	Mean
Language learning software should...	
offer “save” and “return” options from any point of an activity.	3.84
be cost effective.	3.11

Data results from Round Two survey are summarized and presented in Tables 3 -14. Criteria shown in these tables are listed in the same order they were listed on the survey instrument; therefore, no significance can be attached to the order of criteria. Each table provides mean and standard deviations scores of evaluation rating importance given to a list of criteria, by the experts on the panel. The results presented and discussed below will be supported, where possible, by comments made by the panel of experts during Rounds One and Two of the survey.

Table 3 shows the ratings for the first subgroup of criteria, *Conditions For Quality Comprehensible Input*, of Group One, *Group One: Pedagogical Aspect*. Although the rating for this subgroup dropped from Round One to Round Two, it is believed that the criteria listed in it still received a rating between “Important” and “Very Important.”

The first two criteria addressing authenticity of the target language and its linguistic characteristics were skewed towards “Very Important,” while the last criterion, addressing the semantics of the L2 skewed towards “Important” rating on the five-point Likert-type scale.

Table 3

*Results from the Delphi Round Two for the Competencies in Group One:
Pedagogical Aspect, Subgroup One: Conditions for Quality Comprehensible Input*

Criteria	Mean	SD
Language Teaching Software should...		
1 use mostly authentic, non-simplified language for those students with above novice level of proficiency in the target language.	3.79	1.032
2 emphasize linguistic characteristics of the target language.	3.79	0.855
3 include information on semantics of the L2.	3.21	0.918

The high rating of authenticity of the target language of language teaching software was explained by one of the panel members, who stated that, “I believe it is important for language to be in context; cultural thematic, etc.” This indicates that, while those three criteria were not considered essential by the selected panel, at the same time, they were rated above “Important” by the panelists.

Table 4 provides the ratings for the second subgroup of criteria, Conditions For Quality Comprehensible Output, for Group One: Pedagogical Aspect. Criteria #1, 2, and 5 scored lower than the arbitrary mean of 3.95. While Criteria #1 and 5 still scored between “Important” and “Very Important,” both were skewed towards “Very Important.” Criterion #2, addressing opportunities for negotiation of the learner’s output, received a lower rating of 2.95 and, therefore is believed not to be important according to the panel. The following comments by the experts may suggest the explanation for such a significant drop in score ratings.

Table 4

*Results from the Delphi Round Two for the Competencies in Group One:
Pedagogical Aspect, Subgroup Two: Conditions for Quality Comprehensible Output*

Criteria	Mean	SD
Language Teaching Software should...		
1 require communication with the computer program and/or other users.	3.83	1.098
2 provide opportunities for negotiation of the learner's output (i.e., the learner is given an opportunity to explain his/her answer and why he/she believes it is correct in the situation).	2.95	1.224
3 provide opportunities for correction of the learner's output (i.e., in response to learner's mistake offer a hint, rule or another example in place of providing just the correct answer, etc.).	4.32	0.671
4 offer feedback as a summary of L2 performance (i.e., the number of correct and incorrect answers, score in a game, level achieved, overall time, etc.).	3.95	0.911
5 offer feedback as a summary of L2 corrections (i.e., list learner's mistakes with correction, rules, other examples, etc.).	3.74	1.046

“ It is unlikely, at least at this point in L2 software development, that the student will have an opportunity to negotiate meaning with the software – how would this be accomplished in what is essentially binary system?” Thus, although theoretically this criterion might be considered important, it might not be realistic at the present level of language learning software development to be considered as “essential” or even “very important.” Overall in this subgroup, Criteria #4 and 5 were considered essential by the panel of experts, while Criteria #1 and 5 were viewed as being very important.

Table 5 presents mean scores and standard deviations of the results in the third subgroup in Group One. After the resulting analysis in Round One, and the elimination of selected criteria (which had been rated below than arbitrary point of 3.95), there was only one criterion left in this subgroup. In Round Two, this criterion was rated below 3.95, and, therefore, is believed not to be considered essential by the panel of experts. However, its rating lies between “Important” and “Very Important” on the five-point Likert-type scale and was skewed towards the” Very Important” rating. Therefore, it is believed to be classified as being very important criteria by the panel of experts.

Table 6, presenting the rating scores of the fourth subgroup of Group One of the survey, raises the question of the importance of motivation in language teaching software. While two of the four criteria in this subgroup received scores that can qualify them as being essential for language learning software evaluation (Criteria #1 and 3), the two remaining criteria were rated significantly lower. One of the panel

Table 5

*Results from the Delphi Round Two for the Competencies in Group One:
Pedagogical Aspect, Subgroup Three: Application Of The Principles Of Cognition*

Criteria	Mean	SD
Language Teaching Software should...		
1 provide a combination of direct instruction and incidental learning.	3.74	0.733

Table 6

*Results from the Delphi Round Two for the Competencies in Group One:
Pedagogical Aspect, Subgroup Four: Motivation*

Criteria	Mean	SD
Language Teaching Software should...		
1 facilitate success.	4.16	0.765
2 refer learners to other sources for additional learning.	3.11	0.937
3 evoke interest in the language	4.58	0.692
4 foster creativity.	3.58	0.838

members also noted on Criterion #3, which addresses the issue of evoking interest in language, that it is also important to "...evoke interest not only language, but also culture."

Criterion #4, referring to creativity, was given the mean score of 3.58, which puts it slightly above the median between "Important" and "Very Important" ratings. One of the experts' comments might provide the explanation for the lower rating given in the second round: "Where possible and applicable, creativity with the language should be the goal, but not absolute."

Criterion # 2 received a mean score of importance rating, which, although identified as an important criterion, showed that it was considered neither as being essential nor very important by the panel of experts.

Table 7 shows the importance of presentation of educational content in the software. After the elimination of criteria, based on the results of Round One survey, two criteria were included into the fifth subgroup of Group One. Both of the criteria received mean scores of importance rating above 3.0 (3 = Important) from the panel of experts in Round Two, however, Criterion #2, which addressed the relativity between the content of language learning software and students' goals, received a rating above 3.95, and, therefore, was considered essential by the panel of experts. Criterion #1, referring to the integration of software into a curriculum, was rated under 3.95, and therefore, was not considered essential by the panel of experts. Comments offered by the panel members, are believed to provide an explanation for a lower rating of criteria importance in Round Two as compared with Round Two.

Table 7

*Results from the Delphi Round Two for the Competencies in Group One:
Pedagogical Aspect, Subgroup Five: Presentation Of The Content*

Criteria		Mean	SD
Language Teaching Software should...			
1	easily integrate into a curriculum.	3.89	0.994
2	relate to a student's language learning goals.	4.32	0.946

“I would modify your statement a bit – be easily adaptable to a variety of curricula, ... relate to learning goals and styles.”

“...while I still mark it [Criterion #1] important, I felt it is more important that faculty members be trained to do [the] integration themselves.”

Table 7 concludes the analyses of the results of the Round Two survey for Group One: Pedagogical Aspect. As applied to the results for the whole group, a tendency of depreciation of the mean score rating from Round One to Round Two was noted by the researcher. Thus, although in Round One 15 criteria in Group One received ratings equal to or above 3.95, only five of them remained above this arbitrary value in Round Two.

Table 8 presents the standard deviation and mean scores for the first subgroup of Group Two in the Round Two survey. This subgroup refers to the learner’s perspective of applicability issues of language teaching software. All the criteria in this subgroup received high ratings (above 4.0) by the panel of experts and, therefore, are believed to be considered essential by the panel of experts. Moreover, the experts are believed to reach consensus about all five criteria listed in this group. Thus, all

Table 8

*Results from the Delphi Round Two for the Competencies in Group Two:
Applicability Aspect, Subgroup One: Learner Perspective*

<i>Criteria</i>	Mean	SD
Language Teaching Software should...		
1 offer various activities to fit different learning types (i.e., recognition, recall, comprehension, experiential learning).	4.22	0.647
2 offer various activities to fit various learning styles (i.e., dependent/independent learning, deductive/inductive reasoning, visual-graphic/visual-textual, individual/group work).	4.05	0.911
3 require minimum technical knowledge and skills to use.	4.21	0.918
4 provide immediate access to feedback and correction.	4.28	0.752
5 be age, level, and interest appropriate.	4.53	0.612

criteria in this subgroup, as well as the subgroup as a whole, are believed to be considered highly important on the list of criteria by the selected panel of experts.

The results for the next subgroup of criteria are presented in Table 9. Just like the two previous subgroups in Group One, the second subgroup of Group Two was reduced to a single criterion after the analysis of the results of the Round One survey. However, similar to the results received in Round One, this criterion received a high rating of above 4.0 from the panel of experts, and, therefore, is believed to be considered essential on the list of evaluation criteria. It is also significant that in this case, that consensus was reached by the panel, for the criteria importance standard deviation rating of less than 1. These high ratings continue to demonstrate the high level of importance placed by experts on the aspect of language teaching software applicability.

Table 10 presents the results of the Round Two survey for the third subgroup in Group Two. All the criteria in this group received a mean score rating larger than 4.0 and, therefore, all of them were considered to be essential by the experts of the panel. Moreover, this subgroup received the highest score ratings compared with

Table 9

*Results from the Delphi Round Two for the Competencies in Group Two:
Applicability Aspect, Subgroup Two: Teacher Management*

Criteria	Mean	SD
Language Teaching Software should...		
1 be easy to access and try out.	4.37	0.761

Table 10

*Results from the Delphi Round Two for the Competencies in Group Two:
Applicability Aspect, Subgroup Three: General Perspective*

Criteria	Mean	SD
Language Teaching Software should...		
1 be easily readable.	4.32	0.671
2 provide logical and easy navigation (i.e., make it easy to return, move forward and backward, skip a step, enter and re-enter, etc. at any point of the program).	4.89	0.459
3 be easy to get in and out of the program (i.e., it does not require to go through downloads and extra sequences on opening and closing).	4.89	0.315
4 give clear instructions.	4.84	0.375
5 offer easy to use the "help" option.	4.37	1.012
6 have easily accessible user support.	4.11	0.994

other groups and subgroups. Three out of six criteria in this subgroup (#2,3, and 4) received high scores of 4.89, 4.89, and 4.84 out of 5, respectively. Those highly-rated criteria address the issues of easy navigation, operation, and instructions for the software, and are believed to be evaluated as being some of the most important on the list.

It is also noteworthy, that all the standard deviations for the criteria in subgroup three are lower than 1.0 with the exception of Criterion #4, which allows one to believe that the experts reached consensus about every item in this subgroup with the exception of the latter. It is also important to note that this subgroup was the only one that had no eliminated criteria from Round One to Round Two.

The results of the analysis of the third subgroup conclude the series of results for Group Two. Unlike all other groups on the Delphi survey, each criterion selected for Round Two of the survey received a rating larger than the arbitrary value of 3.95 on a five-point Likert-type scale. Therefore, each criterion in this group qualified as being essential on the list of language teaching software evaluation criteria.

The level of agreement among the experts is also noteworthy for Group Two. Out of 12 criteria of Group Two, only one received a standard deviation rating larger than 1.0, which indicates a lack of consensus. Therefore, it was suggested by the experts that applicability aspect was of the highest importance on the list of evaluation criteria for language teaching software.

Groups Three, Four, and Five of the Delphi survey contained fewer criteria than Groups One and Two and, therefore, were not divided into subgroups. Thus, Tables 11 - 13 present the results for each of the remaining groups.

The results presented in Table 11 describe the use of the media aspect of language teaching software evaluation. The criteria in this group received a lower rating in Round Two compared with the results of Round One. Two out of four criteria (Criteria #2 and 4) in the group received a rating equal to or larger than the arbitrary mean value of 3.95. Sound and video components were preferred by experts over still pictures and animation to support the content of language teaching software. While the presence of sound and video were deemed essential criteria for software evaluation, the ratings received by animation and still pictures suggest that although considered important, those types of multimedia use would not be considered essential by the panel.

At the same time, the standard deviation results demonstrate that the panelists have reached agreement on Criteria #1, 2, and 3, while they disagreed in their evaluation of the role of animation (Criterion #3) in language teaching software.

Table 11

Results from the Delphi Round Two for the Competencies in Group Three: Use Of Media Aspect

Criteria	Mean	SD
Language Teaching Software should...		
1 use still pictures to support content.	3.74	0.933
2 use video to support content.	3.95	0.970
3 use animation to support content.	3.26	1.240
4 use sound to support content.	4.42	0.838

Although the ratings for this subgroup show a tendency towards use of multimedia as being an important criterion, some reservations that were stated by the experts in their comments were:

“...where appropriate should be added to each of the four items above.”

“If it makes sense to use media – “essential” is the right answer. If it does not – “not important” is the right answer.”

“It depends on the purpose of the software.”

“If these media are appropriate...”

Table 12 addresses the results of Group Four of the survey. The criteria in this group were rated high by the panel members with the exception of Criterion #3. The panelists rated high the importance of cultural illustrations, multiple cultural perspectives, and a non-discriminatory approach towards different social and cultural groups (Criteria #1,2, and 4 respectively), however, they voiced their doubts about the significance of examples of various accents and dialects (Criterion #2). Yet at the

Table 12

Results from the Delphi Round Two for the Competencies in Group Four: Socio-Cultural Aspect

Criteria		Mean	SD
Language Teaching Software should...			
1	provide authentic examples and illustrations from the L2 culture.	4.26	0.806
2	contain a variety of cultural perspectives.	4.00	0.882
3	give examples of different accents and dialects of the L2.	3.42	1.017
4	be non-discriminatory to social and cultural groups.	4.21	0.855

same time, the standard deviation for the latter criterion demonstrates a lack of consensus about the issue.

The following reasons for underestimation of Criterion #2 were offered by the panelists in their comments:

“I believe that it is important to expose students to various dialects and accents, but it is not that important in an individual piece of software.”

“The usefulness and effectiveness of providing examples of different accents and dialects of the L2 in a particular software program is debatable. In the beginning stages an L2 learner may be more confused and make less progress.”

Finally, Table 13 presents the results of ratings for the criteria included in Group Five: Technical Aspect. This group demonstrates a variety of ratings in both mean scores and standard deviations. Three criteria, Criteria #1, 4, and 3, were considered essential by the panel of experts. These criteria address the availability for both platforms, convenient “help” option, and affordability of the software. As the same time, the issues of extra downloads, detailed manual (“Skip a printed manual. Put necessary info in the help files”), and networking options, while still considered important, were not included on the list of essential evaluation criteria for language teaching software.

The lack of agreement between the experts is also notable for this group of criteria. Thus, 4 out of 6 (criteria #1,2,3, and 6) in this group received a standard deviation score equal to or larger than 1.00, which indicates a lack of consensus among the panelists.

Table 13

Results from the Delphi Round Two for the Competencies in Group Five: Technical Aspect

Criteria	Mean	SD
Language Teaching Software should...		
1 be available in both platforms.	4.06	1.211
2 not require downloading extra supporting software and/or plug-ins.	3.67	1.029
3 have a clear and detailed manual.	3.39	1.335
4 have a "help" option.	4.22	0.943
5 offer an affordable multi-user price.	4.33	0.686
6 be workable with other software (including language teaching and non-language teaching software) in a networked environment.	3.78	1.215

Table 14 presents the aspects which were considered the most significant for language teaching software evaluation by the panel of experts. The group means are provided to demonstrate the importance of each group, while ranking and ordering groups from the one with the largest mean to the one with the smallest suggests an order in which the criteria were considered for evaluation in accordance with the opinion of the selected experts.

The table demonstrates that the applicability aspect of language teaching software evaluation was rated as the most important one, followed by socio-cultural, technical, and use of media, while the pedagogical aspect was considered the least important in language teaching software evaluation. Although all of the five group mean scores were considered important on the adapted five-point Likert-type scale, only groups Two, Four, and Five were considered essential in accordance with the adopted arbitrary mean score value of 3.95.

Table 14

*Comparison of Mean Scores Between the Groups
(the order of group listing is based on their mean ratings for Round Two - from highest to lowest)*

Order	Group	Group Mean
1	Group Two: Applicability Aspect	4.42
2	Group Four: Socio-Cultural Aspect	3.97
3	Group Five: Technical Aspect	3.90
4	Group Three: Use Of Media Aspect	3.84
5	Group One: Pedagogical Aspect	3.80

Table 15 lists criteria within groups and subgroups most highly rated by the same time, the issues of extra downloads, detailed manual (“Skip a printed manual. Put necessary info in the help files”), and networking options, while still considered important, were not included on the list of essential evaluation criteria for language teaching software.

The lack of agreement between the experts is also notable for this group of criteria. Thus, 4 out of 6 (criteria #1,2,3, and 6) in this group received a standard deviation score equal to or larger than 1.00, which indicates a lack of consensus among the panelists.

experts of the panel. (As was mentioned earlier, for a criterion to be essential, it had to have a mean score equal to or above 3.95). These criteria were highly evaluated by the experts and considered as essential language teaching software evaluation criteria. It is important to note that these essential criteria are not listed in accordance with their mean values in Table 15. Instead, they appear in the order of their presentation on the instrument. This order and presentation should assist the reader in visualizing the essential criteria of evaluation of language teaching software.

The list includes 23 criteria out of total of 41 criteria presented in the second round of the study. Therefore, the remaining 17 competencies were considered important, but not essential for language teaching software evaluation.

Discussion of the Findings

In summary, based on the data collected in the study, it can be concluded that a number of aspects, including, pedagogical, applicability, use of media, socio-

Table 15

The List of Evaluation Criteria for Language Teaching Software Considered Essential by the Panel of Experts of IALLT Board and Council

Criteria by Group/Subgroup	Mean (SD)
Language teaching software should...	
Group One: Pedagogical Aspect	
2. Conditions for quality comprehensible output	
provide opportunities for correction of the learner's output (i.e., in response to learner's mistake offer a hint, rule or another example in place of providing just the correct answer, etc.).	4.32 (0.67)
offer feedback as a summary of L2 performance (i.e., the number of correct and incorrect answers, score in a game, level achieved, overall time, etc.).	3.95 (0.91)
4. Motivation	
evoke interest in the language	4.58 (0.69)
5. Presentation of the Content	
relate to a student's language learning goals.	4.32 (0.95)
Group Two: Applicability Aspect	
1. Learner Perspective	
offer various activities to fit different learning types (i.e., recognition, recall, comprehension, experiential learning).	4.22 (0.65)

Table 15

Continued

Criteria by Group/Subgroup	Mean (SD)
Language teaching software should...	
offer various activities to fit various learning styles (i.e., dependent/independent learning, deductive/inductive reasoning, visual-graphic/visual-textual, individual/group work).	4.05 (0.91)
require minimum technical knowledge and skills to use.	4.21 (0.92)
provide immediate access to feedback and correction.	4.28 (0.75)
be age, level, and interest appropriate.	4.53 (0.61)
2. Teacher Management	
be easy to access and try out.	4.37 (0.76)
3. General Perspective	
be easily readable.	4.32 (0.67)
provide logical and easy navigation (i.e., make it easy to return, move forward and backward, skip a step, enter and re-enter, etc. at any point of the program).	4.89 (0.46)
be easy to get in and out of the program (i.e., it does not require to go through downloads and extra sequences on opening and closing).	4.89 (0.32)
give clear instructions.	4.84 (0.38)
offer easy to use the "help" option.	4.37 (1.00)

Table 15

Continued

Criteria by Group/Subgroup	Mean (SD)
Language teaching software should...	
Group Three: Use Of Media Aspect	
use sound to support content.	4.42 (0.84)
use video to support content.	3.95 (0.97)
Group Four: Socio-Cultural Aspect	
provide authentic examples and illustrations from the L2 culture.	4.26 (0.80)
contain a variety of cultural perspectives.	4.00 (0.88)
be non-discriminatory to social and cultural groups.	4.21 (0.86)
Group Five: Technical Aspect	
have a "help" option.	4.22 (0.94)
offer an affordable multi-user price.	4.33 (0.69)

cultural, and technical aspects, are important in the evaluation of language teaching software. It is believed that thematic classification and presentation of these criteria have provided a logical way of organizing the criteria, based on the area of study from which the criteria were drawn.

The criteria listed in Group Two: Applicability Aspect, originates in Second Language Acquisition and has been given much attention in recent CALL research (Chappell, 1997, 1998; Nagata, 1997). It is mostly based on the development of Krashen's (1987) principle of Comprehensible Input and its application in language teaching software. This aspect received the highest importance rating by the panel of experts and, at the same time, was the aspect with the strongest level of consensus in the study.

The criteria listed in Group Four: Socio-Cultural Aspect had the second highest rating of importance on the list of evaluation criteria. This is the item that has recently precipitated a broad discussion in the area of Computer-Assisted Language Learning. Reflecting on the teachings of Hymes (19974) and Halliday (1989), this aspect recently attracted the attention of educators due to some controversies in educational software design (Bishop 2001; Sokolik, 2002). It was also placed high on the list of evaluation criteria by the experts in the field.

The ratings of Group Five: Technical Aspect followed the previous group. These criteria were developed by the instructional technologist (Bitter & Pierson, 1999; Heinich, Molenda, Russell, & Smaldino, 1999). The importance of these criteria have not been widely recognized in the field of Computer-Assisted Language Learning to date, however, in their recent works, some of the CALL researchers

started to address this theme (LeLoup, 1998; LeLoup & Ponterio, 1998; 1999). This group was also rated high by the panel of experts in this study.

Group Three: Use of Multimedia Aspect, received the next highest rating after Group Five. It addressed the on-going discussion in the field about the appropriateness and/or the possibility of distraction caused by multimedia use in educational purposes in general, and language teaching in particular (Kramsch & Anderson, 1999; Wood, 2001). Although opinions on this issue vary in research literature, as well as amongst the experts of the panel in this study, the importance of this aspect was rated relatively lower as compared with the previous one.

Finally, Group One: Pedagogical Aspect received the lowest importance ratings from the panel of experts. This group of criteria targeted general pedagogical principles as applied to the evaluation of language teaching software and was first identified in the works of Salaberry (1999) and Levy (1998). The low rating for this group can be due to the general nature of principles used for the criteria (i.e., several panelists commented on the “vague” nature of the statements). Although the group mean rating for this group was relatively high, these principles are in no way essential for the evaluation of language teaching software, according to the opinion of the selected panel of experts.

In conclusion, although the criteria from every group appears on the list of essential evaluation criteria, the aspects of applicability and socio-cultural aspects were considered to be most essential by the experts of the Board and Council of IALLT, while technical, use of media and pedagogical aspects were given less importance in language teaching software evaluation.

Chapter Summary

Chapter IV presents the results of the analysis of the data in this multiphase study. The chapter gives an overview of the study, provides research questions, presents the synopsis of methodology used in the research, and described the procedures employed in the analysis of the data.

The results of the first and second phase of the study were presented, and therefore, provided answers to Research Question 1. Next, findings of the third phase were presented and analyzed, providing answers to Research Question 2. Finally, the researcher synthesized and discussed the results.

The following chapter, Chapter V, will be the final chapter in this study. It will present the summary and conclusions of the study. It will also discuss the implications for the research as well as provide recommendations for further research.

CHAPTER V

SUMMARIES, CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS

Introduction

The preceding chapters of this dissertation presented a problem to be studied (Chapter I), a review of the literature pertinent to this study (Chapter II), methods and procedures used in this study (Chapter III), and an analysis of the data and the findings of the study (Chapter IV). This chapter (Chapter V) is the concluding chapter of the present study. It summarizes the study, presents conclusions, implications for educators and instructional designers, as well as provides recommendations for further study.

Summary

Due to the its period of existence and a large number and variety of disciplines involved in its formation, the area of Computer-Assisted Language Learning has not yet developed a universal approach to the evaluation of its resources, including language teaching software. At the same time the number of language teaching software resources produced and commercially distributed is rapidly growing, as well as the number of teachers and students implementing those resources into language teaching and learning. While a number of those applications described in the research literature are very successful, there are still some problems associated with the integration of technology into language teaching.

The source of those problems lies not only in the quality of the language teaching resources, which are not pedagogically or technologically appropriate for the

educational goals, but also the absence of reliable guidelines for initial implementation of the technology into language teaching and learning.

Since language learning software is one of the most commonly used means of educational technology by teachers and students, the current research concentrated on its evaluation criteria. A number of evaluation criteria lists were found in a variety of published research literature in the field. Those sources include published software reviews, Instructional Technology textbooks for teachers, and websites for various IT organizations. Additionally, a number of issues related to language learning software evaluation were frequently mentioned in the research articles on Computer-Assisted Language Learning.

Although all the above-mentioned sources address criteria for evaluation of language teaching software, each of them tends to approach the problem from a single angle instead of presenting the global picture. Thus, some sources focus on the important of the technical issues in evaluation of language learning software, while others adopt a pedagogical or second Language Acquisition approach to evaluation. It is believed, however, that in order to effectively incorporate the software into a language learning process, a broad variety of aspects needs to be considered for evaluation.

Thus, the present study has undertaken to bring together a number of items involved with software evaluation into one universal list of criteria for language teaching software evaluation. Moreover, the researcher attempted to provide an expert evaluation of the composed list, and based on this evaluation, condense the

latter list to a list of essential criteria. More specifically, the researcher sought answers to the following research questions:

1. What criteria for evaluation of language teaching software are identified in currently published literature, including textbooks, software reviews, and Instructional Technology and Computer-Assisted Language Learning research?
2. Which criteria are deemed important by the representatives of the Board and Council of the International Association of Language Learning Technology?

In order to address the purpose of the study and to answer these research questions, the researcher employed a three-phase research design. In the first phase of the study, the researcher surveyed a variety of published sources, including educational and language teaching software evaluations and evaluation guides, Instructional Technology textbooks for teachers, containing evaluation lists for instructional media, websites of professional IT and CALL organizations publishing educational software evaluation criteria, and research journal articles addressing various issues related to CALL software evaluation. In phase two, the criteria were extracted from the latter sources and compiled into a continuous list. Once the repetitious criteria were eliminated, 79 criteria were presented on the list. For the purpose of a more logical presentation, these criteria were organized into five thematic groups, depending on the aspect they represented. Due to the large number and a broad variety of subjects in the first two groups, they were further divided into smaller subgroups. In phase three, a modified Delphi study was applied in order to gather expert opinions regarding the importance of each criterion on the list and to

achieve consensus regarding language teaching evaluation criteria considered essential for software evaluation and, therefore, be applicable to the selection and design of language teaching software.

A review of the literature indicated that typically, a modified Delphi study requires a minimum of two rounds and subsequent rounds until consensus is reached or there is enough information to justify the results.

In the present study, two rounds of survey were conducted and met the above criteria. Participants in this study, experts of the Delphi panel, were directors of Language Resource Centers or Language Laboratories in colleges and universities in the U.S. and Canada, who were also members of the Board and Council of the International Association for Language Learning Technology (IALLT). Response rates of 51.4 % and 100% for Round One and Round Two, respectively, were achieved. Taking into consideration the time commitment required of the participants, these return rates can be considered moderately high. Several of the experts who contributed to this study, expressed the support for the goal of current research as well as their interest in the final findings of the study. Taking into account the time commitment and the experts' natural interest, the researcher plans on sending the participants a report of the study at its conclusion.

The data received were analyzed for each round using mean scores and standard deviations for each of the language teaching evaluation criteria identified in the study. After Delphi study responses were analyzed, there were 23 out of 79 criteria with a mean score greater than or equal to 3.95 (3.95 was selected as the arbitrary cut-off point for those criteria that were considered essential). These 23

criteria represent the essential criteria for language teaching software evaluation and, therefore, should be considered in the process of selection and designing of language teaching software.

Group One: Pedagogical Aspect was initially the largest group of criteria (33 items) and, therefore, was further divided into five thematic subgroups. Based on the mean scores received in the Round One survey, 18 criteria were eliminated from Round Two. The analysis of the results of Round Two showed that only three criteria from this group, in subgroups *Conditions for Quality Comprehensible Output*, *Motivation*, and *Presentation of the Content*, were considered essential by the panel of experts. These results were somewhat surprising for the researcher since these criteria are being actively discussed at present in CALL research literature by a number of researchers in the field.

Group Two: Applicability Aspect was initially the second largest group (19 criteria), which was divided into three subgroups. In the first round, this group received high ratings from the experts. Yet, 9 out of 21 criteria in this group received mean scores of importance rating of less than 3.95, and, therefore, were eliminated from the Round Two survey. This group was very highly rated by the experts in the second round and received the highest mean rating as well as the highest level of agreement between the panelists (indicated by the standard deviation scores). All the criteria in this group received mean scores not only above the arbitrary score of 3.95, but also above 4.0, which are equal to “Very Important” on the five-point Likert-type scale used in this study. At the same time, all the criteria in this group, with the exception of one, received standard deviation scores of less than 1.00, which indicates

a general agreement among the experts on those criteria. The only criterion whose standard deviation mean score rating was more than 1.00, was Criterion #5 from subgroup three, which stated “Language teaching software should offer easy to use “help” option.” This indicates a difference of opinion on this criterion among the members of the panel of experts.

Group Three: Use of Media Aspect, due to its homogeneous nature and a small number of competencies, was not divided into subgroups. In Round One, it combined nine language teaching software evaluation criteria. After the analysis of the results of the Round Two survey, it was determined that 5 out of 9 criteria received less than 3.95 mean scores from the panel of experts, and, therefore were eliminated from the Round Two survey. In the second round, the panel reached consensus on 3 out of 4 criteria in this group. Criterion #3 in Group Three (“Language teaching software should use animation to support content”) is the criterion on which the panel did not reach consensus. At the same time, Criterion #3 was one of the lowest rated criteria on the importance rating scale. This phenomenon can be supported by the diversity of opinions concerning the issue of the use of animation in educational purposes. While some of the researchers believe that animation can be an effective tool in presenting educational content to the learners, if its application follows certain pedagogical rules (Wood, 2001), others believe that it only distracts the learners from the content, and therefore should be used at a very minimum or not at all (Chun & Plass, 1997). Although the other three criteria (#1, 2, and 4) in this group received standard deviation scores of less than 1.00, which indicates consensus on these among the panel of experts), only two criteria #2 and 4)

received mean score ratings larger than or equal to 3.95, and therefore were considered to be essential by the panelists.

Group Four: Socio-Cultural Aspect consisted of eight criteria in Round One and was not divided into subgroups. After calculations of the results of the first round, four criteria were eliminated from the current group due to their mean score ratings of less than 3.95. In the second round, a consensus among experts was reached on 3 out of the 4 remaining criteria. The criterion receiving a standard deviation mean score larger than 1.00, which indicated a lack of consensus, was Criterion #3 (“Language teaching software should give examples of different accents and dialects of the L2”). The experts’ comments confirm the difference in opinions on the subject: while some of the experts believed that their students should be prepared for the diversity of a language, others suggested that such an exposure at early stage can cause confusion and misunderstanding. This criterion was also the only one which did not receive a mean score above 3.95, and, therefore, was not considered essential by the panel of experts. All of the remaining criteria (#2, 3, and 4) received score ratings above “Very Important” on the Likert-type scale, and, therefore, were included by the researcher in the list of essential criteria.

Finally, *Group Five: Technical Aspect* contained eight criteria in the Round One Survey. After the analysis of the results of Round One, two of the criteria were eliminated due to a mean score rating of less than 3.95. At the same time, two criteria were added as a result of the experts’ suggestions in Round One. In Round Two, only 3 out of 6 criteria received a mean score ratings above the arbitrary cut-off point of 3.95 (Criteria #1, 4, and 5). All three of the latter criteria address the issues related to

the availability and affordability of the software, which are believed to be important to the experts. Consensus was reached about Criteria #4 and 5, while there was disagreement about Criteria #1. The other three criteria (#2, 3, and 6), although considered to be important by the experts, did not receive a high enough mean score rating on the importance evaluation Likert-type scale to place in the category of “essential” (mean scores equal to or larger than 3.95).

Finally, although all of the initial groups are represented on the final list of essential criteria, according to the panel of experts, the most important group was *Group Two: Applicability Aspect* and the least important group was *Group One: Pedagogical Aspect*. This phenomenon is confirmed by the current research, where applicability issues related to Second Language Acquisition are the most broadly discussed ones. At the same time, it was a surprise to the researcher that the second highest rating was given to socio-cultural issues, which are still not discussed much in the research literature. Finally, the lower rating of the pedagogical aspect can be explained by the general nature of its principles, indicated by the researchers in their comments.

After two rounds of eliminations, the final list consists of 23 criteria out of an initial 79 found in current research literature, which is equal to approximately 31% of the initial number. The process of elimination was found to be productive by the principal investigator, based on the following reason: a list of the essential criteria allows the teachers and researchers to concentrate on the basics in software evaluation and makes the process of selection less laborious.

Conclusions

Based on the findings in this study, the following conclusions were drawn:

- The findings of this study provides evidence for the importance of using perspectives of multiple disciplines in evaluating language teaching software.
- Also, the findings of the study indicate both differences and commonalities among the perceptions of experts regarding the essential criteria for language teaching software evaluation.
- There was agreement among the experts of the identification of 23 criteria (representing the five thematic groups) essential in the design or selection of language teaching software. As a result, the final criteria offer useful and insightful guidance for selecting, designing, or evaluating language teaching software.

Also, it is important to note, that the conclusions in this study were drawn within the limitations and delimitations identified in Chapter 1. However, there was another limitation not mentioned above. It was the arbitrariness of the cut-off point used to make the distinction between essential and important competencies, therefore, limiting the number of criteria.

Implications

This study has provided evidence to support the incorporation of multiple areas of research into the evaluation of language learning software. Furthermore, this study was exploratory and provides an initial list of language teaching software evaluation criteria drawn from the existent research in multiple areas related to Computer-Assisted Language Learning. Therefore, language teachers and students,

as well as language teaching software designers, have evidence in this study of the importance of applying criteria of multiple disciplines to the evaluation/creation of language learning software. The experience of these professionals will determine, however, if additional criteria should require attention, based on specific target population needs.

Also, the findings of this study can be applied to the preparation of seminars, training programs, and workshops for language instructors, which seek to show how to implement computer technology into the language classroom teaching. Also, the results should be useful to those developing language teaching curricula in colleges and universities. Finally, these results might be interesting for the software designers, providing them with additional indicators of how to implement their technology skills into educational software for language teaching/learning to make it effective, applicable, and therefore marketable. Finally, from the results of this study, it is suggested that the compilation of a universal list of language teaching software evaluation from the perspectives of various disciplines should continue to be developed.

Recommendations

The following recommendations are suggested for further study:

1. A replication of the third phase, the Delphi study, is recommended to be conducted with a larger and more diverse population. Since the population of this study was limited to the directors of language resource centers, members of the Board and Council of International Association of Language Learning Technology, the results of this study were confined to this population and,

therefore, may not be generalized to the whole population of language teachers and learners, as well software designers. Thus, a similar, study should be conducted with either of the above groups or with a random sample from the joint pool of participants. In addition, the results could be analyzed controlled for technical, teaching, or learning perspectives.

2. Also, in order to obtain a theoretical approach to the evaluation process, it is recommended that this study be conducted with researchers currently publishing in the field of Computer-Assisted Language Learning.
3. Also, to give a variety of perspectives to the same study, it is suggested that it be conducted with experts from various professional organizations related to Computer-Assisted Language Learning, (i.e., language teaching associations, instructional technology professional organizations, and teacher's professional organizations).
4. It is also recommended that a follow-up qualitative/interview study be conducted with several experts in the field. This research could provide rich in-depth data for a researcher to further analyze language teaching software evaluation criteria.
5. Also, a study concentrating on one of the aspects of language learning evaluation criteria could be conducted. This type of research would provide valuable detailed data on each separate area of Computer-Assisted Language Learning.

It is hoped that this study will augment the current knowledge base of language teaching software evaluation in Computer-Assisted Language Learning, and

will stimulate further research associated with developing a multi-disciplinary approach to the evaluation of language teaching software.

Chapter Summary

The chapter presents a summary of the study, including the purpose, the literature review, methods and procedures employed to answer the research questions, findings and conclusions. Also, the implications of the research were discussed and recommendations for further study were presented.

Following this chapter is a list of References as well as Appendices, including the cover letters used to contact the participants of the study, the survey used in Round One and Round Two of the Delphi study. The researcher's vita appears at the end of the dissertation.

REFERENCES

References

- American Council on the Teaching of Foreign Languages. (1996). *Standards for foreign language learning. Preparing for the 21st century: Executive Summary*. Retrieved on October 21, 2002, from <http://www.actfl.org/public/articles/execsumm.pdf>
- Bishop, A. (2001). An experts guide to products for the multilingual classroom. *Technology and Learning*, 21(9), 39-46. Retrieved on October 21, 2002, from http://www.techlearning.com/db_area/archives/TL/200104/classroom.html
- Bitter, G., & Pierson, M. (1999). *Using technology in the classroom* (4th ed.). Needham Heights, MA: Allyn & Bacon.
- Bond, K. (2002). Perceptions of anonymity in a language learning text chat room. *CALL-EJ Online*, 3(2). Retrieved on March 10, 2002, from <http://www.clec.ritsumei.ac.jp/english/callejonline/6-2/KBond.htm>
- Cameron, K. (1999). Introduction. In K. Cameron (Ed.) *CALL: Media, Design & Applications* (pp. 1-9). Lisse: Swets & Zeitlinger.
- Chapelle, C. (1997). CALL in the year 2000: Still in search of research paradigms? *Language Learning & Technology*, 1(1), 19-43. Retrieved on January 25, 2002, from <http://polyglot.cal.msu.edu/llt/vol1num1/chapelle/default.html>
- Chapelle, C. (1998). Multimedia CALL: Lessons to be learned from research on instructed SLA. *Language Learning & Technology*, 2(1), 22-34. Retrieved on January 25, 2002, from <http://llt.msu.edu/vol2num1/article1/>
- Chapelle, C. (1999). Research questions for CALL research agenda: A reply to Rafael Salaberry. *Language Learning & Technology*, 3(1), 108-113. Retrieved on January 25, 2002, from <http://llt.msu.edu/vol3num1/comment/reply.html>
- Chun, D. (1998). Signal analysis software for teaching discourse intonation. *Language Learning & Technology*, 2(1), 61-77. Retrieved on October 21, 2002, from <http://llt.msu.edu/vol2num1/article4/>
- Chun, D., & Plass, J. (1997). Research on text comprehension in multimedia environments. *Language Learning & Technology*, 1(1), 60-74. Retrieved on October 21, 2002, from http://llt.msu.edu/vol1num1/chun_plass/default.html
- Coates, J. (1978). Technology assessment. In J. Fowles (Ed.), *Handbook of futures research* (pp. 397-422). Westport, Connecticut: Greenwood Press.

- Delcloque, P. (2000). An illustrated history of Computer-Assisted Language Learning. Retrieved on October 27, 2002, from <http://www.history-of-call.org>
- Dunkel, P. (1999). Considerations in developing or using second/foreign language proficiency Computer Adaptive Tests. *Language Learning & Technology*, 2(2), 77-93. Retrieved on March 3, 2002, from <http://llt.msu.edu/vol2num2/article4/index.html>
- Egbert, J. (2001). Active learning through computer-enhanced activities. Teaching English with Technology, *A Journal for Teachers of English*, 1(3). Retrieved on January 25, 2002, from http://www.iatefl.org.pl/sig/call/j_article3.htm
- Ehsani, F., & Knodt, E. (1998). Speech technology in computer-aided language learning: Strength and limitations of a new CALL paradigm. *Language Learning & Technology*, 2(1), 45-60. Retrieved on January 25, 2002, from <http://llt.msu.edu/vol2num1/article3/>
- Encyclopedia Encarta Web site: from <http://www.encarta.msn.com>. Retrieved on January 25, 2002, from <http://www.encarta.msn.com>
- Eskenazi, M. (1999). Using automatic speech processing for foreign language pronunciation tutoring: Some issues and a prototype. *Language Learning & Technology*, 2(2), 62-76. Retrieved on October 28, 2002, from <http://llt.msu.edu/vol2num2/article3/>
- Geissinger, H. (1997). Educational Software: Criteria for Evaluation. Retrieved October 18, 2002 from <http://www.ascilite.org.au/conferences/perth97/papers/Geissinger/Geissinger.html>
- Groot, P. J. (2000). Compute assisted second language vocabulary acquisition. *Language Learning and Technology*, 1(4), 60-81. Retrieved on January 25, 2002, from <http://llt.msu.edu/vol4num1/groot/default.html>
- Halliday, M., & Hasan, R. (1989). *Language, context, and text: Aspects of language in a social-semiotic perspective*. Oxford: Oxford University Press.
- Heinich, R., Molenda, M., Russell, J., & Smaldino, S. (1999). *Instructional media and technologies for learning*. (6th ed.). Upper Saddle River, NJ: Prentice-Hall, Inc.

- Higgins, C. (1993). Computer-Assisted Language Learning: Current Programs and Projects. ERIC Digest. Washington, DC: ERIC Clearinghouse on Languages and Linguistics. (ERIC Document Reproduction Service No. ED355835). Retrieved on October 21, 2002, from http://www.ed.gov/databases/ERIC_Digests/ed355835.html
- Hymes, D. (1974). *Foundations in sociolinguistics. An ethnographic approach*. Philadelphia: University of Pennsylvania Press.
- The International Organization for Language Learning Technology Web site: <http://iallt.org/>
- Jones, J. (2001). CALL and the teacher's role in promoting learner autonomy. *CALL-EJ Online*, 3(1). Retrieved on October 25, 2002, from <http://www.clec.ritsumei.ac.jp/english/callejonline/6-1/jones.html>
- Kasper, L. (2000). New technologies, new literacies: Focus discipline research and ESL learning communities. *Language Learning & Technology*, 4(2), 105-128. Retrieved on October 21, 2002, from <http://llt.msu.edu/vol4num2/kasper/default.html>
- Kern, R., & Warschauer, M. (2000). Theory and practice of network-based language Teaching. Retrieved March 19, 2002, from <http://www.gse.uci.edu/markw/nblt-intro.html>
- Kramsch, C. & Anderson, R. (1999). Teaching text and context through multimedia. *Language Learning & Technology*, 2(2), 31-42. Retrieved on October 21, 2002, from <http://llt.msu.edu/vol2num2/article1/>
- Krashen, S. (1985). *Inquiries and insights*. Hayward, CA: Alemany Press.
- Kung, C. (2002). A framework for successful key-pal programs in language learning. *CALL-EJ Online*, 3(2). Retrieved on March 19, 2002, from <http://www.clec.ritsumei.ac.jp/english/callejonline/6-2/SCKung.htm>
- LeLoup, J. (1998). A communications technology module for the foreign language methods course. Retrieved on October 21, 2002, from <http://www.cortland.edu/flteach/methods/>
- LeLoup, J., & Ponterio, R. (1998). Using www multimedia in the foreign language classroom: Is this for me? *Language Learning & Technology*, 2(1), 4-10. Retrieved on March 19, 2002, from <http://llt.msu.edu/vol2num1/Onthenet/>

- LeLoup, J., & Ponterio, R. (1999). Interactive language exercises on the web: An exemplar. *Language Learning & Technology*, 3(1), 3-11. Retrieved on March 19, 2002, from <http://llt.msu.edu/vol3num1/onthenet/>
- LeLoup, J., & Ponterio, R. (2000). Enhancing authentic language learning experiences through Internet technology. (Report No. EDO-FL-00-02). *Eric Clearinghouse on Language and Linguistics*. Retrieved on October 21, 2002, from <http://www.cal.org/ericcll/digest/0002enhancing.html>
- Levine, A., Ferenz, O. & Reves, T. (2000). EFL academic reading and modern technology: How can we turn our students into independent readers? *TESL-EJ*, 4(4). Retrieved on October 21, 2002, from <http://www.kyoto-su.ac.jp/information/tesl-ej/ej16/a1.html>
- Levy, M. (1997). *Computer-Assisted Language Learning. Concepts and conceptualization*. New York: Oxford University Press, Inc.
- Linstone, H. (1978). The Delphi technique. In J. Fowles (Ed.), *Handbook of futures research* (pp. 273-300). Westport, Connecticut: Greenwood Press.
- Linstone, H., & Turoff, M. (2002). Introduction. In H. Linstone & M. Turoff (Eds.), *The Delphi Method: Techniques and Applications* (p.p. 3-11). Retrieved March 3, 2003, from <http://www.is.njit.edu/pubs/delphibook/index.html>
- Marcos, K. (1994). Internet for language teachers. *Eric Clearinghouse on Language and Linguistics*. Retrieved October 21, 2002, from <http://www.cal.org/ericcll/digest/marcos01.html>
- Nagata, N. (1998). Input vs. output practice in educational software for second language acquisition. *Language Learning & Technology*, 1(2), 23-40. Retrieved March 19, 2002, from <http://llt.msu.edu/vol1num2/article1/default.html>
- O'Dowd, R. (2001). In search of a truly global network: The opportunities and challenges of on-line intercultural communication. *CALL-EJ Online*, 3(1). Retrieved on January 25, 2002, from http://www.clec.ritsumei.ac.jp/english/callejonline/6-1/o_dowd.html
- Ortega, L. (1997). Process and outcomes in networked classroom interaction: Defining the research agenda for L2 Computer-Assisted classroom discussion. *Language Learning & Technology*, 1(1), 82-93. Retrieved on January 25, 2002, from <http://llt.msu.edu/vol1num1/ortega/default.html>

- Paramskas, D. (1999). Computer-Mediated Communication. In K. Cameron (Ed.), *CALL: Media, Design & Applications* (pp. 13-29). Lisse: Swets & Zeitlinger.
- Parra, G. (2000). Learning English through online discussion groups. *Journal of Adolescent and Adult Literacy*, 44(1), 36-38.
- Plass, J. (1998). Design and evaluation of user interface of foreign language multimedia software: A cognitive approach. *Language Learning & Technology*, 2(1), 35-45. Retrieved on October 21, 2002, from <http://llt.msu.edu/vol2num1/article2/index.html>
- Poftak, A. (1999). Technology and language acquisition: An interview with Anna Bishop. *Technology and Learning*, 19(5), 56-59.
- Salaberry, R. (1999). CALL in the year 2000: Still developing the research agenda. *Language Learning & Technology*, 3(1), 104-107. Retrieved on January 25, 2002, from <http://llt.msu.edu/vol3num1/comment/>
- Selinker, G., & Gass, S. (1994). *Second language acquisition: An introductory course*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Shulman, M. (2000). Developing global connections through computer-mediated communication. *The Internet TESL Journal*, 7(6). Retrieved on October 21, 2002, from <http://iteslj.org/Articles/Shulman-CMC.html>
- Sokolik, M. (2000). Alexis, the Encomium TOEIC test preparation system. *CALL@Chorus*. Retrieved September 17, 2002, from http://www-writing.berkeley.edu/chorus/call/reviews/alexis_toeic/index.html
- Stepp-Greany, J. (2002). Student perceptions on language learning in a technological environment: Implications for the new millennium. *Language Learning & Technology*, 6(1), 165-180. Retrieved on January 25, 2002, from <http://llt.msu.edu/vol6num1/STEPPGREANY/default.html>
- Vilmi, R. (1995). Global communication through email: an ongoing experiment at Helsinki University of Technology. Retrieved on January 25, 2002, from <http://www.ruthvilmi.net/hut/autumn93/global.html>
- Williams, H., & Williams, P. (2000). Integrating reading and computers: an approach to improve ESL students' reading skills. *Reading Improvement*, 37, 98-115.

Wood, J. (2001). Can software support children's vocabulary development?
Language Learning & Technology, 5(1), 166-201. Retrieved on October 25,
2002, from <http://llt.msu.edu/vol5num1/wood/default.html>

APPENDICES

Appendix A

Round One: Letter to the Panel of Experts

Dear Mr./Ms./Dr.:

I am a Ph.D. candidate in World Languages/ESL Education and Instructional Media and Technology at the University of Tennessee, Knoxville. I am at the dissertation stage of my program. Please help me with my research by filling out my on-line survey for my dissertation.

I am inviting you, as an experienced language learning technology professional, and a member of the Board and Council of International Association for Language Learning Technology, to participate in this study. This survey is a part of a Delphi study being conducted to evaluate and reach consensus regarding the criteria for the evaluation of language teaching software. Any help that you can provide to ensure this survey is completed and submitted on line will be greatly appreciated. A HIGH PERCENTAGE OF RETURNS IS NECESSARY FOR THIS STUDY TO BE VALID.

Your expertise and knowledge shared through your participation in this study will benefit language teachers, software designers, and researchers of Computer-Assisted Language Learning. In addition, you will benefit from this study, should you be interested in either obtaining the summary or the link to my dissertation upon its completion.

You may be assured that your identity and anonymity will be maintained. All required information is for statistical purposes only. Summary data will be reported in the dissertation and publications following completion of the study. All data will be kept in a locked cabinet by the researcher for a limited period of time and then safely destroyed. Your participation is totally voluntary.

I am asking your commitment to the study by participating in the completion of this survey (Round One) and one or two additional surveys (Round Two and Round Three). The survey is to be completed electronically by going to the link on the bottom of this e-mail and following the instructions. It should not take more than 15-20 minutes of your valuable time. The results will be automatically forwarded to my database. Please complete this survey within one week (by Wednesday, February 5) to ensure that your responses will appear in Round Two of the study. I understand that as a professional, your schedule is extremely tight and demanding, and I sincerely appreciate your time and effort. If you have any questions or concerns about the study, please do not hesitate to contact me via e-mail at vsavchen@utk.edu or phone me, at (865) 936-0759.

Thank you for your time, expertise, and help. I am looking forward to hearing from you at your earliest convenience. Please submit this survey on-line no later than Wednesday, February 5, 2003 if at all possible.

IMPORTANT NOTES:

The software used to design this survey is optimized for Internet Explorer. Please, if it all possible, use this browser to open the link and work on the survey. Filling it out in Netscape will still guarantee the results, however, the graphics will not look good.

The fields on the bottom of each page take up to 255 characters and are designed for specific concise answers to the questions. Should you have any comments about the survey or the study in general, please send them to me via e-mail.

LINK TO THE SURVEY:

<https://webmail.utk.edu/redirect?http://surveys.utk.edu/victoria/index.htm>

Sincerely yours,
Victoria Savchenkova
Ph.D. Candidate in World Languages and ESL Education
The University of Tennessee

Appendix B

Round Two: Letter to the Panel of Experts

Dear Mr./Ms./Dr.:

Thank you for participating in Round One of my Delphi survey. Your participation is extremely valuable for my study. At this time I am asking you to please complete my Round Two survey.

I would like you to note that the Round Two survey has been shortened based on your responses in Round One. At this time, I believe it should take no longer than 10 minutes to fill out and submit your response. I know how extremely busy you might be at this time of the semester, and I appreciate your time, effort, and expertise. Please help me finish my research – a high return rate of replies in Round Two is vital for the study to be complete.

In order to fill out Round Two of the survey, please go to the link at the bottom of this e-mail and follow the instructions. Please, if at all possible, submit your response no later than a week from today, Friday, March 7, 2003.

Thank you again for your participation. Should you have any questions or comments, please contact me via phone at (865) 936-0759 or e-mail at vsavchen@utk.edu

I am looking forward to hearing from you.

IMPORTANT NOTES:

The software used to design this survey is optimized for Internet Explorer. Please, if it all possible, use this browser to open the link and work on the survey. Filling it out in Netscape will still guarantee the results, however, the graphics will not look good.

The box on the bottom of each page on the survey only allows a maximum of 255 characters and, therefore, your comments need to be very specific and concise to fit in the boxes. Should you have any additional comments about the survey or the study in general, please send them to me via e-mail at vsavchen@utk.edu.

LINK TO THE SURVEY: <http://surveys.utk.edu/victoria3/index.htm>

Sincerely yours,
Victoria Savchenkova
Ph.D. Candidate in World Languages and ESL Education
The University of Tennessee

Appendix C

Round One Survey Instrument

**Evaluation Criteria for Language Teaching Software:
A Delphi Study of Professional Perspectives**

Round One Survey

Directions:

Please read each evaluation criterion, then indicate your response by checking the radio button that corresponds to the category that most accurately reflects your expertise and assessment of the item. In addition, please feel free to make comments on any particular statement and/or add new competency statement(s) in the fields provided. Upon completion, please click the “Submit” button. Your response is vital for this study and I appreciate your input. Thank you for your time and your thoughts.

Please use the following legend to rate the statements:

NI = Not Important SI = Somewhat Important VI – Very Important
I = Important E = Essential

GROUP ONE: PEDAGOGICAL ASPECTS

1. Conditions for quality comprehensible input

Language teaching software should...

use mostly authentic, non-simplified language (some exceptions, such as slower rate of speech for the novice learners, etc., may apply).

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

emphasize linguistic characteristics of the target language.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

help learners comprehend semantics of the L2.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

help learners comprehend syntax of the L2.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

present the content from a variety of perspectives.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

use correct grammar, spelling, and sentence structure.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

be applicable in a variety of situations.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

use examples based on content

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

Other:

--

2. Conditions for quality comprehensible output

Language teaching software should...

require communication with the computer program and/or other users.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

provide opportunities for correction, negotiation, and repetition of the learner's output.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

offer feedback as a summary of L2 performance.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

offer feedback as a summary of L2 corrections.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

offer feedback as a summary of all the performances on the task.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

give opportunities for multiple practice/use of newly acquired language skills.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

provide open-ended activities.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

offer activities requiring problem solving.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

Other:

--

3. *Application the principles of cognition*

Language teaching software should...

provide a combination of direct instruction and incidental learning.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

balance structural and integrative approaches.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

relate new material to previous knowledge.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

require in-depth processing: association, comprehension, and generalization.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

encourage students to be strategic learners.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

involve higher-order thinking skills.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

Other:

--

4. Motivation

Language teaching software should...

facilitate success.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

refer learners to other sources for additional learning.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

evoke interest in the language.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

foster creativity.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

develop independence.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

Other:

--

5. *Presentation of the content*

Language teaching software should...

relate to the curriculum.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

relate to a student's language learning goals.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

be logically arranged.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

be adaptable in a variety of situations.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

fit the target audience.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

Other:

--

GROUP TWO: APPLICABILITY ASPECT

1. Learner Perspective

Language teaching software should...

offer various activities to fit different learning types (i.e., recognition, recall, comprehension, experiential learning).

NI ☐ SI ☐ VI ☐ I ☐ E ☐

offer various activities to fit various learning styles (i.e., dependent/independent learning, deductive/inductive reasoning, visual-graphic/visual-textual, individual/group work).

NI ☐ SI ☐ VI ☐ I ☐ E ☐

require minimum technical knowledge and skills to use.

NI ☐ SI ☐ VI ☐ I ☐ E ☐

provide immediate access to feedback and correction.

NI ☐ SI ☐ VI ☐ I ☐ E ☐

be age, level, and interest appropriate.

NI ☐ SI ☐ VI ☐ I ☐ E ☐

keep the same level of difficulty within an activity.

NI ☐ SI ☐ VI ☐ I ☐ E ☐

feature an adjustable pace of learning.

NI ☐ SI ☐ VI ☐ I ☐ E ☐

offer *save* and *return* options from any point of an activity.

NI ☐ SI ☐ VI ☐ I ☐ E ☐

present appropriate learner's features (i.e., thesaurus, spell check, URLs with additional information, note-taking option).

NI ☐ SI ☐ VI ☐ I ☐ E ☐

provide good and easy instruction.

NI ☐ SI ☐ VI ☐ I ☐ E ☐

Other:

--

2. Teacher Management

Language teaching software should...

be easy to test and assess.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

provide control for rate, sequence, and level of difficulty.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

keep a record of activities.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

allow content development.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

allow student assessment.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

allow student tracking.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

allow comparison of student progress.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

Other:

--

3. General Perspective

Language teaching software should...

be easily readable.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

provide logical and easy navigation.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

be easy to get in and out of the program.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

give clear instructions.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

be easy to use the *help* option.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

have easily accessible user support.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

Other:

GROUP THREE: USE OF MEDIA ASPECT

Language teaching software should...

use still pictures to support content.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

use video to support content.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

use animation to support content.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

use sound to support content.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

have content-related multimedia only.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

use multimedia, which does not distract from content.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

have well thought-out organization of multimedia.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

present visually appealing multimedia.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

offer a variety of multimedia.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

Other:

--

GROUP FOUR: SOCIO-CULTURAL ASPECT

Language teaching software should...

provide authentic examples and illustrations from the L2 culture.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

contain a variety of cultural perspectives.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

give examples of different accents and dialects of the L2.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

be non-discriminatory to social and cultural groups.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

relate to the learners' life.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

be age appropriate.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

encourage intercultural communication.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

provide sources of additional information on L2 countries and cultures.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

avoid stereotyping (e.g., race, sex, gender).

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

Other:

--

GROUP FIVE: TECHNICAL ASPECT

Language teaching software should...

be available in both platforms.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

not require downloading extra supporting software and/or *plug-ins*.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

have a clear and detailed manual.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

have a *help* option.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

load easily and without error.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

be cost effective.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

Other:

--

OTHER ASPECTS:

*At this point, please verify that you have checked all the categories listed above.
If you are ready to submit your survey, please click on **SUBMIT ANSWERS** button.
Thank you for your participation!*

SUBMIT ANSWERS

Appendix D

Round Two Survey Instrument

**Evaluation Criteria for Language Teaching Software:
A Delphi Study of Professional Perspectives**

Round Two Survey

Directions:

Please read each evaluation criterion, then indicate your response by checking the radio button that corresponds to the category that most accurately reflects your expertise and assessment of the item. In addition, please feel free to make comments on any particular statement and/or add new competency statement(s) in the fields provided. Upon completion, please click the “Submit” button. Your response is vital for this study and I appreciate your input. Thank you for your time and your thoughts.

Please use the following legend to rate the statements:

NI = Not Important SI = Somewhat Important VI – Very Important
I = Important E = Essential

GROUP ONE: PEDAGOGICAL ASPECTS

1. Conditions for Quality Comprehensible Input

Language teaching software should...

use mostly authentic, non-simplified language (some exceptions, such as slower rate of speech for the novice learners, etc., may apply).

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

emphasize linguistic characteristics of the target language.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

help learners comprehend semantics of the L2.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

help learners comprehend syntax of the L2.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

Other:

--

2. Conditions for Quality Comprehensible Output

Language teaching software should...

require communication with the computer program and/or other users.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

provide opportunities for correction, negotiation, and repetition of the learner's output.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

offer feedback as a summary of L2 performance.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

offer feedback as a summary of L2 corrections.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

offer feedback as a summary of all the performances on the task.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

Other:

--

3. Application the Principles of Cognition

Language teaching software should...

provide a combination of direct instruction and incidental learning.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

Other:

4. Motivation

Language teaching software should...

facilitate success.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

refer learners to other sources for additional learning.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

evoke interest in the language.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

foster creativity.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

Other:

--

5. *Presentation of the Content*

Language teaching software should...

relate to the curriculum.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

relate to a student's language learning goals.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

Other:

--

GROUP TWO: APPLICABILITY ASPECT

1. Learner Perspective

Language teaching software should...

offer various activities to fit different learning types (i.e., recognition, recall, comprehension, experiential learning).

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

offer various activities to fit various learning styles (i.e., dependent/independent learning, deductive/inductive reasoning, visual-graphic/visual-textual, individual/group work).

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

require minimum technical knowledge and skills to use.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

provide immediate access to feedback and correction.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

be age, level, and interest appropriate.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

Other:

--

2. Teacher Management

Language teaching software should...

be easy to test and assess.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

Other:

--

3. General Perspective

Language teaching software should...

be easily readable.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

provide logical and easy navigation.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

be easy to get in and out of the program.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

give clear instructions.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

be easy to use the *help* option.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

have easily accessible user support.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

Other:

--

GROUP THREE: USE OF MEDIA ASPECT

Language teaching software should...

use still pictures to support content.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

use video to support content.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

use animation to support content.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

use sound to support content.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

Other:

--

GROUP FOUR: SOCIO-CULTURAL ASPECT

Language teaching software should...

provide authentic examples and illustrations from the L2 culture.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

contain a variety of cultural perspectives.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

give examples of different accents and dialects of the L2.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

be non-discriminatory to social and cultural groups.

NI ☐ **SI** ☐ **VI** ☐ **I** ☐ **E** ☐

Other:

--

GROUP FIVE: TECHNICAL ASPECT

Language teaching software should...

be available in both platforms.

NI ☐ SI ☐ VI ☐ I ☐ E ☐

not require downloading extra supporting software and/or *plug-ins*.

NI ☐ SI ☐ VI ☐ I ☐ E ☐

have a clear and detailed manual.

NI ☐ SI ☐ VI ☐ I ☐ E ☐

have a *help* option.

NI ☐ SI ☐ VI ☐ I ☐ E ☐

offer an affordable multi-user price.

NI ☐ SI ☐ VI ☐ I ☐ E ☐

be workable with other software (including language teaching and non-language teaching software) in a networked environment.

NI ☐ SI ☐ VI ☐ I ☐ E ☐

Other:

--

OTHER ASPECTS:

*At this point, please verify that you have checked all the categories listed above.
If you are ready to submit your survey, please click on **SUBMIT ANSWERS** button.
Thank you for your participation!*

SUBMIT ANSWERS

VITA

Victoria Savchenkova was born in Kharkov, Ukraine on July 18th, 1974. She matriculated her elementary, high school, and initial higher education in Ukraine. She enrolled in G. Skovoroda Kharkov State Pedagogical University in Kharkov, Ukraine, and in 1996, graduated with the Diploma (which corresponds to an American master's degree) in Linguistics and Russian and Foreign Language Education, with a concentration in Teaching English and Russian. While studying at the university, she started working as a translator for business and religious organizations in Ukraine. She also conducted private classes in English and Russian for people relocating to other countries. In 1997, she was offered a position of Instructor/Counselor in the U.S. Space and Rocket Center, NASA's Official Visitor's Center, in Huntsville, Alabama, as a part of an international exchange program. In August of 1998, she began matriculating her Doctor of Philosophy in Education degree program with a concentration in Foreign Language Education and a specialization in Computer-Assisted Language Learning.

During her career at The University of Tennessee, she served as a graduate assistant in an administrative capacity, first, as an advising assistant in the College of Education Graduate Advising Center, and next, as a recruitment database manager with the Office of Graduate Admissions at The University of Tennessee. Her academic advising, research, and technological experience, in addition to her proficiency in several languages and in multiple computer applications, support her professional interests well.