Incorporating ESOL Learners' Feedback and Usability Testing in Instructor-Developed CALL Materials

Greg Kessler and Lia Plakans

IN THE FIELD of TESOL, a long tradition exists of instructor-developed materials. These materials appear in the forms of textbooks, video- and audiotape recordings, Web sites, and, more recently, CD-ROMs. The need for materials developed by teachers in our field is essential because these individuals are the most immediate experts on the needs of ESOL learners, the cognitive abilities of different age groups, and the language learning process of their specific learners. As developers, however, teachers need to consider creating and implementing a systematic process for ensuring the usability of their materials. In this process of evaluating materials, learners must be included, as they are also experts on their learning as well as the benefactors of well-developed materials.

This article proposes the use of *usability testing* to systematically evaluate instructor-developed materials. Usability testing is a method of observing how a user interacts with and experiences materials in order to identify those characteristics that simplify or confound the use (Dumas & Redish, 1993). We will discuss the application of this method in the field of computer-assisted language learning (CALL) using our own project, a multimedia CD-ROM that was designed to improve the skills of international graduate students at Ohio University.[1] The students we worked with in revising our materials were adult international students who had fulfilled the university's English language requirements for nonnative English speakers, but who still felt the need to improve their oral communication skills. Our particular application, however, may be adapted and applied to a range of other ESOL contexts.

The need for materials developed by teachers in our field is essential because these individuals are the most immediate experts on the needs of ESOL learners, the cognitive abilities of different age groups, and the language learning process of their specific learners.

CALL Materials Development

CALL in ESL began as a small group of individuals who attempted to use computers to enhance the traditional classroom. This integration of technology initially resulted in the development of many early CALL materials that resembled traditional print materials. Though benefiting from delivery through a system of computers, the materials offered much the same content as textbooks. With the advent of the World Wide Web and other multimedia systems of delivery, CALL materials exploded into a media-rich extravagance that some worried might jeopardize the language focus of materials. The field has matured and now offers teachers an array of technology-based pedagogical materials. CALL instructor-developers today attempt to integrate the best aspects of the attention-getting media and the language objectives contained in traditional textbooks. However, this integration of varied media can unnecessarily complicate their use.

As instructors become more familiar with commercially available CALL options, they are creating their own contextually appropriate CALL materials. Consequently, teachers are designing materials that address specific courses, contexts, and groups of students. The commitment of resources, along with the benefits of these teacher-developed CALL materials, demand careful consideration to create useful and usable materials.

Throughout the dynamic evolution of CALL materials, developers have designed materials based on a set of design principles for all software (Leshin, Pollock, & Reigeluth, 1992; Thorn, 1995). The following principles presented by Boling and Soo (1999) as "essential elements of good software design" (p. 443) provide valuable guidance:

- The interface and terminology are consistent from screen to screen.
- The layout of each screen makes good use of space.

- Legibility and readability are high.
- The software makes good use of contrast, repetition, alignment, and proximity.
- Serious navigational errors are prevented.
- Undesired actions are easily reversed.
- Audio and video playback (where applicable) are of good quality.

In addition to these design principles, instructor-developers should consider the effects of learner variables (e.g., age, language proficiency, learning styles, previous experience with computers) when designing CALL materials. Although a needs assessment may identify some potential concerns, specific computer-related assessment offers additional insights into the effects of these variables (Bishop, 1999; Ngeow, 1999; Soo, 1999) on learners' use of computer-based materials.

Rationale for Student Input

There are many reasons why learners should be involved in the development of all educational materials, particularly those incorporating computers.

- **Learners are stakeholders.** Because learners are the audience for the material, their ability to use and learn from it must be considered. In ESOL, there is also the issue of language appropriateness. By including students in the development process, course designers can reliably determine the terms, phrases, and vocabulary that learners will find accessible. In addition, most learners are very aware of what helps them learn and what does not. Thus, they are more willing to use products they have helped develop.
- **Learners are individuals.** Much attention has been given to different learning styles, strategies, and multiple intelligences. Obtaining feedback from a variety of learners provides insights into how individuals approach a particular language learning task, CALL software, and the specific content of the material.
- **Learners are often computer literate.** Today's learners typically have an increased awareness of computers, software, and their combined potential. In fact, often in academic English programs, many of our students are enrolled in technology disciplines. Including them in the process of CALL materials development not only improves the quality of the project but also empowers the students by recognizing their expertise. However, computer expertise is not a prerequisite for participating in usability tests.
- **Learners are affected by the environment.** Variations exist across computer platforms, networks, and laboratories. Given this variation, it is crucial to observe learner use in the environment in which CALL materials will ultimately be used. It is also necessary to be aware of learners' interactions with the relevant CALL environment.

Our Project

In this article, we describe our experience developing a multimedia CD-ROM that was created as a "virtual textbook" for an on-line course in oral communications to illustrate a method for incorporating learner feedback in materials developed by instructors. We began by addressing some of the issues noted by others in second language acquisition (SLA) regarding features of CALL materials (Chapelle, 1998; Plass, 1998).

First, we spent time determining language features that would be the focus of our course. To arrive at these, we considered the aspects of oral communication that were the focus of the course, and those we wanted to add to supplement the course. The language features highlighted in our CD-ROM were: listening skills (prediction, main idea, details, note taking, inference, summary), language skills (spoken grammar, idioms, communications skills, rhythm and intonation, vocabulary), and presentation skills (level of formality, nonverbal communication, audience, introduction, visual aids, connection, logic, conclusion).

Next, we considered tasks that would require learners to analyze these features. On the CD, a variety of exercises focused on oral communications skills and strategies were used to analyze clips of lectures on eight different academic topics. The exercises were divided into the three language features categories: listening skills, language skills, and presentation skills. The content of the CD gave learners direct instruction in the features and immediate feedback to create a program with greater learner control (Boling & Soo, 1999). Content was also provided on each feature through an "advice" section, which included written explanations, samples, and 1-minute video lectures for each of the skills and strategies.

In addition to the features and tasks, an important part of the development process was to consider our purpose for the CD and to plan its integration into the course. Our intent was to use the CD to create a CALL/on-line component of the Oral Communications course. Through the tasks on the CD, learners were given the opportunity to recognize important characteristics of oral communication to produce written and oral responses that are distributed through the World Wide Web. On the Internet, compressed audio, using Purevoice[2] (Qualcomm, 1998) technology, allowed learners to share their feedback orally from a distance. By interacting with these materials, learners developed the ability to incorporate these necessary strategies into their own presentations for the three face-to-face meetings of the on-line oral communications course.

Once we had created a framework for our virtual textbook, we used Hyperstudio[3] (Knowledge Adventures, 1998) to develop a template of our program using the clip of just one lure creating tasks for each lan-

Usability Testing

Usability testing is common in the field of product development. In usability testing, data are collected on the interaction between users and prototypes of a product. Much information exists on usability testing in a variety of fields, from engineering to educational technology to medical technology (Dumas & Redish, 1993; Hom, 1996; Lindgaard, 1994). It is considered a critical step in production, which often leads to revisions, and is usually repeated throughout the development cycle.

There are many methods that can be employed in usability testing. The overall process is to get users to perform tasks on the materials being developed and to observe what occurs as they do so. During observation, the developer collects data on the actions, particularly the problematic actions, of the user. There are many possibilities for proceeding with usability testing. Each developer should consider the materials, the users, and the environment and then adopt the most appropriate method. An article by Sugar (1999) looked at novice designers' typical mistakes in usability testing and then provided guidelines to avoid them. We adapted the following guidelines from Sugar (1999, p. 42), which we found especially critical in producing quality results from our usability testing.

During testing

- Consider any difficulties encountered by users to be indicative of flaws in the design of the material and not the fault of the users.
- Do not ignore users whose actions differ from those of other users; these actions represent individual differences that must be considered by the developer.

After testing

- Consider the problems revealed in the testing and avoid leaping to the single, easiest solution.
- Determine what may underlie a problem.
- Consider at least two possible solutions to each problem.
- Reflect on solutions and consult other developers before making any final decisions.

Steps for Usability Testing

Participants and Recruitment

Developers should consider who will be using the end product—their language level, nationality, specific skill level, age, and familiarity with material type (especially important in CALL testing). The participants in our usability testing were students who were enrolled in, or who had taken, the oral communications class for which the CD-ROM was being developed; thus, they represented the target audience for our multimedia textbook. The study began with four participants, which was enough for productive feedback, given the scope of the project. These students came from a variety of fields and cultural backgrounds.

Determining the Process

Developers should consider exactly what will take place during testing. For example, what will the users do first? Where will the testing take place? How long will each task take? It is essential at this stage to consider how to collect the data from the usability session. Although there are numerous possibilities (Hom, 1996), we describe three methods below, which seem most appropriate to use with ESOL learners.

1. The *think aloud protocol* is the most common method for collecting data in a usability testing session because it is an easy way to obtain a good deal of qualitative data. In this protocol, users are asked to say what they are thinking as they complete the tasks. Users' comments and actions are recorded in writing and on video, if possible, by the test facilitator.
2. In the *co-discovery method*, two users work together on the prescribed tasks. This interaction is also observed and recorded. The users are asked to vocalize their thoughts and interaction, which then become the data analyzed by the developer.
3. The *self-reporting log method* is basically a paper and pencil journal in which users record their actions and thoughts while performing tasks with the material. This may be a good method, particularly for students with limited oral proficiency in English. After testing, the developers collect the logs and analyze the written data.

In our project, we used the think-aloud protocol, in addition to observing students' use, to inform our investigation. We chose this method because students eligible to take the communications skills class have fairly strong oral skills. Also, it allowed us to collect a sizable amount of data in a short time. In addition to the think-aloud, we asked participants some general questions after the testing. This short interview was used to collect additional thoughts which the participants might not have articulated during the testing and also to allow them to make some general comments about the materials.

Designing Tasks for Participants

Before testing, developers should consider what students will be expected to do with the materials. Then they should draft clear tasks to represent these expectations. It is also necessary to bear in mind the length of time each task will take and what background knowledge the participant will need to complete it. For our testing, we prepared tasks that resembled the kinds of assignments the students would have in their on-line oral communications class using the program.

Location for Usability Testing

Developers should be sure to arrange the space where the usability testing will take place. Ideally, the room should be similar to where the material will be used. Also, recording equipment should be placed carefully so as not to distract the participant. Our testing took place in the classroom lab where learners would eventually use the software. Learners were seated at a computer while one observer sat beside them taking notes and the other recorded the session on a digital video camera, located diagonally behind the participants.

USABILITY TESTING PROCEDURE

(Total time about 1 hour and 15 minutes)

- **At the beginning of the test, make the participants comfortable. Explain their role in the testing; be sure they know that it is the material that is being tested, not them. Have them read and sign a release form. (20 minutes)**
- **Explain "think aloud," which means to talk about what one is doing and thinking while completing the task. Then, give participants one task and take notes while they "think aloud." (30–40 minutes)**
- **Ask interview questions. First, get clarification on any confusing parts of the "think-aloud." Then ask the following questions. (15 minutes)**
  1. **What did you find the most difficult in using this program?**
  2. **What did you find most beneficial about this program?**
  3. **Could you suggest any improvements?**
- **At the end, thank them for participating and ask if they have questions. Be sure to give them contact information, in case they have questions later. (15 minutes)**

Conducting the Test

When the participants arrive, the developers should explain what will take place during the session. Also, depending on the project, developers should obtain the participants' written permission (or that of their parent or guardian) to participate through a release form that clearly states the purpose of the session and the rights of the participants. In addition, is should be made clear to the participants that it is the material that is being tested, not them. This assurance can alleviate some testing anxiety that could otherwise impact the session (Hom, 1996). Before concluding the session with the participants, it is important to answer any questions they have and, perhaps, ask some general follow-up questions. In our project, we followed a detailed procedure during testing to assure that each session maintained a consistent structure. (See the sample procedure in the sidebar on p. 17.)

Analyzing Data Collected From the Usability Test and Implementing Changes

Again, numerous possibilities exist for this step, depending on the goal for the testing and the product being tested. Following our testing, we compiled all the comments from the think-aloud protocol and the interviews and grouped them by their foci (e.g., design, navigation, or content). It is common to conduct usability testing with a predetermined focus in mind (Hom, 1996). However, because we did not set our criteria initially, we chose this inductive process to look for patterns in the data. This was to avoid overlooking any aspect of our material that might need revision. By allowing the foci to emerge from the sessions, we were able to find areas with problems in the program that we had not predicted before testing. A *usability matrix* was developed to systematically organize feedback (see a sample excerpt from our matrix below). Following our discussion of the students' comments in each area, we determined appropriate changes and implemented them, resulting in a revised version of our multimedia textbook.

Findings

As instructors developing CALL materials, we found student feedback not only useful, but critical in creating quality materials. Numerous important considerations, which might not have been uncovered otherwise, were revealed. The following is a description of how usability testing affected our perception of the courseware and led to further development.

- **Assessing learners' expectations.** Through the testing, we could see how learners negotiated computer software. For example, one student commented that he

Usability Matrix Sample

Focus	Issue	Student Input	Developer Action
Design	Program Introduction	**S1:** "Need to set apart Introduction more." **S2:** On Bulletin Board, make Introduction larger and in red. **S3:** Could not find the Introduction on main page. **S4:** Could not find the Introduction.	–Add bar across the top for Introduction. –Plan for student orientation. –Develop users' manual.
	Main Menu	**S1:** "Bulletin board is confusing." **S1:** Suggested more separation of flyers on bulletin board. **S2:** "On bulletin board, I can't tell which is lecture." **S3:** Very confused by bulletin board. **S4:** Could not figure out the bulletin board.	–Revise main menu. –Plan for student orientation. –Develop users' manual.
Navigation	Video	**S1:** Suggested having a start button for videos instead of a bar. **S2:** In Advice section, could not tell how to start the video. **S3:** Had difficulty starting the video on the advice page. **S4:** Could not tell how to start the video, suggested a button to the right.	–Change text on start button. –Allow user to click on video picture to start. –Hold an orientation for students on using materials. –Develop a users' manual for students' reference.
Content	Instructions	**S2:** "Too much information in the instructions." **S3:** Had a hard time reading exercise instructions and question. Eyes went first to the video on the right side, not to the question at the top of the screen.	–Pare down the written instructions. –Adjust the color contrast, making the questions darker. –Train students in an orientation. –Develop users' manual.
	Vocabulary	**S3**: Could not understand the words, "instill confidence."	–Create a link to an on-line dictionary. –Add a discussion board to the on-line support for students to ask vocabulary questions. –Have a glossary in the manual for technical and content terminology.

Note: Test takers' actual statements recorded by the test facilitator are indicated by quotation marks in the Student Input column. All other statements are observations made by the test facilitator about the test takers.

preferred a wide-screen format for viewing the lectures because he was used to that size when watching television. The same student was also observed clicking on every piece of text that we had colored blue, expecting them to be hyperlinks to the Web.

- **Determining major problem areas.** The testing highlighted areas that required solutions. The think-aloud protocol allowed students to express their problems with the software and also their suggested solutions. For example, all four of the participants struggled to navigate the main menu of our program, which was a significant problem. In addition, they all experienced difficulty locating the introduction video linked to the main page. Apparently, it was too similar to the links to lectures. This led us to revise the format of the main page.
- **Testing changes based on initial informal input.** By conducting both the informal input sessions and the usability testing, we were able to track changes to see if they were effective. In the initial test with a class, a student had recommended we include instructions, which were added. But in the usability testing, we found that students had difficulty locating the instructions, and that, once located, they were too detailed. This observation argues for not only usability testing, but testing material prototypes several times during the development process.
- **Identifying language appropriateness problems.** Although we assumed this particular audience possessed a high level of language ability, we still found some areas where language confused users. One participant struggled to complete the exercise because she could not understand certain content vocabulary. This led us to develop a hyperlink to an on-line dictionary.

Conclusion

As more instructors attempt to develop computer-based materials, it becomes increasingly important to assure the appropriateness and usability of these materials. The quality of the materials tested in this project improved much more as a result of usability testing than it had during nearly a year of seeking suggestions from colleagues and developer-initiated revisions. The involvement of students in the development process proved to be valuable. The usability testing procedures employed in this project represent one of many approaches that may be used to have students assess the usability of materials.

By employing methods of usability testing similar to those presented in this article, instructor-developed materials can be made more efficacious and student friendly. We encourage all ESOL instructor-developers to consider incorporating usability testing into the materials development process. Methods of evaluation should also be shared with the TESOL community to increase our understanding of how such feedback and evaluation can serve us all.

Notes

[1]For samples of usability testing and access to downloads, please see the companion Web site to this article at http://gregling.net/tj/.

[2]Purevoice technology allows the recording and compression of audio for e-mail distribution. It is available for free for both Windows and Macintosh platforms from Qualcomm (1998).

[3]Hyperstudio is a classroom multimedia authoring program available from Knowledge Adventures (1998).

References

Bishop, A. (1999). CALL issues: Multicultural considerations in a CALL environment. In E. Hanson-Smith & J. Egbert (Eds.), *CALL environments: Research, practice, and critical issues* (pp. 315–325). Alexandria, VA: TESOL.

Boling, E., & Soo, K. S. (1999). CALL issues: Designing CALL software. In E. Hanson-Smith & J. Egbert (Eds.), *CALL environments: Research, practice, and critical issues* (pp. 442–458). Alexandria, VA: TESOL.

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

Dumas, J. S., & Redish, J. C. (1993). *A practical guide to usability testing*. Norwood, NJ: Ablex.

Hom, J. (1996). *The usability methods toolbox*. Retrieved February 2, 2001, from the World Wide Web: http://www.best.com/~jthom/usability/.

Knowledge Adventures. (1998). Hyperstudio® (Version 3.3 for Windows) [Computer software]. (Available from http://www.hyperstudio.com.)

Leshin, C. B., Pollock, J., & Reigeluth, C. M. (1992). *Instructional design strategies and tactics*. Englewood Cliffs, NJ: Educational Technology.

Lindgaard, G., (1994). *Usability testing and system evaluation: A guide for designing useful computer systems*. London: Chapman & Hall.

Ngeow, K. Y. W. (1999). Classroom practice: Enhancing and extending learning styles through computers. In E. Hanson-Smith & J. Egbert (Eds.), *CALL environments: Research, practice, and critical issues* (pp. 302–314). Alexandria, VA: TESOL.

Plass, J. L. (1998). Design and evaluation of the user interface of foreign language multimedia software: A cognitive approach. *Language Learning and Technology, 2*(1), 35–45.

Qualcomm. (1998). Purevoice™ (Version 2.0 for Windows) [Computer software]. (Available from http://www.qualcomm.com).

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

Sugar, W. (1999). Novice designers' myths about usability sessions: Guidelines to implementing user-centered design principles. *Educational Technology, 39*(6), 40–44.

Thorn, W. J. (1995). Points to consider when evaluating interactive multimedia. *The Internet TESL Journal, 2*(4). Retrieved January 25, 2001, from the World Wide Web: http://www.aitech.ac.jp/~iteslj/Articles/ThornEvalueConsider.html.

Authors

Greg Kessler teaches in the Ohio Program of Intensive English at Ohio University, in Athens, in the United States. He coordinates ITA and graduate oral communication courses and manages a self-access computer lab. He is currently serving as coordinator of TESOL's Electronic Village, a member of the TESOL CALL Interest Section Steering Committee, and as OhioTESOL Webmaster.

Lia Plakans is an instructor in the ESL program at The University of Iowa in Iowa City, in the United States. She teaches courses in the intensive English and the international teaching assistant (ITA) preparation program. The project addressed in this article was based on work completed through a Technology Incentives Package grant at Ohio University.

Fulbright Teacher Exchange Program