

FROM LANGUAGE LAB TO MULTIMEDIA LAB: BRINGING ORAL LANGUAGE ASSESSMENT INTO THE NEW MILLENNIUM

In C.M. Cherry (Ed.) *Dimension 2002: Cyberspace and Foreign Languages: Making the Connection*. Valdosta, GA. 33-42.

In most foreign language classes, teachers work to help students develop competency in each of the four skills: listening, speaking, reading and writing. As part of their teaching program, they must also assess their students' progress in these areas. The evaluation of listening, reading and writing is a fairly straightforward process. It can readily be accomplished by pen and paper tests using true/false, multiple choice, and/or essay-type questions (Baker, 2001). The assessment of speaking skills, however, is much more challenging, especially given the fact that it should be done on a regular basis (Gonzales, 1989). In some cases, it is actually neglected because of the amount of time that oral testing requires (Egan, 1999).

It would be pedagogically unsound to attempt to measure speaking proficiency with a pen and paper test (Flewelling, 1996). Therefore, other approaches are required. Traditionally, teachers have relied on classroom observation, oral interviews, and, in some cases, language lab testing to obtain oral proficiency marks for their students. Technology now offers us a new option: computerized oral testing. This paper will consider the strengths and weaknesses associated with more traditional oral evaluation techniques and will then explore how oral assessment via computer may make the task of assessing a student's oral proficiency more efficient and more effective.

Classroom Observation

This is a frequently used technique that lends itself well to formative evaluation since it can be done on a daily basis. Teachers can decide ahead of time to monitor the oral work of a limited number of pre-selected students in each of their classes (Parrott, 1993). Performance can be documented through the use of checklists that outline the criteria to be considered. Teachers may consider the performance of students who participate orally in whole class situations, they may observe students as they work in small groups, or they may ask individual students leave their desks and to perform specified tasks while the teacher observes. In most cases, classroom observation marks would supplement marks from other forms of oral assessment.

Interviews

Interviews between the student and teacher are an effective way of focusing on a student's oral capabilities. There are, however, problems associated with this strategy. Conducting oral interviews with one student at a time is a very time-consuming process. Given that teachers can have 200 or more students in a given school year, scheduling interviews can be a nightmare.

Teachers can cut the amount of time required for oral interviews in half by asking two students to communicate with each other on given topics, but there are problems associated with this technique as well. If one student is stronger than the other, he or she may tend to dominate the conversation. Additionally, it is difficult for a teacher to focus equally on both students at once (Hughes, 1989).

Another problem is that once the student has finished speaking, there is no opportunity for teachers to listen to the response again. Therefore, teachers must remain carefully attuned to what the student says and they must be able to quickly translate what they hear to the evaluation criteria which have already been established for the interview. Some teachers may wish to remedy this problem by tape recording the interviews so that play back can occur.

Giving oral interviews is tiring. Typically, teachers start out fresh, able to remain focused on student answers. As they go on, however, fatigue can set in and their tiredness can result in loss of focus. The result may be evaluations that are not as fair or as accurate as they should be (Haggstrom, 1994). Some teachers may be able to arrange to have interviews monitored by two instructors: one who can lead the discussion and one who can grade student performance. This may better ensure fairness since the evaluator can concentrate completely on student responses but it still does not guarantee that the interviewers' evaluations won't be affected by fatigue. Furthermore, in some schools there may be only one instructor who teaches the target language, or, in cases where several instructors teach the same language, it may not always be possible to arrange for two to be freed up at the same time for interview sessions.

Evaluating Large Numbers of Students at Once

Perhaps the best-case scenario for teachers wanting to evaluate the oral skills of their students is to find a way to evaluate a whole class at one sitting. One way of conducting group assessment is to use a language lab. Apart from the obvious advantage of giving teachers a means of evaluating many students at once, there are other benefits

associated with this method of evaluation. Pre-recorded questions help to ensure that all students receive the question in the same fashion. Teachers can pre-establish the length of time allocated for a response, again ensuring that all students are tested in the same fashion. Another advantage is that teachers can evaluate student responses at a time and place that are convenient to them. They can take breaks as often as needed to ensure that fatigue does not negatively influence their evaluations. Furthermore, because student responses are recorded, teachers can save recordings as documentation of their marks. This documentation can be particularly useful at parent-teacher interviews or if a student's mark is called into question. Recorded answers can be played back as many times as necessary, and teachers can replay certain sections of a response. They are, therefore, better able to focus on individual evaluation criteria as opposed to having to give a more holistic assessment.

Although there are many advantages associated with language lab assessments, there are major drawbacks. Interview situations tend to be very artificial and they impose stresses on students that may affect performance. Pre-recorded questions are inflexible in that, unlike in an interview situation, there is no way for the instructor to prompt, rephrase, or follow up on student responses (Hughes, 1989). Furthermore, technical problems can arise, resulting in frustration for the teacher and the students. The worst-case scenario would be equipment that malfunctioned en masse, thus forcing the teacher to cancel the evaluation session. Less serious but equally annoying would be if one or two recorders acted up, preventing students assigned to those stations from participating in the test. If a student did not speak clearly and directly into the microphone, teachers might be frustrated by inaudible or barely audible responses. There is also, of course, the

problem of cassette tapes breaking during recording or play back. Tapes can be spliced, but often this remedy results the loss of a section of the recording.

The Move To Computerized Oral Language Assessment

The author decided to move to computer-based oral assessment because her university closed its language lab and refocused its interest on the university's multimedia lab. This closure did not create a problem for instructors who had used the language lab for instructional purposes only, but it did create a problem for the author, who had used it for student evaluation.

Rather than return to one-on-one interviews with students, the author approached the university's Manager of Instructional Research, Design, and Development to see if a software program could be written that would allow her to evaluate groups of students in the same way that she had in the language lab but now in the new multimedia lab. Use of an early version of the resultant software proved to be so successful that the author decided to explore how other foreign language teachers would respond to the idea of using the software for oral language assessment. Input from local teachers resulted in the software's being reworked several times until it reached its present form. And it is this version that is now the subject of a pilot study throughout Canada, the United States, and Europe.

Evaluation and Computerized Testing

Whenever student performance is to be evaluated, instructors will need to decide in advance the method they will use to record and score test results. They will also need to determine the criteria on which the evaluation will be based. (Gronlund, 1998). Many instructors find it useful to make use of an evaluation rubric designed to help them to focus on specific aspects of a student's performance. The rubric might take the form of a checklist, a rating scale or anecdotal comments. In some instances, teachers may choose to apply pre-established criteria such as those of the ACTFL-OPI (American Council on the Teaching of Foreign Languages Oral Proficiency Interview) (ACTFL, 2002). The ACTFL criteria are also used in COPI (Computerized Oral Proficiency Interview) (Center for Applied Linguistics, 2002) and SOPI (Simulated Oral Proficiency Interview) (Center for Applied Linguistics, 2002b). Many teachers, however, may find the ACTFL rubric better suited to placement testing rather than for testing performance related to specific curriculum content. Teachers may prefer to create their own rubrics in order to ensure that the criteria reflect the material taught in class and the specific competencies to be evaluated. Teachers may also prefer to tailor their rubrics to the grading procedures required for reporting at their institution.

The Software

A problem associated with much of the educational software on the market is that it has been designed by programmers with little or no pedagogical expertise. (Flewelling,

1989). The present software, however, was developed by an expert in second language pedagogy and experts in software design. It is, therefore, both pedagogically and technically sound.

The software is network-based, and it provides teachers with a reliable and easy-to-use method of creating and conducting oral language assessments. It is well suited for use with one student at a stand-alone computer, a group of students in a computer lab, or for students learning in distance education programs.

The system uses streaming media technology to deliver oral language tests over a network. Streaming media technology uses special data compression methods to allow large audio and video files to be distributed over the Internet. This method of delivery is more efficient and more reliable than standard web page delivery methods. It requires users to have a program such as RealPlayer or Windows Media Player on their computer so that the media files can be decompressed and played.

The software records oral test responses made by students and then uploads the responses to a network server. Teachers can access student files, evaluate the responses and record oral feedback for students, who will later be able to listen to their original responses and their teacher's recorded comments.

The software is a shell that contains no resources. Teachers create their own questions, thus ensuring that the content will, unlike much of the commercially-available foreign language software, reflect the vocabulary, grammar, and themes presented in class.

The software uses a browser interface to provide teacher and student portals. The portal strategy organizes and maintains teacher and student resources within a school

community. The teacher portal provides access to tools and resources required for the creation of class lists used to authenticate students during log-in, for assembly of test questions and tests, for the evaluation of student submissions, and for recording oral feedback. The student portal gathers and displays tests and other resources that have been assigned to an individual student by one or more teachers. The portal method gives students one place to go to get their assignments and to retrieve their assessments.

Question resources are imported to the program from within the teacher portal. Currently, standard test questions are created using one .jpg graphic file and one audio recording file in .wav format. Questions can be stored in pools or folders that represent different themes, levels of difficulty or skills. Tests are generated by selecting questions from one or more question pools. Teachers may choose to have questions delivered either in a preset order or randomly.

Students can access questions in one of two modes: practice mode or test mode. In the practice mode, students may select test questions at random from a test menu bar. After a question is presented, students may record an oral response, listen to their recording, and then compare their response to a sample answer provided by the teacher. The sample answer allows teachers to provide an example for students and parents, thus clarifying performance expectations and encouraging learning (Heide & Henderson, 2001). In practice mode, work is self-paced, and student recordings are not sent to the server. In test mode, questions are presented in a timed fashion. Each question has a preset answer preparation time, and answer response time that have been pre-determined by the teacher during test production. In this mode, student responses are uploaded to the server and stored there for evaluation.

The software will run on any Pentium-class, multimedia- equipped PC computer with 32MB of RAM running Windows 95, 98, NT, or 2000. At the present time, a Java version for Macintosh computers is not available, but it is in development, and, once available, plans are to create a version of the oral testing software for Macintosh computers as well.

A sound card, microphone, headset or speakers, and network connection are required for software use. There is no need for additional software purchases. The current release requires Microsoft Internet Explorer 4.0 or higher. The free browser plug-ins listed below are required. Users who do not already have them installed on their computer will easily be able to download them at no cost.

- Java 1.3 from Sun Microsystems
- Shockwave Player from Macromedia
- RealPlayer from Real Networks

Hypotheses Regarding Computerized Oral Assessment

Use of the software has suggested a number of possible benefits of computerized oral evaluation to the author. Pilot testing of the software will, it is hoped, provide concrete data about the possible strengths of computerized oral testing in general and this oral testing software in particular.

A common complaint from foreign language teachers is that commercial software frequently doesn't reflect the vocabulary, grammatical structures, and themes taught in class. This software avoids that problem, since teachers may author their own questions,

tailoring them to specific material taught in class. Thus teachers may find more uses for this software than for software with fixed content.

Students may tend to feel more relaxed responding to an inanimate computer than in front of the teacher. Less stress may result in better student performance.

The software has video as well as audio capability. Because teachers can include a picture or text with each question, a context for questions is established. This capability is important because research shows that students learn best when they use language in context (Claybourne, 1999). Northrup and Tracy (1998) comment that video and audio are useful tools for all second language learners, and Brown (1987) points out that audio and visual prompts establish a context for the questions, thus helping to ensure that test questions have content validity. This feature may increase teacher confidence in the software's pedagogical soundness.

The software is interactive and may appeal to students with different learning modalities: auditory, visual and kinesthetic (Grasha and Yangarber-Hicks, 2000). Additionally, Bitter, Camuse and Durbin (1993) point out that one of the most promising roles of computers in education is their ability to assist the handicapped. Although the software is primarily intended to assist teachers with oral testing, it may also be useful with students who have difficulty writing since they may benefit from being able to record answers orally.

Teachers can mark the responses at their convenience and from any location with a networked computer. They are not obligated to continue marking when fatigue sets in. Additionally, teachers can play the same response as many times as they wish without downgrading the quality of the response, thus allowing them to focus on different aspects

of the student's speech at different times. The result may be a fairer and more accurate assessment.

Teachers can attach oral feedback to each student answer so that students can play back not only their own answers but also the teacher's reactions. Linking feedback so intimately with student responses makes the assessment more meaningful for students (Hall, 2000). The result may be enhanced student learning.

The software can either be used in a computer lab setting or on a single computer by one student at a time. This flexibility allows teachers who have access to a computer lab in their school to test a group of students at once or, alternatively, they can send an individual student to a classroom computer to do a test while the regular lesson is taught by the teacher. This may result in more frequent evaluations than if teachers were using more traditional evaluation methods.

The volume of students' recorded work can be increased or decreased even after the recording has been completed according to the marker's preferences. The result may be a fairer and more accurate assessment.

Research suggests that parents and educators alike are concerned with knowing on what oral assessments are based (Crawford, 1996). Buckley (1992) contends that the best way to evaluate oral language is to use taped samples. The software provides teachers with recorded language samples that, because they are stored in files, can provide teachers with documentation that will support marks they give for oral language proficiency. The availability of this documentation may result in teachers feeling more confident about their accountability to students, parents and administrators.

Because students can access their recorded answers and teacher feedback, they are in a position to assess their own competency. Crawford (1996) suggests that by reviewing their work, students will realize that the more they practice speaking, the more proficient they will become. This practice may have a positive effect on student learning.

Pilot Study

The pilot sites include elementary, secondary, and post-secondary students studying a variety of foreign languages: French, Spanish, German, Mandarin, and English as a Second Language, to name a few. The pilot study has been made possible through a grant from the Office of Learning Technologies, a branch of the federal government of Canada.

Research questions to be addressed are as follows:

- What are the evaluation practices currently used by foreign language teachers in assessing the oral competency of their students?
- Do teachers find that use of the software makes evaluating student oral work easier than traditional evaluation practices?
- Do teachers find that use of the software makes their evaluations more accurate and fairer?
- Do teachers find that use of the software increases their confidence in their level of accountability with regards to the marks they assign students for their oral work?
- Do teachers find that use of the software helps them better meet the needs of special needs students who may have difficulty in writing responses to questions?

- Do teachers like using the software?
- Do students like using the software?
- Does the use of the software have an effect on student motivation with respect to oral participation?
- Does the use of the software have an influence on students' oral marks?
- Does the use of the software help teachers to better integrate technology into their teaching programs?
- Is the software useful in distance education situations?

The pilot study will also provide the researchers with feedback concerning the reliability and functionality of the software and whether any technical or pedagogical changes in the software would be advised.

Piloting will continue through the spring of 2002, and results of the study will be available once data analysis has been completed. Data from the pilot study will allow the software developers to modify the software, if necessary, to better meet the needs of teachers and students and to make claims about the software that will satisfy concerns of a pedagogical and technical nature. This evidence will be important if teachers and students are to use the software with confidence.

Conclusion

More and more, teachers are being encouraged to integrate computers into their teaching programs (McLafferty, 2000). Some encourage their students to use software designed to enrich the curriculum; others have their students use computers for research. Fewer are using computers for testing purposes. However, as Gronlund (1998) suggests,

“as computers become more widely used in the classroom, we can expect computer-assisted testing...to play an increasingly important role” (p. 131). This belief is echoed by Hall (2000) who states that “computer-based testing is the wave of the future” (p. 15).

By using computers for oral language assessments, foreign language teachers will be able to demonstrate the integration of technology into their teaching programs – something that more and more education administrators are encouraging (McLafferty, 2000). They may also be able to enhance student learning and proficiency, increase instructor accountability, and make the task of evaluating oral skills less burdensome, fairer, and more accurate.

Teachers interested in further exploring this option should consult the web site: www.netresources.ca where they will find a sample test along with detailed information about the software. They may also contact the author at flewell@uwindsor.ca. It is the author’s intention to publish the results of the pilot study in a subsequent article, and any readers interested in these results are welcome to contact the author in this regard.

REFERENCES

- ACTFL. (2002). Proficiency Testing: Testing For Proficiency. *ACTFL Proficiency Guidelines – Speaking (Revised 1999)*. Retrieved from www.actfl.org. Jan. 8, 2002.
- Baker, C. (2001). *Foundations of bilingual education and bilingualism*. Clevedon, UK: Multimedia Matters.
- Bitter, G., Camuse, R., & Durbin, V. (1993). Using a microcomputer in the classroom. Boston: Allyn and Bacon.
- Brown, H. (1987). *Principles of Language Learning and Teaching*. Englewood Cliffs, New Jersey: Prentice Hall Regents.
- Buckley, M. (1992). Oral language: A curriculum yet to come. *English Journal*, 84, 41-45.
- Center for Applied Linguistics. (2002). *Computerized Oral Proficiency Interview*. Retrieved from www.cal.org/projects/copi.html. Jan. 8, 2002.
- Center for Applied Linguistics. (2002b). *Foreign Language Test Development – Speaking*. Retrieved from www.cal.org/tests/fltests.html. Jan. 8, 2002.
- Claybourne, T. (1999). The status of foreign language and technology. *Media and Methods*, 36, 6-7.
- Crawford, R. (1996). Documenting and evaluating oral language development in the classroom. *Reading Horizons*, 36, 285-296.
- Egan, K. (1999). Speaking: a critical skill and a challenge. *CALICO Journal*, 16, 277-293.

- Flewelling, J. (1989). *The evaluation of foreign language software: Teacher perceptions*.
Unpublished Doctoral Dissertation, University of Toronto, Toronto.
- Flewelling, J. (1996). Instructional leadership and the core French program: A guide for principals. *The Canadian Administrator*, 35, 1-8.
- Gonzales, P. (1989). Prochievement testing of speaking. *Foreign Language Annals*, 22, 487-496.
- Grasha, A. & Yangarber-Hicks, N. (2000). Integrating teaching styles and learning styles with instructional technology. *College Teaching*, 48, 2-10.
- Gronlund, N. (1998). *Assessment of student achievement*. Boston: Allyn and Bacon.
- Haggstrom, M. (1994). Using a video camera and task-based activities to make classroom oral testing a more realistic communicative experience. *Foreign Language Annals*, 27, 161-175.
- Hall, M. (2000). A streamlined future for assessment. *Thrust For Educational Leadership*, 29, 15.
- Heide, A. & Henderson, D. (2001). *Active learning in the digital age classroom*. Toronto: Trifolium Books.
- Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press.
- McLafferty, C. (2000). Technology revolution = education evolution. *Kappa Delta Pi Record*, 37, 22-25.
- Northrup, B. & Tracy, C. (1998). Using technology in foreign language and ESL programs. *Media and Methods* 34, 12-13.

Parrott, M. (1993). *Tasks for language teachers*. Cambridge: Cambridge University Press.