Georgia State University
Digital Archive @ GSU

Modern and Classical Languages Faculty Publications

Department of Modern and Classical Languages

2008

Digital Recordings and Assessment: An Alternative for Measuring Oral Proficiency

Peter B. Swanson
Georgia State University

Patricia Early
Georgia State University

Follow this and additional works at: http://digitalarchive.gsu.edu/mcl_facpub

Part of the Curriculum and Instruction Commons, Educational Assessment, Evaluation, and Research Commons, Educational Methods Commons, Modern Languages Commons, and the Other Languages, Societies, and Cultures Commons

Recommended Citation
Swanson, P., & Early, P. (2008). Digital recordings and assessment: An alternative for measuring oral proficiency. In A. Moeller, J. Theiler, & S. Betta (Eds.), CSCTFL Report (pp. 129-143). Eau Claire, WI: Central States Conference on the Teaching of Foreign Languages. Available at: http://digitalarchive.gsu.edu/mcl_facpub/3/

This Article is brought to you for free and open access by the Department of Modern and Classical Languages at Digital Archive @ GSU. It has been accepted for inclusion in Modern and Classical Languages Faculty Publications by an authorized administrator of Digital Archive @ GSU. For more information, please contact digitalarchive@gsu.edu.

8

Digital Recordings and Assessment: An Alternative for Measuring Oral Proficiency

Peter B. Swanson
Patricia Early
Georgia State University

Encouraging and motivating student engagement in the foreign language classroom is challenging on multiple levels, not the least of which is overcoming perceptions of irrelevance in real-world applications and the affective barriers, such as public-performance anxiety within a group of peers and learner struggles to convey authentic representations of the self and learner abilities. In order to address these and other de-motivating elements with regard to oral proficiency development, a closer evaluation of student values in language learning is required. In a recent survey of undergraduate elementary Spanish students, learners overwhelmingly reported that they place a higher value on speaking and listening proficiency as opposed to reading and writing skills (Swanson & Early, in press), and yet current practices of oral proficiency assessment do little to empower student ownership of language skills.

One way to approach the task of engaging students in the tasks related to oral proficiency development is to encourage students the opportunity to create out-of-class recordings in order to demonstrate their proficiency, thus allowing students to self-select the recordings they believe best represent their true level of accomplishment. However, prior to the decision to begin such assessments, considerations must me made in regards to the most appropriate technology, the specific indicators of proficiency to be assessed, the design of the assessment tool, and the creation of meaningful and authentic tasks.

Oral language assessments and current technology

Communicative second-language instruction at every level focuses on the development of language proficiency in four distinct skills: written language, reading proficiency, listening comprehension, and oral language production (American Council on the Teaching of Foreign Languages, 1999). The first three

skills are routinely evaluated within the classroom as well as through formal assessments, whereas the challenge to assess spoken language ability has resulted in more frequent formative assessments in the classroom, but fewer formal assessments. This is due primarily to the challenges presented by oral assessment, namely the difficulty inherent in the development of useful and flexible rubrics for scoring (Foster, Tonkyn, & Wigglesworth, 2000) and instructors' time required for individual learner assessment (Flewelling, 2002).

In addition to these challenges, traditional, formative oral assessments conducted in the classroom rarely leave an assessment artifact. The creation of assessment artifacts contributes greatly to evaluation, in that they can be archived for future reference and can be used for comparison between-subjects to measure overall progress towards proficiency goals. Digital technology and the conversion of analog language lab systems to digital recording capability are advancing the capabilities for whole-class, concurrent archival recordings (Flewelling, 2002). Researchers in language learning and instruction are beginning to investigate the uses of emergent digital technologies for the potential benefit they promise when incorporated into the language curriculum for the purpose of oral proficiency development and assessment (Chan, 2003; Egbert, 1999; Volle, 2005).

Rapid advances in personal digital technology and the availability of both hardware and software resources for individual recording may provide instructors with the capabilities to collect digital oral production artifacts, while at the same time reducing the amount of class-time required for oral assessment. Each year new digital tools are introduced into the interactive web environment for the use of bloggers, podcasters, amateur (and increasingly, professional) artists, and multimedia aficionados, and although primarily created for the non-educational market, these tools are easily adapted for use in the language curriculum. To begin, we will briefly outline the functionality, challenges, and advantages of digital tools in three distinct groups: portable hardware (Sanako™ mp3 recorder, Creative Zen V™, and the Phillips 1210™), software (Windows Sound Recorder™ and Audacity™), and webware (YackPack™ and gCast). Then, we outline a nine-step procedure to create rubrics to assess students' oral language proficiency. Afterwards, we discuss an oral language assessment research project and the instructors' opinions of digital voice recordings.

Hardware and software resources

Portable, personal hardware

With the widespread diffusion of digital music technology, the prices for personal, portable devices have fallen within a comfortable range for educational purchases. Although the large capacity iPods are still among the digital elite, it is possible to find mp3 recorders with built-in microphones for prices ranging between $35 and $120, depending upon the features and the storage size of the unit. (The iPod was not evaluated as part of this research due to the requirement of an accessory microphone in order to facilitate recording. Only devices with integrated microphones were included.) The underlying premise of using a portable device is that instructors could issue a written prompt to the class or

prerecord an audio prompt onto the devices, check out the units to each student, who then record their responses outside of class. The students would then return the device to the instructor, who could either offload the recordings onto a master archive, or simply evaluate the recordings at their leisure. The primary challenges inherent in this approach could be the transfer of the prompts to each unit and the administration requirements of checking the equipment out to students (considering the possibility of loss or damage).

The lowest-priced unit investigated was the Phillips SA1210 ($35), a basic 1GB mp3 player and voice recorder with push-button recording and an integrated microphone. Although the quality of the recording had a distinctly mechanical tone to it, the articulation was clear and comprehensible. The midrange recorder was a Creative Zen V ($55), also with 1GB of storage and an integrated microphone. The process of recording was rather simple, with "microphone" selected from a list of resources on the main menu, and the recording quality was clearer than that of the previous device. An additional advantage to the Creative recorder is the ability for the instructor to transfer, not only an audio prompt to the students via a prerecorded message stored on the player, but also deliver images as prompts, by transferring digital images to the player and having them called up by the student.

At the upper end of the price range, the Sanako mp3 recorder ($120) was evaluated. Although equipped with only 512 mg of storage capacity, this recorder, designed specifically to serve the needs of language learners and teachers, does have the advantage of featuring a dual track recording system, in which the student can record their voice while concurrently listening to a teacher-track. This recorder expands possibilities for question and answer assessments or simulated, asynchronous "interviews". The recording quality was excellent; however, one significant drawback to the Sanako recorder is that, in ease of use comparisons with other products, the Sanako recording process was not intuitive. Therefore, significant training or detailed user guides would need to be provided to the students in order for them to complete their recording assignments.

Software

Although application software exists in many forms and environments, for the purposes of this article, software is defined as an executable computer application that is directly installed on an individual workstation. Through a search of software download sites, it is possible to identify dozens of shareware and freeware digital recording programs, each with its own interface and features, but all capable of recording oral production in one or more recording file formats, the most common formats being .wav and .mp3. For more information regarding these file types, refer to http://en.wikipedia.org/wiki/Audio_file_formats. Whenever recording via software (or webware, to be discussed next) a minor investment in microphones and headphones will be required. These accessories are easily purchased from any electronics or discount store and can be as low as $10 for a reasonably durable and functional model.

For the purposes of this article, we dispense with the discussion regarding the issues surrounding the digital divide and acknowledge that instructors must evaluate their students and consider whether or not most students will have access to computers outside of the school environment. If students are requested to produce recordings via a personal computer at home, it is highly recommended that parents be informed or included in the process prior to the assignment. It is ethically essential to be certain that the recorder installation requested is free of adware, spyware, or license limitations, and that the tool itself will not monopolize computer processing and storage resources.

The free Audacity recorder (Mazzoni & Dannenberg, 2000), available at http://audacity.sourceforge.net/, is an open-source recorder (available to the public with relaxed or non-existent intellectual property restrictions) that meets these requirements. Its familiar buttons and interface contribute to ease of use, and for the more technically proficient user, the software also allows relatively sophisticated editing capabilities. Sound files are recorded in the .wav format, but if .mp3 recording is required due to file storage limitations, an additional LAME encoder can be easily downloaded and installed from an associated website.

Every computer that utilizes the Windows operating system comes already equipped with the Windows Sound Recorder™. This program is accessible via the Start Menu by clicking on Programs > Accessories > Entertainment > Sound Recorder. One main disadvantage inherent in the Windows Sound Recorder™ is the limited recording time available (60 seconds). In addition, the only file format available with the Sound Recorder is the .wav format, but the limited functionality of the recorder can also contribute to its ease of use, as users do not have to download an additional file encoder.

Webware

Webware encompasses online applications of software that do not require downloads and installation of software on individual computers. As such, these tools are available from any web-enabled computer provided it is capable of sufficient connection and processing speed. An immediate advantage presented in these tools is the non-dependence on computer operating system, making them accessible to all platforms: Windows, Apple, and Linux. An administrative, and potentially legal, concern in using webware for student assessments is the fact that these recordings are created, and stored, via third-party servers, raising questions of confidentiality and reliability. However, in each of the systems presented below, it is possible for instructors to limit access to the accounts and the student recordings to only themselves and others they delegate, with the exception that internet service providers and webware developers maintain access and archives of the recordings for security purposes.

A popular free web tool for voice recording is YackPack (Fogg, 2005), obtainable at www.yackpack.net. Using this software, instructors can establish class "packs", or groups of students, and then interact asynchronously with the students. Prompts and responses can be recorded via the online interface and delivered to either an individual or the entire class, and ongoing discussion threads

can be created to share information and create truly communicative exchanges. One disadvantage of YackPack is that teachers would need to create a "pack" for each class, and then invite the students to join the "pack" via email accounts. As a result, it is necessary for each student to have an active email account prior to joining the class "pack".

For optimal results, instructors may care to utilize the language lab environment, where instructors or media specialists assist students in creating their accounts and joining the group. However, once the initial setup has been completed, recording and submitting recordings is intuitive and the interface is easily accessed and utilized. An additional advantage is the ability to personalize the recording environment by uploading student images to the "pack", reducing the impersonality of the digital environment. The greatest advantage offered by the YackPack, however, is its ability to accommodate open class discussions and threaded conversations, bringing the opportunity to archive and assess interactive discussion skills between students, and not inauthentic responses to programmed prompts.

The final tool to be discussed is gCast, developed as a tool to make podcast production and distribution easily accessible to bloggers, and accessible at www.gcast.com. While categorized as a web tool, it holds a distinct advantage over the other tools in that it requires no student computer in order to record student voice. gCast is unique in its ability to record input via telephone, and archive it on an established web account. In order to utilize gCast, the instructor must first create a gCast account. Again, it is highly recommended that separate accounts be created for individual classes to facilitate organization of recordings. Once the account is created, a gCast web page is created for that user (the instructor) and a PIN number, or access code, is identified for that account. Instructors may then distribute a toll-free telephone number indicated by gCast, and the access code, to their students.

Using any telephone, students can call into the gCast account, record their responses, review them, and then submit them using simple commands that are now familiar to anyone who has used an electronic voice mail system. By logging into the gCast account, instructors can review and evaluate their student recordings. Because the microphone technology in telephones is quite sophisticated, the high quality and clarity of recordings is remarkably consistent. One disadvantage of this system is that the filenames as they appear on the account website do not indicate the name of the caller, so it would be necessary for students to state their names orally at the beginning of each recording. Of course, the primary advantage for this system of recording is that it does not make presumptions regarding student access to digital technology; any student with access to a telephone can record their voice.

In conclusion to the technical section of this article, it is important to note that the tools mentioned here are the tools that are available at the time of this writing; new digital recording tools are developed and existing tools refined each year, adding greater capabilities and user interfaces that are easier to navigate. These tools, although created for the general web population, add functionality

and practicality to both oral production and listening comprehension development and assessment in the language curriculum. What remains is to establish reliable measures of language ability that can be both generalized to a student population, but specific enough to be useful as analysis tools for individual student oral production.

Rubrics as assessment tools

For years, rubrics have become one of the standard tools to measure student achievement because "rubrics can help teachers analyze and describe students' responses to complex tasks and determine students' levels of proficiency" (National Council of Teachers of Mathematics, 2000, p. 22). Defined as a set of scoring guidelines for evaluating student performance that classify performance into different categories that vary along a continuum, rubrics provide educators a means to evaluate student performance. Additionally, rubrics can help inform students of what is expected in terms of assessment criteria and can help improve student performance, especially if given to the students prior to assessment.

The advantages of using these scoring guides are manifold. Rubrics allow for more consistent and objective assessment as well as allowing teachers to clarify the specificity of assessment criteria. Additionally, these scoring guides show students how performance-based activities will be evaluated and help promote student awareness about the criteria. Further, rubrics can provide benchmarks for educators to measure and even document student progress over time. Lastly, they can function as a useful vehicle for structured feedback to students and for measuring instructional effectiveness. Clearly, rubrics can serve as an integral educational tool especially when implemented alongside the framework of Backwards Design, where instructors first determine the learning outcomes, agree on acceptable evidence of competency for the outcomes and results (assessment), and then plan instruction based on the performance objectives (Wiggins & McTighe, 2001).

There are two distinct types of rubrics: analytic and holistic. Teachers who select a holistic rubric attempt to describe the overall quality of the task to be evaluated. Mertler (2001) summarizes researcher sentiment that holistic rubrics are utilized when errors in some part of the process can be tolerated provided the overall quality is high (Chase, 1999), that use of holistic rubrics is probably more appropriate when there is no definitive correct answer (Nitko, 2001), and that holistic rubrics offer a somewhat quicker scoring process than analytic rubrics since holistic rubrics focus on the overall quality and proficiency students demonstrate on specific tasks in order to get an "overall" sense of what the student was able to accomplish. Additionally, he posits, "only limited feedback is provided to the student as a result of scoring performance tasks in this manner" (p. 2).

However, when a fairly focused type of response is required, analytic rubrics are usually preferred (Nitko, 2001). Here, the instructor identifies important elements of a certain performance task (grammar, fluency, or vocabulary use for example) and assigns a point value for each criterion. Students are evaluated based on performance on each criterion and a summary score of all the different criteria

is obtained. Such inspection of multiple criteria may require more precision and even more time on behalf of the instructor. In fact, Mertler (2001) recommends, "an individual's work should be examined a separate time for each of the specific performance tasks or scoring criteria" (p. 3). From such detailed scrutiny, the degree of feedback for both students and teachers can be tremendous. Students can receive specific feedback on their performance on each of the individual criteria and teachers can adjust instruction as needed based on the results. By design, the analytic offers more detailed information on students' specific strengths and weaknesses while holistic rubrics do not (Nitko, 2001). (Refer to Appendix A and B for an example of each type of scoring guide.)

Viewed as one assessment tool among many, holistic and analytic rubrics use different perspectives to evaluate student performance. Even though rubrics can be designed in a variety of formats, they contain three common features: a stated objective, a range to rate performance, and specific performance characteristics arranged in levels indicating the degree to which a standard of performance has been met (Shrum & Glisan, 2005). Typically, numbers are assigned in ascending order to indicate better performance. In order to further the notion of what constitutes a rubric, we now present specific suggestions for FL educators to construct quality rubrics for oral language assessment. Our step-by step strategy offers instructors a means to create rubrics that dovetail with instructor-determined learning objectives for oral language proficiency.

Rubric construction guidelines

While scores of rubrics are only a click away on the Internet, their integrity can remain problematic due to an array of issues from lack of precision to determine differences in student ability to a lack of congruence between learning objectives and assessment of those objectives. Additionally, many of these easy-to-find rubrics lack any certainty of peer-reviewed approval and seemingly appear to belong in the category of assessment where one rubric serves all purposes. We advocate a 10-step procedure that FL educators can use that not only improves rubric integrity but also increases the accuracy of measuring student oral ability.

1. Determine and state learning outcome(s).
2. Align outcomes to national and state standards for FL education.
3. Determine assessment objective(s) and decide if an analytic or holistic rubric would best measure student achievement.
4. Work collaboratively with others from different schools to develop assessment criteria.
5. Select succinct titles for the performance levels.
6. Articulate quality definitions for each criterion.
7. Assign a numerical scale that is congruent with overall grading measures.
8. Solicit student and colleague opinion and revise as necessary.
9. Share the rubric with students before assessment is administered.
10. Following assessment, encourage students to archive rubrics as a means to document oral language development and progress.

To begin, the FL teacher should determine the learning outcomes. These outcomes should be written as statements regarding what teachers expect students will be able to do as a result of a learning activity. We recommend stating the outcomes using brief statements or phrases utilizing verbs from Bloom's Taxonomy. For example, students will be able to describe their families using the vocabulary from the chapter. Or perhaps, students will be able to compare Christmas traditions in France and the United States. Note, that the outcomes are written in terms of student performance. Additionally, when composing outcomes, FL teachers must determine where students reside regarding the development of language skills. Clearly novices will not be able to express the same levels of competence as advanced students.

Next, once the learning outcomes are established, we encourage FL educators to take into account state and the national standards for FL learning and align the designated outcomes to the standards. The national standards (American Council on the Teaching of Foreign Languages, ACTFL, 1999) revolve around the five goal areas of communication, cultures, connections, comparisons, and communities. The standards support the notion that FL students should function in three modes (the interpersonal, interpretive, and presentational) that serve as a framework for describing language performance at the Novice, Intermediate, and Advanced levels. Further, these three modes are intended to provide a more integrated and natural way of examining communication rather than the traditional approach of teaching and testing the four skills (listening, speaking, reading, and writing) in isolation (Ohio Department of Education, 2007). Therefore, we strongly urge FL educators to align the standards to the objective(s) for the evaluation.

Once the learning objectives and accompanying standard(s) are determined, the assessment objective(s) need(s) to be articulated in a manner that is consistent with the learning outcomes. For example, if the learning outcome deals with being able to describe one's home, the assessment should deal with the important details surrounding the description such as vocabulary. Once decided, the FL educator needs to make a decision regarding appropriate rubric format, analytic or holistic, to evaluate student performance.

As mentioned earlier, analytic rubrics identify and assess individual components of a performance task whereas holistic rubrics assess student performance on the basis of an overall impression of student performance (Pomplun, Capps, & Sundbye, 1998). As Montgomery (2001) notes, one type of rubric is not inherently better than the other. Instructors should utilize a format that works best based on their purposes for assessment. For example, the decision to use a holistic or analytic rubric can be viewed on how the assessment results will be viewed. If an overall, summative score is desired, a holistic scoring approach may be more appropriate. However, if formative feedback is the required, we recommend an analytic scoring rubric. Additionally, the choice of rubric design can also rest on time requirements to create assessment tools, the nature of the performance task itself, and any specific performance criteria the instructor chooses to isolate.

In the fourth step of our rubric construction model, we suggest working with colleagues at different schools, even at distance if the FL educator is the sole

member of the department, to determine the performance criteria (grammar, vocabulary, etc.) for the assessment. Penuel and Riel (2007) reported that getting help from outside one's immediate circle is valuable for obtaining new information and expertise. Many times information shared among a teacher's close circle of colleagues, especially those who have worked together for many years, may not be sufficiently diverse. Sharing ideas with those outside of one's school may develop new ideas for performance criteria that could be perceived as useful.

However, before writing the performance criteria, we recommend labeling the performance levels with succinct titles. Gradient titles such as "Exemplary, Excellent, Acceptable, Unacceptable" or "Distinguished, Proficient, Apprentice, Neophyte " are common and applicable. Equally, "Superior, Good, Fair, Needs improvement" functions well too. The titles do not require lengthy catchy labels. We propose using titles that promote student confidence, that are a reflection of teachers' expectations and titles that avoid using negative wording for the lower levels of performance. Instead, FL educators should opt for titles that encourage students.

Once the criteria have been named, we recommend using between three to five distinct criteria for analytic rubrics. Too many criteria can become overwhelming for students and concentration on several distinct aspects can garner ample understanding of current student performance. For oral language assessment purposes, criteria such as vocabulary, pronunciation, grammar, fluency, and references to culture should be considered appropriate. As a matter of importance, FL educators should continually reflect on the purpose of the assessment when selecting assessment criteria. For example, if the instructor is working primarily with new vocabulary along side well-known grammar skills, perhaps less focus should be placed on grammar whereas vocabulary knowledge should be the assessment target. Additionally, the criteria should be compared constantly with the names for each of the performance levels as a way to ensure descriptions match the appropriate titles.

Once the criteria have been established, we recommend placing the criteria on the left side of the rubric grid. Positioned in such a fashion it allows the reader to view the performance level descriptions from left to right, which is congruent with textual layouts of books or even this article. Once the criteria are placed on the grid, quality definitions for each criterion need to be developed. It is crucial that distinct descriptions are composed for each level of performance. That is, if four different levels of quality are assigned, each level contains accurate descriptions that clearly discriminate between levels. Popham (1997) reminds educators that excessive rubric length is problematic and we concur. Quality definitions need to be constructed in such a manner that indicates performance differences clearly.

Therefore as an example, we suggest avoiding words such as "several, few, some" when describing number of errors students make during performance tasks. It is difficult to discern what distinguishes these words whereas using numerical indicator (less than five, more than three) are much more specific. However, if educators are not interested in knowing precise numbers for criteria such as vocabulary, we recommend descriptors such as "Broad, Adequate, Limited, Very

Limited" for a rubric containing four levels of description. Regardless of choice of wordage, we strongly urge FL educators to be consistent and clear throughout the descriptions and not to blend qualitative and quantitative measures.

The seventh step addresses the numerical scale that works along a continuum. Many times rubrics progress from 4, 3, 2, 1, where students earn minimal points for substandard performances. As researchers, we advocate that the lowest rating still be assigned a positive point value. By doing so, differences between students who participated in the assessment and displayed low levels of performance can be differentiated from those students who did not even participate in the assessment. Additionally, we recommend a total numerical value for the assessment to be directly aligned with teachers' grading schema.

For example, if the archetypal scale (90% A, 80% = B and so forth) is used to determine students' overall grades, the rubric should reflect the same scale such as one that totals 10 points or multiples of 10. Thus, an eight on a performance task would easily reflect a B performance. If teachers prefer to maintain rubrics that do not mirror overall grading categories, Shrum and Glisan (2005) present specific procedures to convert rubric scores to grades. In either case, we recommend placing the rubric's grading scale (A=18-20 points and so forth) near the bottom of the page for student reference since our informal conversations with secondary and collegiate students revealed that students become perplexed by having to use multiple mathematical steps to arrive at a letter grade for their presentation.

Once the rubric is constructed, it is advisable to solicit student and colleague opinion during a pilot test. Since students will be evaluated using the rubric, it is important for students to have the opportunity to review the document and provide feedback. Additionally, prior knowledge of the assessment characteristics allows students to focus their attention specifically on the rubric criteria in hopes of improving performance. Further, once teachers have established a FL colleague community with whom to share issues related to language instruction, expert feedback can help polish quality language assessments.

The final two steps of the rubric creation process include giving it to students prior to assessment and encouraging students to archive graded rubrics as a means to document student second language progress. We recommend giving the rubric to students at the beginning of units so students can begin to prepare for specific upcoming oral assessments much like a teacher-created study guide for written examinations. Once presented to students, we urge teachers to remind students about the upcoming assessment(s) and to emphasize its linguistic components during instruction and activities.

After evaluation has taken place, we strongly advise FL educators to implement student portfolios because they can provide students with opportunities to display quality work, serve as a vehicle for critical self-analysis, and demonstrate progress toward mastery of a foreign language. Besides archiving the rubrics and showcase pieces of linguistic achievement, we suggest adding voice and video recordings to DVD. Additionally, teachers can archive students' recordings to document second language proficiency and progress. Such documentation can be displayed during parent-teacher conferences as well as during accreditation visits. Further, archived

recordings can serve as a metacognitive strategy where students reflect and act on the knowledge of mental processes to improve learning.

Now that the rubric has been constructed and tested, minor modifications can be made for diverse student bodies in today's FL classroom. For gifted and talented students, rubrics can be modified to include additional criteria and possibly more stringent descriptions. Conversely, modifications can be made for special needs students. Additionally, FL teachers can implement strategies to weight some criteria more than others by doubling, for example, the impact vocabulary usage has on an oral assessment. Finally, teachers should leave room on the page for students' and teachers' names, unit plan or chapter designation, the date of assessment, and space for teachers' comments.

As a matter of formatting, ideally the scoring guide should be comprehensive enough to fit on a single piece of paper. At times, text size may need to be reduced to accommodate margin requirements depending on the size and shape (portrait vs. landscape) of the scoring guide. As a suggestion for assessing oral language proficiency during paired activities, we recommend placing two rubrics on the same side (one next to the other) of a sheet of paper. Using this side-by-side strategy, FL educators do not have to move from sheet to sheet when evaluating student performance, eliminating possible performance rating errors. Additionally, not only do we feel that a one-page rubric can be a comprehensive tool to determine students' oral language proficiency, when given prior to assessment, it can help reduce student anxiety since students will be aware of performance expectations.

Activity and accompanying rubric examples

Clearly, using digital technology to assess student oral language ability has serious implications for increasing valuable classroom instructional time. In this section we showcase several examples of oral language activities that FL teachers can be assigned to students as an out-of-class assignments and accompanying rubrics that can be used to assess the performance task. Many tasks that are currently performed in FL classrooms can be adapted for use as out-of-class recorded oral assessments. In the following examples, students are instructed to record responses to learning objectives. While these objectives could be evaluated using different criteria or rubric format, these examples serve as ideas to stimulate FL teacher thought and ultimately FL teacher assessment practices.

The first task is for first-year students at the beginning of the semester where L2 ability is emerging. A two-point rubric was designed to measure students' pronunciation ability and their fluidity of speech.

Task 1	Read a short paragraph (30-40 words) in the target language.
The teacher gives students a sheet of paper with the paragraph (30-40 words) in the target language. They are told that 10 words from paragraph will be selected for assessment of pronunciation and fluidity that the students demonstrate while reading. Students have two days to record the paragraph and email the recording to the teacher.	

	10 points	9 points	8 points	7 points
Pronunciation	Pronounced all selected words correctly.	Pronounced 9 selected words correctly.	Pronounced 8 selected words correctly.	Pronounced 7 or less selected words correctly.
Fluidity	Speech sustained throughout without pauses or stumbling.	Speech sustained most of the time and contains some stumbling.	Speech is choppy with frequent short pauses.	Speech contains long pauses.

The next speaking task is a common first-year assignment where students are learning to describe people. Here, the instructor is evaluating students' ability to use appropriate vocabulary, correct grammatical usage of noun-adjective agreement, and the completion of the task. While the educator could choose to evaluate students on other criteria, such as those from Task 1 (above), the assessment focus is different.

Task 2	Describe a friend in the target language.
The teacher asks students to select a friend to describe using at least 10 descriptive adjectives. The teacher carefully shows students the vocabulary from the textbook that he/she expects students to use in their recordings. Further, the instructor models an example by playing a recording he/she made where he/she describes a friend using at least 10 descriptive adjectives from the chapter. Students have two days to record the description that lasts a minimum of 30 seconds and upload their recording on the teacher's web page.	

	10 points	9 points	8 points	7 points
Vocabulary usage	Wide range of vocabulary that is appropriate for task.	Adequate range of vocabulary and generally appropriate for task.	Limited range of vocabulary that is sometimes inappropriate for task.	Inadequate range that is appropriate vocabulary for task.
Noun-adjective agreement	Student correctly used at least 10 descriptive adjectives when describing the friend.	Student correctly used at least 9 descriptive adjectives when describing the friend.	Student correctly used at least 8 descriptive adjectives when describing the friend.	Student correctly used less than 8 descriptive adjectives when describing the friend.

Task completion	Student speaks for at least 30 seconds.	Student speaks for 25-29 seconds	Student speaks for 20-24 seconds.	Student speaks less than 20 seconds.

The final example may be an assessment for an upper-level course. Note that the teacher has opted to measure student performance using a holistic rubric design to get an overall impression of the students' ability to articulate a progression of events instead of concentrating on discrete elements.

Task 3	Students are to compose a story that narrates events shown in a series of pictures.
The teacher gives the students a sheet of paper that contains seven pictures in cartoon layout that shows a person ordering food at a restaurant and the customer experiences a few problems. Students are requested to narrate the conversation between the customer and the waiter. Students have one day to record their response and email the recording to the teacher by 10pm.	

Exceeds Expectations	Superior completion of task, progression of events is readily understandable with very few pauses or hesitations, rich use of vocabulary without grammatical errors.	10
Meets Expectations	Completes task with limited difficulty, progression of events are understandable with minimal interpretation needed, hesitates during presentation, sufficient use of vocabulary, and demonstrates limited grammatical errors.	9
Almost Meets Expectations	Partially completes task, progression of events are understandable requiring moderate interpretation by the listener, uses frequent short pauses and speech is erratic, inadequate use of vocabulary, and demonstrates emerging use of grammatical structures.	8
Does Not Meet Expectations	Does not complete task, progression of events are not understandable and much interpretation is required by the listener, speech stops accompanied with long pauses, insufficient / inaccurate use of vocabulary and grammatical structures.	7

What is important to note from these three examples is how not only the activity but also the rubric is tailored to address the language task. Again, the tasks are written as learning objectives using suggested verbs from Bloom's Taxonomy.

Conclusions

In this article we discussed the usefulness of out-of-class student digital recordings as a means to both empower and motivate students in their oral language proficiency achievement and to facilitate the creation of oral language assessment artifacts. To that end, several inexpensive or free technology tools were highlighted that educators can easily use in the classroom. As with any educational resource, the tool itself is only as beneficial as the pedagogical foundation and instructional objective upon which it is employed. Each tool presents its own unique advantages and disadvantages dependent upon the educational culture and environment. Within the spectrum of technology resources, educators can locate an appropriate application that will fit their needs and budget as well. Once the technology tool has been identified, the educator can then turn to creation and implementation of the evaluative tool, the rubric, and the design of meaningful and authentic oral proficiency tasks.

Clearly, digital voice recordings can have an integral place in FL classrooms, but also hold the potential to reap valuable benefits outside of class instructional time. We believe that once the educators begin to work with the technology, they will not only share the opinions of the instructors from the study, they will also find new and innovative uses to broaden the information shared here.

References

American Council on the Teaching of Foreign Languages. (1999). Standards for Foreign Language Learning in the 21st Century. Lawrence, KS: Allen Press, Inc.

Chan, M. (2003). Technology and the teaching of oral skills. CATESOL Journal, 15(1), 51-57.

Chase, C. I. (1999). Contemporary assessment for educators. New York: Longman.

Egbert, J. (1999). Classroom practice: Practical assessments in the CALL classroom. In J. Egbert and E. Hanson-Smith (Eds.). CALL environments: Research, practice, and critical issues (pp. 257-271) Alexandria, VA: TESOL.

Flewelling, J. (2002). From language lab to multimedia lab: Oral language assessment in the new millennium. Dimension: Proceedings of the Southern Conference on Language Teaching, 2002, 33-42.

Fogg, B.J. (2005). YackPack [software]. Palo Alto, CA: Stanford University.

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. Applied Linguistics, 21(3), 354-375.

Mazzoni, D., & Dannenberg, R. (2000). Audacity [software]. Pittsburg, PA: Carnegie Mellon University.

Mertler, C. A. (2001). Designing scoring rubrics for your classroom. Practical Assessment, Research & Evaluation, 7(25), 1-10. Retrieved November 24, 2007 from: http://pareonline.net/getvn.asp?v=7&n=25

Montgomery, K. (2001). Authentic assessment: A guide for elementary teachers. New York: Longman.

National Council of Teachers of Mathematics. (2000). Principles and standards for school mathematics. Reston, VA: National Council of Teachers of Mathematics.

Nitko, A. J. (2001). Educational assessment of students (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Ohio Department of Education. (2007). Ohio Foreign Language Model Assessment Project. Retrieved November 24, 2007 from: http://www.ode.state.oh.us/GD/Templates.

Penuel, W.R., & Riel, M. (2007). The "new" science of networks and the challenge of school change. Phi Delta Kappan, 88(8), 611-615.

Pomplun, M., Capps, L., & Sundbye, N. (1998). Criteria teachers use to score performance items. Educational Assessment, 5, 95–110.

Popham, J. (1997). What's wrong — and what's right — with rubrics. Educational Leadership, 55(2), 72-75.

Shrum, J.L., & Glisan, E.W. (2005). Teacher's Handbook: Contextualized Language Instruction. Boston, MA: Thompson Heinle.

Swanson, P., & Early, P. (in press). Technology for Oral Assessment: Recapturing Valuable Classroom Time. In C.

M. Cherry and C. Wilkerson (Eds.), Dimension: Proceedings of the Southern Conference on Language Teaching. Valdosta, GA: SCOLT Publications.

Volle, L. (2005). Analyzing oral skills in voice e-mail and online interviews. Language Learning and Technology, 9(3), 146-163. Retrieved May 17, 2007, from http://llt.msu.edu/vol9num3/volle/

Wiggins, G., & McTighe, J. (2001). Understanding by Design. Upper Saddle River, NJ: Prentice Hall.