

# **DIGITAL VOICE RECORDINGS IN ONLINE LEARNING ENVIRONMENTS**

Gordon Gamlin

Loyola Marymount University, USA [ggamlin@lmu.edu](mailto:ggamlin@lmu.edu)

## **Abstract**

This paper explores the increasing relevance of digital sound files in foreign language acquisition. Several exercise types are listed along with practical teaching tips. Throughout we consider the role of voice recordings in online learning environments along with the challenges and the advantages they may present.

## **1 Introduction**

Digital media has aided foreign language learners in improving reading, writing and listening skills for over two decades, yet attention to oral proficiency has lagged behind the other skills in many online learning environments including those that use system management tools. Solutions to this problem need not be technically elaborate or go beyond readily available software applications. Thus, this paper focuses on the important role of learners' voice recordings in combination with certain electronic course tools designed to enhance language learning and oral competency. A number of exercise types, activities, and learning tasks are included to provide concrete examples for this discussion. Digital voice recorders surpass analog cassette recorders chiefly in convenience, mobility, and sound file sharing capability. The sound files also may become content elements for course tools in learning platforms such as *Blackboard* or *WebCT*. For example, students may upload these sound files to integrate them with Blackboard discussions. Once digital sound becomes part of the online learning environment it may also be considered as an oral text within a visual interface. Pronunciation and creative exercises may focus on sound and speech but they also rely on related listening, as well as reading and writing tasks. Still, the need to make the learner's voice heard within online environments ought to be obvious. In light of the overall sophistication of interactive computer games and software applications for corporate enterprise, a continued perceived inability of academic technology to adequately assist learners with speaking a target language reflects poorly on computer assisted language learning (CALL) as a whole. It is therefore time to find new ways to improve and research the technology to further the integration of digital recordings into the language curriculum.

## **2 Why Voice Files?**

The attraction of originally created voice files can be explained in several ways: On a rudimentary level, voice recordings are an opportunity to personalize content and thus make it more meaningful to learners. The act of recording also lends permanence to learners' creative, though not always, original performance. The objectification of the recorded word allows manipulation and heightens self-awareness, while the technology has an intrinsic playful appeal to the learner. For the language instructor, recorded voice files restore one aspect of the tutor relationship to the student. Instructors can once again listen to each individual student output regardless of class size or time, while students can be sure that their effort will not go unnoticed. Beyond these simple and logistic reasons for the appeal of digital voice files in language teaching, there are also socio-political and economic forces at work. Technological advancements in multimedia for language learners have always been used to recruit students or justify expenditures under the banner of progress. More interestingly, these innovations shape the theoretical discussion in language acquisition. The technical capabilities to include recordings of the target language in multimedia have been with us for some time now. Parallel to the increased availability of media for language instruction, we have witnessed a growing demand for authentic speech in language learning, or as Belz and Kinginger put it: "The emphasis on authenticity within communicative approaches to language education has evolved in response to a rapid expansion in access to resources and people afforded by telecommunications technology" (Belz & Kinginger, 2003). Original multimedia production in language acquisition and the need for authenticity are inseparable at this point in the field's development. The transition from analog to digital voice media at this stage paradoxically brings with it increased individualization for the learner along with an expansion of the learner's community.

## **3. The Separation of Skills**

Technological innovations often present excellent opportunities to challenge existing widely held beliefs. Online learning environments will thus likely reshape the field of teaching foreign languages continuously with every new application. Within the online learning platform, for example, the distinction between four paired, yet separable skills (listening, speaking, reading, and writing) is more of a convenient academic construct than an adequate reflection of real language learning. Online learning environments that merge written and spoken modes in "chat rooms" emphasize the interrelation between skills as they dissolve former distinctions. Similarly, digital audio files appear as documents on discussion boards with all the permanence and authority of the written word. Formally, they are true oral texts. In another example, listening comprehension was first held to be a "passive" skill though it engages the learner actively in meaning making. A "silent period" of listening is still thought to proceed speaking in language acquisition though one might wonder if researchers have been listening carefully enough to their subjects and their first utterances. Campbell, for example, clearly identifies "problems with phonology" as the source of learners' aural comprehension difficulties (Campbell, 2005). The inability to produce or speak a word thus hampers the learner's ability to comprehend its meaning and vice versa. The two are inseparable and should be treated as such within a coherent multimedia online learning environment that combines visual and aural components in a meaningful yet challenging way for learners.

#### **4. Activity Types and Examples**

Currently most language lesson activities with digital recordings are still modeled on analog technology, and several popular exercise types using audiotape recorders date back to the emergence of academic technology in language acquisition. This may be a good opportunity to revisit and update many of these lesson activities and tasks that may be divided into student talk, collaborative tasks, sound effects, mimicry and response, and teacher talk. The first group here entitled “student talk” may include student centered individual tasks such as descriptions, definitions, and sound portfolios.

##### *Descriptions*

This first activity takes advantage of the portability and convenience of voice recorders as students describe an object, a dynamic situation, or an environment such as a market, for example. The format could be similar to that of a live radio broadcast. “Here we are live at...” In preparation for this activity students need to draw up lexical fields with relevant vocabulary, and feedback criteria need to be established. While the activity can be easily shared online, feedback may be sent individually or collectively to students’ private folders. Thus, “Live at the Farmers’ Market” may include names of fruits and vegetables, other food groups, as well as methods of packaging and even preparing them.

##### *Definitions*

Students define their favorite word, concept, or phrase (from the course) and post it in a designated course area. Some learning environments, for example, provide a custom glossary, which could be extended as the course progresses. Successive groups of students may also be part of a continuous database that is “passed on” from one section to another. As they compose their definitions, they consult a variety of sources beyond dictionaries to shape their definitions. In order to ensure that course participants listen to each other’s postings, a “top ten” list could be drawn up or students could be asked to use the selected phrases in one of the activities described below.

##### *Sound Portfolio*

Individual students collect a weekly digital diary of short narratives, impressions, or answer discussion questions on current topics. They post these narratives in the homepage area of Blackboard or similar programs. (Access to these sound files can be limited to the instructor or specific groups of students.) Since this activity may cover the entire length of a course, this is an excellent way to assess individual learner progress not only in oral proficiency but also in improved syntax, increased vocabulary and overall achievement. In large multi-level classes instructors can easily examine individual accomplishments relative to the initial language level at the outset of a course.

### *Collaborative Tasks*

The individual tasks mentioned above are creative yet may be structured to fit the individual curriculum. The second group of activities with digital voice recorders, here is entitled “collaborative tasks” and involves two or more students or, in immersion environments, learners and native speakers. This group includes chain stories, interviews, and dramatic readings.

### *Chain Stories*

The basic model of a chain story typically serves to review the past tense and relies heavily on formulaic structures from the oral tradition: “Once upon a time, there was a [protagonist] in a land far way...”

Students continue the story during lab hours one student at a time until all have had their turn to advance the narrative. Throughout, learners listen to the chain of events their peers have created and contribute their own short segment. The collaboration is posted in its entirety as an online discussion and students suggest endings (in writing). Depending on the popularity of this activity, the story could be continued soap opera style as a radio play. Typically, such stories are especially well received if they include references to the course participants’ environment. Hence, the instructor could provide a number of terms such as the ones mentioned above under “definitions” above to guide the activity.

### *Interviews*

After some initial preparation on a given topic, students interview each other, guest speakers, visitors, or native speakers at school and outside the classroom. Preparatory work includes formulating questions, researching a certain lexical group, topic or content field. Interviews are conducted individually, in pairs, or in small groups. Follow-ups could include transcriptions, summaries, quoted highlights, short articles, and further discussions. Where willing chat partners are available at corresponding institutions, the entire activity could be conducted online. The key to a successful interview is careful preliminary preparation on all parts. Guest speakers, for example, can be an excellent source once students are able to address the topic at hand.

### *Dramatic Readings*

Individual students practice short dramatic reading passages, tongue twisters, jokes, monologues, excerpts from (online) newspapers, works of fiction, and poems. Throughout, (peer) correction ought to occur before the recording. Dramatic readings ought to be closely linked to the course topic at hand. Once the finished recordings are posted in assigned course areas more feedback could be posted alongside the original sound file in the form of more general commentary on the content.

The next category of activities also belongs under the heading of “collaborative tasks” but it includes sounds rather than speech and is intended to spark conversations and provide a context

at the same time. Motion picture quotes and sound effects, music collage, and sounds from different live environments, all share an intriguing quality that prompts an initial curiosity if sufficiently personalized.

### *Sounds from the Environment*

Students record a number of sound effects from a commonly shared environment. Other groups then guess the context and source of the sound. Often, one need not go very far beyond the classroom to come across the hum of an escalator, or the more dramatic sounds of a coffee shop, or vending machine, for example. In a variation on this, groups of students produce a series of sound effects that should form a narrative. Other groups guess and interpret the sound sequence and the story that supposedly goes with it. Actual everyday occurrences such as waiting for and boarding a bus, for example, work better than fictionalized accounts. Groups take turns producing sounds and narrating. The results may be posted along with a written narrative that serves as an answer key.

### *Music and Motion Pictures*

Digital voice recorders easily record original adaptations that can be superimposed on specific parts in songs and films. However, international copyright law discourages the integration of commercial music and video clips into online language presentations. Traditionally, students have enjoyed superimposing their own lyrics on popular songs and inserting their own dialogues in well-known motion pictures. Digital technology affords them the opportunity to sample and manipulate original sources with increased sophistication. Before attempting a project of this kind, instructors must familiarize themselves with current copyright issues that govern their source material.

The final student centered category is entitled “mimicry and response” and includes the recording of formal presentations such as lectures in the target language. This is, perhaps, the most common usage of digital voice recorders in academe though not necessarily in language teaching.

### *Recording a Lecture*

Students may place a digital recorder next to the lecturer’s podium to review the lecture (in part) at a later date. Here, follow-up exercises with specific learning goals are all-important for (language) learners. For example, students may be asked to give a graphic representation of the lecture content in the form of a diagram or a table. Ideally the instructor is familiar with the lecture and its content to guide the review. Overall, the digital recorder proves to be an excellent tool for content based language instruction in immersion settings where digitally recorded lectures are not provided by the institution.

### *Other Lecture Formats*

Aside from academic lectures, guest speakers and performers may also agree to have parts of their presentations recorded. The same considerations apply to these talks as to the formal

academic lecture and careful preparation and debriefing is an integral part of any such lesson. Instructors should be equally aware that performers are typically very hesitant to have their original work recorded in any way.

### *News Broadcasts*

Online news broadcast are available with and without transcripts in both audio and video format. While copyright laws govern these broadcasts just as they apply to commercial music and video mentioned above, there is an activity for language learners that can make convenient use of the digital recorder. After listening to the original broadcast, learners can use the transcripts provided to record or “mimic” the newscast on the digital recorder. Learners then are able to post their own news online. Course participants could thus receive highlights of the (weekly) news that is of most interest to them.

Finally, digital voice recorders are equally applicable to “teacher talk” as they are useful for student-centered activities.

### *Correction and Repetition*

In response to sound files produced by learners, instructors can post feedback, corrections, and customized pronunciation models that are then available to the learner on demand.

### *Teachers’ Talk*

Instructors can post spoken instructions, examples, and questions online. These can be paired with written instructions to increase exposure and help learners.

### *Assessment*

Sound files may be used in online quizzes that can be computer graded instantly. Course tool packages increasingly offer their own software to integrate speech more fully into their learning platform. At the same time, digital voice recorders provide a convenient alternative.

### *Writing Prompts*

Instructors may arrange for native speakers to record question sets. Students later answer these questions (in writing). In the case of the Sony IC voice recorder, voice recognition software (*Dragon Naturally Speaking*) converts digital voice recordings into a text document. This has a number of advantages since “automatic speech recognition makes it possible to assess active vocabulary and word-based pronunciation in a fairly reliable manner” (Jurafsky & Martin, 2000). Students can thus dictate a text into the recorder and walk away with a printout of their spoken text immediately thereafter. In this scenario, the printed text serves as a form of pronunciation assessment. Thus generated learners’ text could then be further used for peer editing and related tasks.

## **5. Integrating Sound File Activities**

From the examples and lesson suggestions above, it becomes apparent that although the focus here is on activities with voice recordings, the described tasks need to be embedded in lessons that equally incorporate all other skills. Thus students write out notes in preparation for the activities above, they listen to their peers' projects and read and write comments related to them. At the same time, digital sound files are always presented within the visual interface of a multimedia online course environment. Instructors and designers of the course tool packages, therefore, face the challenge of integrating sound files and indeed all content within a meaningful structured learning environment that will complement learning experience instead of frustrating it by drawing unnecessary attention to the technology itself.

### *How to Integrate Voice Files with Graphic Representations*

With respect to the exercise types listed above, it becomes necessary to support the content and learning goals and outcomes both visually, as well as aurally. In the "chain story" mentioned above, for example, the narrative sequence initially should be posted in a clear linear fashion that follows the story's beginning, climax, and end. This layout can be further supported by pictorial descriptions of the narrative or a simple timeline that highlights and labels the conflicts within the plot. This does not preclude variations or multiple endings that may branch out (visually) from any point in the story. Such visual graphic representations of knowledge are advocated and well documented by Richard E. Mayer (2001) and others and can especially be applied to vocabulary building, which, in turn, increases oral proficiency in foreign language acquisition. Thus, different web designs also differ in their effectiveness in mediating and preparing for specific oral tasks.

It may be redundant to observe that embedded media delivery through course tool packages differs significantly from print-based and live exercises and tasks in its spatial and temporal dimensions. Online courses thus appeal to different ways of learning, and digital voice files within these courses offer the learner a certain command of the moment, that is tempered with deliberation and increased self-awareness. As graphic representations of course content aid online media delivery in general, they also help foreign language learners with oral tasks and overall fluency. Only a clearly structured interface and content area that is in turn linked to assessment makes for a lucid course presentation. For an extensive discussion of the role of visual and graphic representation in learning see the findings of Bernhard Mohan (1986) and Mayer (2001). Both Mohan and Mayer argue that careful web design furthers student learning. Ideally, technology, learning, and cognition can combine to include different learning modes (Beckett & Mohan, 2003). Thus oral and visual language learners may substantially benefit from currently available online course tools.

## **6. Pronunciation Models**

The International Phonetic Alphabet was perhaps the first concerted effort to mediate and explain the sounds of the target language through the visual medium. For the most part, however, this attempt is cumbersome because it requires the novice to learn the new (phonetic) alphabet along

with the sounds of the target alphabet. The process proves to be somewhat convoluted and in some ways mirrors the now abandoned European attempt to create the simplified language of Esperanto. Simplification turned into distraction from the original goal of communication. Thanks to the successful creation of advanced online learning environments, we have since seen new effective solutions to the problem of pronunciation. In their seminal paper entitled “Sound spellings: Online pronunciation enhancement in a computer assisted language learning environment,” Culhane and Ito, for example, pioneer a sound spelling method based on pronunciation methods employed by radio and television journalists (Culhane & Ito, 2003). Culhane and Ito further demonstrate how the instructor and students’ use of electronic tools enhances oral proficiency using a combination of online tools in Etextweb combined with the class text *Your World*. Throughout, the overall effect of the learning environment is one where online graphic representations of knowledge enhance oral proficiency in foreign language learning.

## **7. Anatomic Models**

Computer generated graphic representations need not be limited to content and phonology alone. A significant amount of research and development resources is allocated to phonological real-time visual representations of the human speech tract. The *Sony SANS Virtuoso* language lab, for example, features an image of a human mouth visual mirroring the recorded utterance. Pronunciation exercises found in language learning software such as *Auralog*, or *Rosetta Stone* feature spoken sample phrases along with animated anatomical renderings of the human speech tract. More ambitious projects rely on three-dimensional models that require the use of specially colored “3-D” spectacles to give a close-up view of the mouth in an attempt to show learners how to pronounce certain sounds (Horii, Inoguchi, Kawazoe, Ogawara, Saita, & Takahashi, 2005). An emphasis on anatomic graphic representation of speech opens new possibilities for language teaching and, in its own way, surpasses the human model any language educator could provide in a live classroom. Klauser points out that “technical effects would allow one to visualize particular points or to exaggerate the pronunciation for learning purposes” (Klauser, 2003). Learners, of course, might not be aware of the model’s exaggerated aspect and may mistake it for the norm. This in turn, raises the problem that the entire notion of a norm in pronunciation is questionable and culturally bound to begin with.

## **8. Conclusion**

Improved voice recording capacities within online learning platforms and the spread of MP3 players with recording capabilities both point to a growing role of digital voice recordings within the field of second language acquisition. The costs related to the implementation of this technology are rapidly declining. Sophisticated portable digital voice recorders retail anywhere from US\$80 while MP3 players, notebooks, and handhelds like the *PalmOne LifeDrive* mobile manager include further recording and file sharing capabilities. The question of how best to use digital sound files for language teaching and learning is one that lies at the core of CALL. If today’s academic technology cannot adequately assist in teaching and learning how to speak a target language then the field is heading for a crisis. A comparison of language activities that use analog recording with those employing digital voice recorders reveals how technological advance

prompts an overall shift towards a more integrated synergetic teaching and learning approach. Individual skills are no longer clearly delineated but rely on each other instead. Future related studies would likely examine the implications of this paradigmatic shift and examine the notion of normative phonology in autonomous online learner communities, to name only one related area of investigation.

## **References**

- Beckett, G.H., and Mohan, B. (2003). A functional approach to research on content-based language learning: Recasts in causal explanations. *The modern language journal*, 87(3), 421-432.
- Belz, J.A., and Kinginger, C. (2003). Discourse options and the development of pragmatic competence by classroom learners of German: The case of address forms. *Language learning*, 53(4), 591-648.
- Campbell, A. R. (2005). The development of a simple stress-time training software. *PacCALL*, 1(1), in this issue.
- Clark, R. C., and Mayer, R. E. (2003). *E-learning and the science of instruction*. San Francisco: Jossey-Bass.
- Culhane, S.F., and Ito, H. (2003) "Sound Spellings: Online Pronunciation Enhancement in a Computer Assisted Language Learning Environment," *Research Bulletin*, Kagoshima University, Vol. 55, 65-76
- Horii, H., Inoguchi, Y., Kawazoe, Y., Ogawara, Y., Saita, I., and Takahashi, A. (2005) Fostering better oral Japanese via e-learning, National University of Singapore: *Proceedings of CLaSIC 2004*.
- Jurafsky, D., and Marin, H.J. (2000). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Upper Saddle River, NJ: Prentice Hall.
- Klauser, F. Hard facts about software. (2003). *Management and training*, (2), 43-47.
- Mayer, R.E. (2001). *Multimedia learning*. Cambridge: Cambridge UP.
- Mohan, B. (1985). *Language and content*. New York, NY: Addison-Wesley.

**Keywords:**

CALL (Computer Assisted Language Learning)  
Digital Media & Foreign Language Learning  
Speaking Proficiency  
Oral Proficiency  
MOODLE (Modular Object-Oriented Learning Environments)  
WebCT  
Blackboard  
Virtuoso & Soloist Instructional Control Software for Multimedia Language Learning Systems  
Graphic Representations of Knowledge  
Online Communities

**Websites:**

Etextweb  
<http://www.etextweb.com/>

**Examples:**

Your World-online Classroom  
<http://www.etextweb.com/Onclass/moodle/login/index.php>

Pacific CALL Association – <http://www.pacall.org/>