PARTICIPATORY DESIGN: CLASSROOM APPLICATIONS AND EXPERIENCES

Pamela Connors, Alice Davis, George Fortier, Kerry Gilley, Brian Rundle, Chris Soland, and Alice Tarachow
Tucker-Maxon Oral School, Portland, Oregon, USA

ABSTRACT

Using the participatory design process, a team of programmers, researchers and a school community have come together to plan and design a computerized speech toolkit to meet the needs of children with profound sensorineural hearing loss. Along the way there have been many challenges. Despite these, the process is succeeding. The speech toolkit features and Baldi are being improved. Applications designed with the toolkit are used by deaf children to improve their listening and speech skills, expand their vocabulary, practice language skills and review concepts in academic areas. This paper describes the challenges this project has faced, how the toolkit has developed and some specific applications created by grant team members.

1. INTRODUCTION

The speech toolkit began as general-purpose software for building interactive computer-based speech systems [4]. The toolkit currently uses speech recognition, speech synthesis and an animated face, called Baldi, developed at the Perceptual Science Laboratory [1]. The speech toolkit's intuitive drag-and-drop graphical interface initially allowed programmers to build complex spoken dialogue systems for research and business applications. But could this toolkit be used for educational applications? Could classroom teachers use the tools to create speech-driven computer lessons for students? The toolkit's developers knew that although programmers easily used the toolkit, much research and modification would be needed before it was teacher-friendly.

In September 1997, developers of the speech toolkit (from the Center of Spoken Language Understanding at the Oregon Graduate Institute) in collaboration with developers of Baldi (from the Perceptual Science Laboratory at the University of California, Santa Cruz) were awarded an NSF Challenge grant. The purpose of this three-year grant was to improve Baldi and the toolkit as well as build learning and language applications for profoundly deaf children at the Tucker-Maxon Oral School in Portland, Oregon [3]. Profoundly deaf speakers benefit from as many language-input opportunities as possible. Since Baldi had been developed to be very accurate in his visible speech movements [2], he would greatly enrich the language experience of these deaf children.

This grant was to be guided by principles of participatory design. Participatory design is an active team process in which all team members participate in the planning and design of a product or program. In the NSF grant, programmers, researchers and educators have come together, to be involved in all phases of the software development process using the speech toolkit.

2. SOFTWARE DESIGN CHALLENGES

There have been several challenges to ensuring the toolkit's success in the classrooms at Tucker-Maxon. The first of these was to teach teachers who have little to no technical expertise how to use the toolkit and to teach programmers and researchers about methodology and curricula in deaf education. Another challenge was to improve the user interface and core speech technologies, so they would be more appropriate for classroom environments. A final challenge was coordinating members of the grant team. Team members include toolkit system and application programmers, researchers (both graduate students and professors), teachers, and to some extent even the students themselves. They necessarily work at several different locations (OGI, TM, UC and UCSC). Consequently, there are limited opportunities for face to face whole group meetings. Through training courses, classroom observations, discussions, and electronic communication, the teachers at Tucker-Maxon and the developers from OGI and UCSC have worked through these challenges.

2.1. Continuous Education

The integration of the toolkit in the school began by teaching the educators how to use the toolkit in an intensive weeklong training course. In this course, programmers learned about the teaching of speech and language to children with profound hearing loss. Since the toolkit has been set up in the classrooms, on-site researchers observe classroom activities and learn from the teachers the types of applications they want to create. The researchers then help the teachers and tutor them on how to make the target lesson using the toolkit. They also communicate to the programmers about the most commonly requested functions, design features, and software limitations. This continual education has been bi-directional. While the toolkit facilitates the rapid development of spoken language dialogues, more advanced functions have needed additional coding. Novice computer users, like teachers, cannot be expected to learn a programming language to make daily lessons. While researchers and programmers have provided tutorials on how to include these advanced modules in daily lessons, programmers have learned from teachers and researchers about how to package these commonly requested functions into intuitive and simple user interfaces.

2.2. The hardware interface

At Tucker-Maxon the toolkit and Baldi are installed on Pentium

Pro 200 machines with 64 megabytes of RAM. Each has a SoundBlaster compatible audio card and OpenGL capable video card. The operating system of all of the machines is Windows NT. The audio output is fed through a peripheral amplifier. Use of the amplifier not only allows for greater volume control, but enables students to connect their audio input devices directly into channels containing the computer's audio output. This ensures that the students receive the best audio signal possible and that this signal is not in competition with any background noise in the classroom.

2.3. Improvement of the software interface

In the speech toolkit language lessons center around dialogues with Baldi. These lessons are made with a graphical authoring tool, called the Rapid Application Developer (RAD). When RAD is launched, both a blank canvas and a tool palate are presented (Figure 1). The palate consists of "objects" that can be dragged to and dropped on the canvas. These objects invoke the distributed technologies in the toolkit, such as speech synthesis, recognition, and animation. A dialogue author creates a lesson by placing objects on the canvas and connecting them with lines to indicate sequence. When students run the lesson, they listen to and respond to Baldi. Their utterances are processed by the recognition technology. These events occur in the sequence and manners determined by the dialogue author and are represented by the underlying arrangement of toolkit objects on the canvas.

When RAD was introduced to the classrooms at Tucker-Maxon it looked quite different from the way it does today. As the result of participatory design, the software has improved dramatically. This is best exemplified by describing the first uses of the toolkit in the classrooms and how the toolkit has changed to meet the needs of the teachers.

After having been exposed to the speech toolkit in a short course, the fifth grade teacher knew what he wanted the toolkit to be able to do in his classroom. He envisioned a toolkit lesson that reinforced class content in a dialogue between the animated agent and the student. At the time, the students were studying state names in the U.S. He wanted to display a map with one state circled and have the computer ask, "What State is this?" After the student responded, the computer would evaluate the response. If correct, the lesson would progress to the next question. If incorrect, the computer would provide a practice session where the student could listen to and watch Baldi articulate the word.

The design of this toolkit application consisted of a number of dialogue states, each containing a question, speech recognition vocabulary, and some code that displayed and removed the picture. This application was very large in terms of the number of objects the user had to "drag-and-drop" on the toolkit canvas. Basically, there had to be a dialogue object for every question in the application. Creating toolkit lessons with these dialogue states were time consuming and rigid. The teacher said that editing each object took a lot of time and that, since the order of the lesson was always the same, students quickly learned the sequence and were not being challenged as much as he hoped. He also observed that, if the students missed an aural cue indicating that the response was incorrect, they assumed they had proceeded to the next question when they were actually in practice mode.

After discussion with the on-site researchers, a programmer added coded features to address the teacher's concerns. A randomization function was added to mix up the presentation of questions; eventually, the question and answer practice modes became more obvious to the students by having Baldi move to the other side of the screen when the response was incorrect. Yet, these improvements also had a negative side. For example, the teacher had to be very careful when editing the programming language, so critical spaces, commas or quotation marks were not accidentally removed. Additionally, the image display and randomization features added to this lesson were only available in the template for this specific application. Teachers wanted these functions available for any application made with the toolkit.

Now, these toolkit features and many more are bundled into various new objects that the dialogue author can drop into an application. For example, a media object provides a graphical interface for presenting images, sounds, text, and image maps. A list object allows the teachers to enter all the questions, answers, images and sounds in one state object, making applications much quicker to build and edit. Finally, the author can easily modify the conversational agent's placement and appearance via a graphical tool.

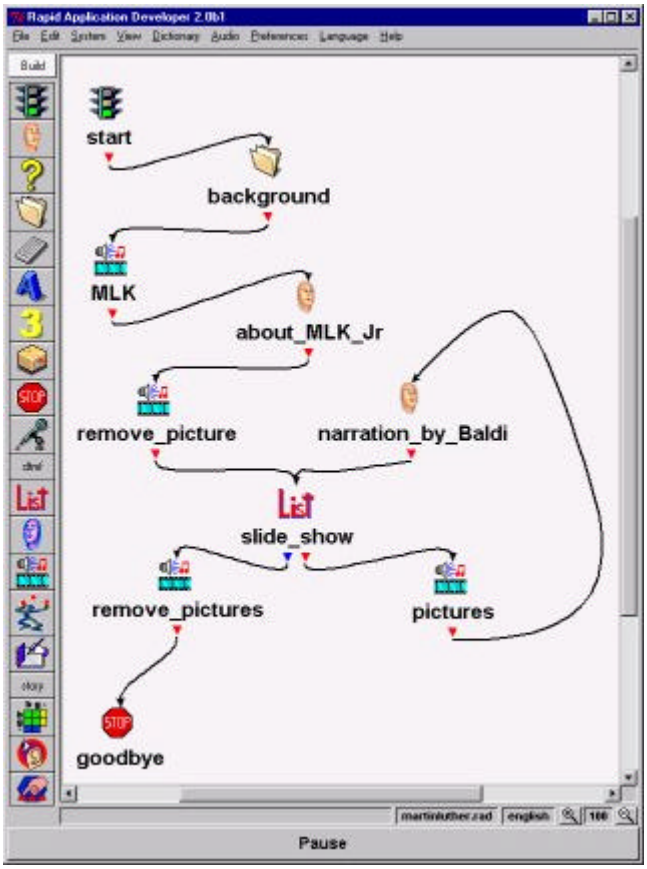

Figure 1. The Rapid Application Developer

2.4. Communicating with team members

A critical component of participatory design is the communication between the developers of the technology and the users of the technology. By November 1997, TMOS had two researchers on-site to serve as liaisons. At this point, teachers began meeting individually with researchers to plan and create applications for their classrooms. These meetings were usually once or twice a week after school or during planning periods. In addition, all of the teachers involved in the grant, the TMOS director and the on-site researchers met weekly as a large group to discuss new applications wants and needs. The researchers then passed this information on to the programmers. The programmers occasionally came to these meeting to review goals, discuss problems and provide information on the most current toolkit revisions. These individual and group meetings continue and are currently the major way for facilitating continuous education and smooth communication. The on-site researchers have linked programmers and teachers by filing weekly progress reports, keeping a "wish list" of the school's wants and needs, and tracking when and how these wishes have been met. These reports and lists are e-mailed to all team members and published on the project Web site. When new features are added to the toolkit, the teachers and researchers provide feedback to the programmers. Therefore, at various levels, communication loops are in place to handle the different aspects of the participatory nature of the project.

3. TODAY'S TOOLKIT IN THE CLASSROOM

There are currently fifteen computers installed with the speech toolkit at Tucker-Maxon. They are being used daily in the classrooms of first through fifth grades.

3.1. Fifth Grade Classroom

The students in this class use the speech toolkit every day in many different ways. They use Baldi to learn new vocabulary, practice individual speech segments and review new concepts in science, math and social studies. Students also program their own Baldi dialogues for use by other students in their classroom.

During a typical week, the teacher determines the vocabulary students need to practice. These vocabulary words are primarily from the literature they are reading. The teacher then types these words, definitions and sentences onto a notepad and saves into his vocabulary folder. When the student runs the toolkit, this list is read by the program. Baldi helps the students say an individualized list of words, the definitions and sentences.

The teacher also designed a speech practice application called SPLUSH. Every week he updates the speech sounds he wants each student to get additional practice with using the toolkit. In this application, Baldi presents target syllables and words for the student to repeat. The students' utterances are processed by the recognition system. If their production receives a passing recognition score, Baldi gives them positive oral feedback and presents the next target. If the utterance is scored below a passing threshold, the application switches to a practice tutorial. In the practice subdialogue, students compare their recorded production to Baldi's model by clicking on the buttons "me" and "Baldi". Here, students can listen and practice the target as long as they wish and then reenter the tutorial to have Baldi check their progress. Use of the toolkit has become an integral part of the curriculum in this classroom. The teacher simply updates the program each week with new syllables, words and sentences.

3.2. First Grade Classroom

Students in this class use the speech toolkit and Baldi to practice vocabulary and language throughout the school day. They make "Daily News" applications where, with the help of a teacher, they scan images and create a question and answer dialogue about current events. For example, Baldi will ask "What is STUDENT NAME's news for today?" After Baldi evaluates the response, he first asks the student to "Tell me about that," waits for their response and then moves to the next student's news item. Throughout the week, students use this application for review and vocabulary practice.

Students also learn to recognize and pronounce vocabulary they encounter throughout the day. In these applications, students click on various pictures on the canvas listen to and watch Baldi's say the word and then record their own production (Figure 2). In a similar activity, students practice their auditory skills by discriminating between words such as shoot vs. sheet. At the end of both practice activities, students are rewarded with praise from Baldi, a song and stickers.

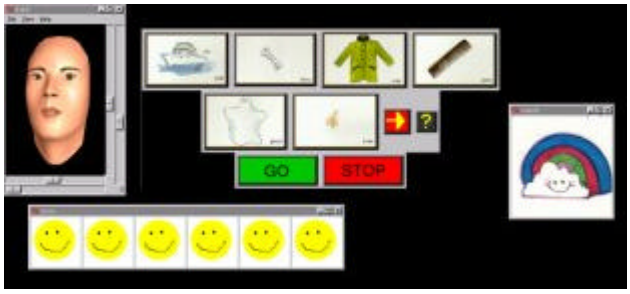

Figure 2. Daily vocabulary lesson

3.3. Third Grade Classroom

The five students in this class use Baldi and toolkit lessons to supplement classroom curriculum. They use applications made by the teacher and learn how to create their own. A particular favorite of the students is a concentration game where they match an auditory vocabulary word to its picture.

Recently, students use the toolkit to make reading comprehension and vocabulary activities from their own stories. In the application this month, Baldi reads a story the students wrote about dinosaurs. He then asks comprehension questions that the children helped write, and evaluates their pronunciation of the response. Next, Baldi provides definitions of key vocabulary in the story and the students supply the word. These comprehension and vocabulary questions are presented randomly and Baldi tells students whether they received a passing recognition score on the response. This feedback lets students know, first, whether they have answered the question correctly, and, second, whether their pronunciation received a passing recognition score. Exercises like these are valuable to

the students as they not only provide an opportunity for students to publish and present their own class work, but also design and work through oral language lessons.

3.4. Speech Therapy

Listening Tutor is a speech perception and production application. It consists of three activities that can be completed by students independently. The first activity is an auditory discrimination task in which minimal word pairs, like "bag" and "wag" are presented. Students click on buttons to indicate whether the two words heard are the "same" or "different." If the students are not able to discriminate between the two words, they receive visual speech input from Baldi. They have the opportunity to see Baldi say the words as many times as they like before returning to the auditory only discrimination task. Baldi's help will encourage students to transfer information from the audio-visual model to the auditory only model. Along the way, visual reinforcement, in the form of stickers, is given for each student response.

The second activity is an auditory identification task in which two pictures of minimal word pairs are presented. An audio file of one of the words is played and students click on the picture of the word they hear. Visual reinforcement (i.e., a sticker) is also provided. In the third part of the Listening Tutor application, students record their own speech productions of the word targets. Once recorded, the words can be played back by the student, re-recorded or saved for the teacher. The recording of each student's productions are placed in a file that can be played by the teacher at anytime. The Listening Tutor is based on a hierarchy of auditory perception skills. Initially students discriminate between pairs differing in place, manner and voicing (e.g., first consonants of "man" and "can"). Students gradually progress to discrimination activities that are more sophisticated. That is, word pairs differing in one aspect (e.g., the first consonants of "tan" and "can").

Transparent Speech Tutor is a speech production application that can be completed by students independently or with the teacher. In this application, Baldi asks the student to imitate single words that have a speech target in common (e.g. final /l/ words, like "bell" and "whale"). After a picture of a bell is presented, Baldi says, "bell." Students are then given buttons for recording, playing back or re-recording their own speech production of the word target. They can also listen and watch a transparent Baldi produce the target word again. Transparent Baldi has transparent skin so that the internal articulation is clearly visible. Students or teacher can manipulate Baldi to produce the target word in slow motion, rotate his head to observe the production from different angles and drop his skin and teeth to better observe the tongue movement of the talking head. Like the production activity of the Listening Tutor, once students record their own speech production of the word target, these words can be played back, re-recorded or saved for the teacher. Speech recognition technology limits the possibility to play students utterances aligned with Baldi's animation in the Transparent Tutor. However, with this application, students can observe the placement and movement of Baldi's articulators during the production of word targets. Baldi's abilities to 1) show these placements and movements from inside the oral cavity and from a variety of angles in exactly the same manner each time and 2) slow and stop these articulatory movements to whatever rate and at whatever place the student desires are not humanly possible. That is why the transparent Baldi feature is a very valuable addition to the armentarium of all speech teachers and a very useful aide for students refining their speech production.

4. CONCLUSION

Participatory design has coordinated the talents of members for improving technology and meeting challenges. Continuous education has facilitated understanding between autonomous team members and consolidated multiple intelligences in the design of speech technology software for the classrooms of children with profound hearing loss. Under this successful model, research and cooperation will continue to improve the toolkit for educational uses and language instruction.

ACKNOWLEDGEMENTS

This work was supported in part by NSF grant ECS-9726645, NSF Challenge Grant CDA-9726363, a joint grant from the Office of Naval Research and DARPA, and Intel Corporation.

REFERENCES

[1] http://mambo.ucsc.edu/psl

[2] Massaro, D., Perceiving Talking Faces: From Speech Perception to a Behavioral Principle, MIT Press, Cambridge, 1998.

[3] Stone, P. (1999). Revolutionizing language instruction in oral deaf education. In *Proceedings of the 14th International Congress of Phonetic Sciences*; San Francisco, August 1999.

[4] Sutton, S., Cole, R., et al. "Universal Speech Tools: the CSLU Toolkit," In *Proceedings of the International Conference on Spoken Language Processing (ICSLP)*, pages 3221-3224, Sydney, Australia, November 1998.