



BRINGING SPOKEN LANGUAGE SYSTEMS TO THE CLASSROOM

Sutton, S., Kaiser, E., Cronk, A. and Cole, R.

Center for Spoken Language Understanding
Dept. Of Computer Science & Engineering
Oregon Graduate Institute of Science and Technology
PO Box 91000, Oregon 97291-1000, USA.

sutton@cse.ogi.edu
<http://www.cse.ogi.edu/CSLU/>

ABSTRACT

Currently, there are few opportunities for people to learn about and experiment with the latest spoken language technology. Furthermore, most research and development activities are restricted to a handful of academic and industrial labs. In order to make the technology less exclusive, it must become more accessible to the general population. This is now feasible with the development of the CSLU Toolkit which combines easy-to-use authoring tools with state-of-the-art human language technology. In this paper, we focus on the educational role of the toolkit and describe how it is being used in several local schools.

1. INTRODUCTION

Research and development of spoken language systems is currently limited to relatively few academic and industrial laboratories. Because building such systems requires multidisciplinary expertise, sophisticated system development tools, language resources (e.g., pronunciation dictionaries), substantial computer resources and advanced technologies such as speech recognition and text-to-speech synthesis, few undergraduate institutions or high schools are able to offer students the opportunity to develop and use spoken language systems.

We are thus faced with a dilemma: advances in the development of spoken language systems are needed to allow ordinary people to participate in the information age, yet we cannot offer our students opportunities to learn about and use these systems.

Our goal is to engage a generation of students in interactive systems technology by bringing the technology into the classrooms. We aim to provide the necessary tools and infrastructure that will make it

possible to offer instruction on spoken language technology in schools and universities everywhere.

In this paper, we outline the spoken language software environment that is under development and describe participatory design activities involving short courses taught to various populations, including middle school and high school students.

2. TOOLS, RESOURCES AND INFRASTRUCTURE FOR LEARNING

We are developing the CSLU Toolkit, an integrated software environment that incorporates some of the latest research advances in spoken language understanding, text-to-speech synthesis, dialogue modeling and the rendering of realistic talking faces (Sutton et al., 1996; Schalkwyk et al., 1997).

The Toolkit includes easy-to-use graphical authoring tools that allow the rapid prototyping and testing of spoken language systems for a wide range of tasks. Figure 1 shows the graphical authoring environment (CSLU Rapid Prototyper) and talking face (Baldi) in use.¹

From an educational standpoint the Toolkit provides the opportunity for “hands-on” learning, exploration and experimentation with speech technologies.

The Toolkit is available for educational and non-commercial purposes via the world-wide web (<http://www.cse.ogi.edu/CSLU/>). In addition, on-line documentation, support and teaching material, including manuals, tutorials, frequently-asked questions pages (FAQs) and a discussion board are available from this web site.

¹ The talking face technology (Baldi) was developed by Dominic Massaro and Michael Cohen at the University of California, Santa Cruz.

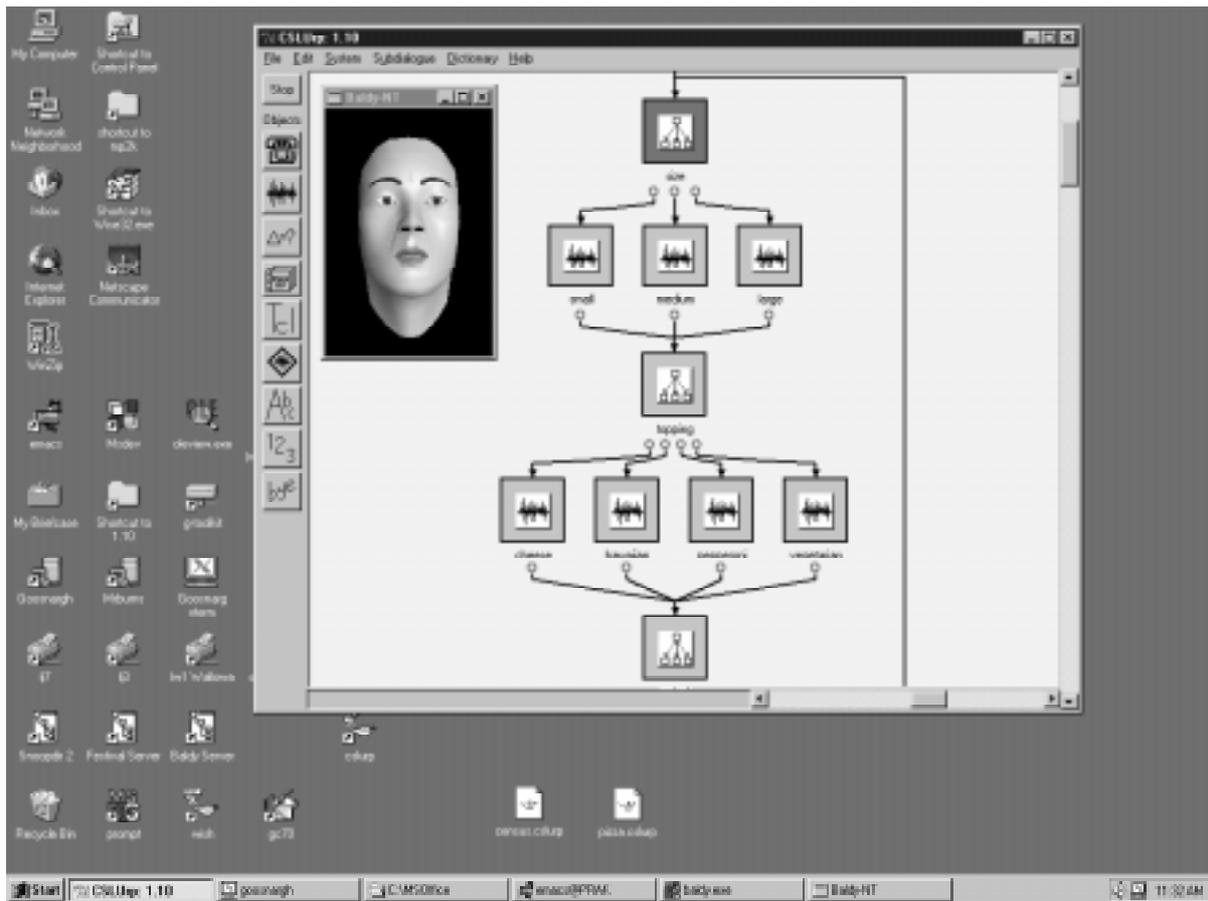


Figure 1: Components of the CSLU Toolkit running in Windows95

3. CLASSROOM EXPERIENCES

We employed a participatory design methodology in developing the Toolkit—the notion that users should play an active part in every phase of the design process (Muller & Kuhn, 1993). Accordingly, we have offered courses to students at different levels. By observing how students use the software, what difficulties they encounter and by collecting feedback, we are able to identify and improve on any problem areas.

To date, we have held short courses for a number of populations, including:²

- Professionals (Summer 1995—see Colton et al., 1995—and Summer 1996),
- High-school students and teachers (Winter 1996),
- Middle-school students (Spring 1997)
- High-school social studies class. (Spring 1997)

² Courses involving local schools were organized in collaboration with the NW Regional Educational Service District and the Saturday Academy.

Each course incorporated a set of assignments intended to introduce participants to the basic toolkit capabilities and skills needed to develop spoken language systems. The assignments walked participants through building increasingly more complex systems, such as pizza ordering, telephone surveys and web access by voice.

The final part of each course involved participants designing and developing their own project. This was an opportunity for students to apply what they had learned in a creative manner and to use speech in a way that might enhance the performance of ordinary tasks. The projects exhibited several notable properties:

1. the diverse range of tasks covered;
2. the short amount of time available for the implementation—typically less than eight hours; and
3. the limited amount of training time that took place to learn how to use the tools.

Participants gained insights into a range of technical processes, including speech recognition fundamentals and dialogue design issues such as, how to phrase prompts, how to select and model the recognition vocabulary, and how to detect and recover from

miscommunications. Details of the various courses are given in the following sections.

3.1 Course for Professionals

In Summer 1996, we offered a trial course to a group of professionals from a range of backgrounds. These included teachers, marketing executives and software consultants. All had basic computer skills, although only a handful had any experience in programming a computer.

Since this was early on in the development of the CSLU Toolkit, we learned many useful things about the design and usability of the authoring tool interface. The kinds of problems encountered included interface issues arising from differences between Unix and Windows95. Having identified this problem we have now changed the interface to adopt the native look and feel of the operating system being used. There have also been several architecture issues associated with supporting multiple users.

The student projects included:

- Airline Reservations: simulates reserving a flight using real airline information from the WWW;
- Homework Help-Line: telephone-access to homework assignments on a WWW page;
- Wetlands Data-Entry: enter water-quality data onto a HTML page by voice;
- Personal Assistant: manages appointments, takes messages, etc.
- MovieLine: provides movie information including show times and brief descriptions;

3.2 Course for High School Students and Teachers

In October 1996, we offered a three-day short course to students and teachers from ten local high schools in Oregon. The invited students were typically above-average and fairly familiar with computers. The purpose of this experiment was to determine the feasibility of using the Toolkit with a younger age group. As it turned out, the partnering teachers mostly played an auxiliary role since the students dominated the computer.

Example projects included:

- Elroy: a telephone-speech version of a computer psychoanalyst program;
- Speech-Tech: provides speech therapy and articulatory training;
- Class Registration: enrol for classes;

- VoiceE-mail: record and retrieve structured E-mail messages;
- Traffic Reports: access and submit reports on local traffic conditions;
- WWW Voice Browser: enables access and navigation of the WWW by voice;
- Computer Help-Line: troubleshooting for hardware and software problems;
- Science Trivia: question-answer game supports levels of difficulty and awards prizes; and
- Election On-line: provides the latest ballot-measure results from the WWW

3.3 Course for Middle School Students

In April 1997, an after-school enrichment workshop was held for a middle school students. This was a 16-hour course divided into four classes. Most of the students had general experience using computers (e.g., word processing and chat sessions) but no programming experience. Classes consisted of three kinds of learning activity: interactive lectures, laboratory exercises and a self-selected project.

The lectures reviewed basic topics relevant to speech, such as the propagation of sound, the human hearing anatomy, the psycho-acoustic phenomenon of hearing and sampling theory. Lab exercises were designed to familiarize students with the graphical authoring tool. The tool serves as a kind of visual programming interface which is ideal for introducing younger, inexperienced students to programming and system development. The student projects involved building spoken language systems that provide voice-access to particular world-wide web sites by voice. The projects included:

- Weather report: provides the current weather conditions and the week's forecast for major cities around the country;
- MTV top 20: provides the titles and artists of the current top 20 music videos, posted by MTV; and
- Cats: provides a description of numerous types of cats.

3.4 Course for High School Social Studies

Also in April 1997, a social studies class from a high school in rural Oregon took part in a course to create interactive dialogues with historic figures. The course took place over several weeks. The motivation for the course was to engage a group of average students in a novel learning environment. Many of the students had very little experience with using computers.

Students chose either a notable historic event or historic figure as the topic for their interactive system. Examples included Neil Armstrong, Martin Luther King, and the Vietnam War. They conducted research about their chosen topic—using both books and the Internet as a resources. The resulting projects allowed the user to interact with and learn about the particular person or event by asking simple questions.

4. DISCUSSION

While the practice of making the latest spoken language technologies available in schools may seem appealing, there are issues concerning its educational value. For instance, how can we know whether the technology enriches the learning experience or improves a student's ability to learn?

In our early experiments, we gauged the success of the courses: 1) by observing the degree to which participants were able to complete the assignments; and 2) through feedback from participants in the form of an evaluation questionnaire.

We have learned valuable lessons from our participatory design experiments, in terms of what works and what does not, what technologies need to be improved, what features need to be added and which aspects of the teaching materials are effective. Some good things that arise from using the Toolkit as a learning environment are that it fosters critical thinking and hands-on learning about how spoken dialogues should be used to perform tasks. It allows students to experience and become comfortable with new technologies such as speech recognition and text-to-speech synthesis. It enables students to express themselves through the systems they create, in the way that they design prompts and dialogues to carry out a task. In contrast, some bad aspects are that there is still a lot of crafting and tuning that must take place in order to make a system work well and that such practices are often hard to learn.

As a measure of the interest generated by the courses, ten of the participating schools now each have a donated computer running the CSLU Toolkit. Several of the original students have continued to work with the Toolkit and are sharing their knowledge and experiences with other new students.

The Toolkit was recently ported to Windows95 and Windows NT. This will make it possible to introduce the Toolkit into more classrooms in the NW Regional Education Service District and allow us to monitor and evaluate the benefits of this on-going effort.

In addition to our involvement with the local school district, we have established partnerships with other

community leaders who serve diverse populations including disabled people, under-represented minorities and under-privileged people.

5. ACKNOWLEDGMENTS

Special thanks go to Mark Fanty, Jacques de Villiers and Johan Schalkwyk, co-developers of the CSLU Toolkit. The work described in this paper was sponsored in part by grants from NSF, ONR, DARPA, and the membership companies of the Center for Spoken Language Understanding. The courses were made possible with computers donated by Intel.

6. REFERENCES

Cole, R., Sutton, S., Fanty, M., Kaiser, E., Schalkwyk, J., de Villiers, J., Cronk, A., Colton, D., "Cyberspeech: Password to Cyberspace," Proceedings of the DIAC Workshop: Community Space and Cyberspace, Seattle, WA, March 1997.

Colton, D., Cole, R., Novick, D., Sutton, S. "A Laboratory Course for Designing and Testing Spoken Dialogue Systems," Proceedings of the 1996 International Conference on Acoustics, Speech and Signal Processing, Atlanta, GA, 1129-1132, May, 1996.

Muller, M., Kuhn, S., (eds.) "Special Issue: Participatory Design," Communications of the ACM, June 1993.

Schalkwyk, J., de Villiers, J., van Vuuren, S., Vermeulen, P., "CSLUsh: An Extendible Research Environment," To appear in Proceeding of EuroSpeech97, Rhodes, Greece, 1997.

Sutton, S., Novick, D.G., Cole, R., Vermeulen, P., de Villiers, J., Schalkwyk, J. and Fanty, M., "Building 10,000 Spoken-Dialogue Systems," Proceedings of the 1996 International Conference on Spoken Language Processing, Philadelphia, PA, 709-712, October, 1996.