ROUGH EDITED COPY

RIT/NTID

An International Symposium
Technology and Deaf Education

"An Analysis of Text Generation
And Information Display Methods"

Presenters:

Peter Lalley
Daniel Bogaard
Donna Hager

Session T10E
June 24, 2008

* * * * *

* * * * *

>>MODERATOR: Well, we're all ready. Why don't we go ahead and get started. Good morning. This is workshop session T10, Development of Display Modes for Including Text And Non-text information. So hopefully everyone's in the right place. We have four presenters.

We also have evaluation forms that we'll be handing out right at the very beginning just to make sure that you have them so that once we're all done here, you can go ahead and fill those out for us. And make sure that you include the number T10E.

Okay. You can also fill these out online if you would like. The interpreters today are Sarah and Jim, and I think we're going to go ahead and get started. So thank you.

>>PETER LALLEY: Thank you to Sarah and to Jim. My name is -- I'll let you go -- my name is Peter Lalley, and I have a group with me from -- Dan Bogaard, Cathy Beaton from information of technology department and Donna Hager from Monroe Community College.

At the end of this, we hope we have enough time to demonstrate the speech-to-text and the use of what we call heads-up display.

This study started from a discussion with Leanne West at Georgia Tech and Cathy Beaton and myself who had a concern that a large number of students arrive here on campus without any sign language skills.

So -- and they -- they have grown up in high schools, mainstream high schools, and what we wanted to do was work on ways that we could provide them with choices in the classroom on what kind of access services they would prefer.

So what we decided to do was to set up a program, a research program, to evaluate three different speech-to-text systems, and to develop ways of presenting this information in a unified screen so that students sitting in a classroom would not have to be looking at multiple inputs of information.

For example, the interpreter, PowerPoint or video, some other thing that's going on. So if we could unify all of that into one screen with a text coming up on that screen, we think that might provide more choices to the individual students.

We know that one system of access doesn't fit all people so we're just looking for ways to broaden their opportunities.

We start, we decided to look at CART, C-Print and automatic speech recognition. And what we're using for the automatic speech recognition is Dragon Version 9. Earlier versions were not adequate to use in the situation, but with developments of their software and so forth, we felt Version 9 could be used.

And after we evaluated those three speech-to-text systems, what we wanted -- we had students would come in. We would use all three of these systems and get their feedback on it.

And test their comprehension. And then we would pick one of those systems to go on to the display mode.
This is a very brief description of the way we did the first part. Students would come in to the room. In each session we had 18 students participating. We divided them up into three groups of six each. We set up the three different systems and Donna Hager presented various lectures about a, what would you say, first, second-year college?
>> Yes. Absolutely. First-year. Intro.
>>PETER LALLEY: Intro courses in business.
We would have a ten-minute lecture, and then we would stop, give them a comprehension test and get their feedback on a particular speech-to-text system that they used. And then we rotated the students and each one moved to another tech system, we gave them another lecture, comprehension test, feedback on the system, and at the end of the process, we had a focus group of all of the students and got lots and lots of feedback on the different systems.
These are the demographics for the group that we did on speech-to-text. Oops. Wrong one.
We have background information on what was the first, second or third language. You can see we have a fair diversity.
We even had students, Arabic and Farsi was their first language.
Many of whom English was first, ASL second, so forth.
So diversity of communication needs in this group.
We know, we have the background on the age that they became deaf. Most of them became deaf very early in life.
And information on what kind of assistive technologies they used in high school before they came here. So we know a lot about these students. And the students we had in this were students that had been on campus for probably two to three years. They were in bachelor programs, most of them. There were a few in the associate degree programs. They had experienced other kinds of access services here, and they were in a wide range of majors, probably representative of probably about 30 or 40 different majors.
What I'm going to show you now is examples, and if you have the copies of the PowerPoint, examples from the lectures of what the students saw with CART, and you can see CART is very, very good. Very good. Very accurate. And it comes up on the screen rapidly.
The next one is C-Print, which is the same section of the lecture that you saw previously, and that's what the students actually saw on the screen. I would say the accuracy in general is well above 95% on this.
And here's the same section of the lecture with ASR. If you read quickly through here, what you see is not in there is a lot of the punctuation. There's ways of putting that

punctuation in. There's two ways. One, you can say,

"Period, comma" so forth. The second way is you can work
with pausing your speech. It will automatically put a
period in or a comma. Any of the other punctuations, you
have to add speaking the word.
But what we saw from this was that the text from all three
systems were fairly -- or were accurate. And fairly close
in quality.
This is a screen shot of a dragging caption in one of the
classes. Do you want a picture?
So if you were a student sitting in class with a faculty
member using Dragon, this is what you would see if you just
had a text-based display.
During the focus groups, we asked the students, would you
use any of these systems in class. And what you see is that
almost I'd say it's two-thirds of the students said they
would use CART, said they would use C-Print and very low
number that wouldn't use ASR.
And what had happened was that on a first test, we found out
that just before the first test, we found out that Dragon
only goes one to one. It doesn't broadcast to multiple
computers. So we had to build a software bridge using
Python and that was done by two students working with us, Ed
Brannin and Chris Muench who are now off and working. And
at various times, the ASR would collapse. All of a sudden
it just rolled through it and it was a very bad impression
on the students.
The second test, after we put that software in to fix it,
was much, much better.
And when we had Donna doing a lecture, it was much, much
better.
When we gave them the comprehension test, miscellaneous
meant these were three lectures, kayaking, Rochester, oh,
and web cookies. And that was the miscellaneous. The
academic lectures were on three different areas of business,
corporation --
>> Employee motivation, forms of business and small business
enterprise.
>>PETER LALLEY: When they did the comprehension test,
scores were very, very good on the academic one. And ASR
actually finished the highest on those.
Now, we finished that part. We're proceeding to develop the
display modes. And what we decided to do was that we
decided to use ASR with the display modes for accuracy and
speed, CART and C-Print are excellent. We know what they

will do. We are not sure what ASR by a faculty member
would -- how that would work out. We knew we could get very
accurate text, but we wanted to use that.
And the reason, C-Print and CART -- C-Print's not so
expensive, but CART is very expensive on an hourly basis.
So we were looking for a system that was less expensive,

actually cheap, and could be used over a broad spectrum, and
the faculty members who wanted to take the time to train.
>>AUDIENCE MEMBER: Do I understand you correctly to be
saying that faculty didn't train their voices on the ASR
yet? You really haven't tried that?
>>PETER LALLEY: Yes, we have.
>>AUDIENCE MEMBER: You have?
>>PETER LALLEY: Yeah. The first time we used ASR, it was
like ten minutes of training -- 15.
>> 15 is what they said out of the box.
>>PETER LALLEY: And Donna trained it for --
>> About an hour.
>>PETER LALLEY: About an hour. And trained it according to
the directions and so forth.
>> And the first test was the results that were 39% correct.
The 89 -- or 81, whatever, we had from the second test,
that's when she trained it for an hour.
>>AUDIENCE MEMBER: And how do you encourage faculty to take
an interest in doing this?
>> Well, the interesting -- if they actually want to use
this, all right, and that's a hurdle, but if they want to
use it, the more they lecture, the more they train it. You
don't have to go through this every time you use it it gets
better.
>>PETER LALLEY: You have your voice profile, your
vocabulary and so forth put into the system.
The other thing that's --
>>AUDIENCE MEMBER: We tested the students to see which
system they preferred, but do you have their preferences in
your -- you know, in the computer, like can they choose what
their preference is? Is that another part of your process?
>>PETER LALLEY: We hope that as we go on in this that we
can build it so that if there are C-Print operators
available, if CART is available, they can switch and pick
their choice of text generation.
>>AUDIENCE MEMBER: In ASR, morning voice can be different
from afternoon voice. Is that controlled for?
>> Yeah.
>>PETER LALLEY: And then we move to Phase 2, which we're in

the process of doing now, looking -- we did the same kind of thing. We recruited students, divided them into groups, exposed them to three different ways of seeing the non-text information. One in a composite screen with a PC where the text comes up and then they have the PowerPoint.
Two, projecting the whole thing to a screen from the class, and then three, we used a system of a heads-up display that looks -- it looks a little -- Dan says geeky, but this is where the collaboration with Leanne west at Georgia Tech comes in. She has developed a system which is powered by wireless PDA which displays the text on the heads-up. But

when you look at the heads-up display you can look straight through it and you'll see whatever is on the screen.
I've tried it and after about 15 minutes, it was okay. It takes time to get used to.
That's the third way we're doing this display.
And what -- this is just a mock-up of what we're trying to do is to bring the PowerPoint together, the person presenting the lecture. A lot of students depend on looking at the presenter, and the text coming up some place.
This is just a mock-up. When we first showed, some of the students didn't like it this way, so we have a lot of developmental work to do here. This banner will not show up. We take that off.
And this is how the heads-up display looks. They have over the face. This is right in front, powered by this PDA.
And actually, this was done at a baseball game where they were getting the announcements and so forth and what's going on in the game. He's using his PDA and he's using the heads-up display.
Now, after we used this and the other systems, we got a spread of opinions on it. Now, these students didn't have time, more than three minutes, to get used to the heads-up displays, so some of them had some difficulty with it. This student said it strained his eyes, he didn't like the head set component, but he loved the PDA. That's good information to get.
Another student said he gets disoriented looking through it, and that was my first impression, getting a little bit disoriented.
And then this may be one of the geeky students? This is actually awesome. If we knew who said that, we would bring the student back and help us improve it.
On the display, some liked it, others didn't like it. What this is kind of telling us is supporting, at least our

rationale for doing this study, that students need choices,
which system to use that fits them better.
And again, that might be the same student.
I have to say that we used the room upstairs to do these
tests. The projection on the screen was not very good. It
was very dim. So I don't think we got a good test in that
first round. But it keeps coming up, they like the portable
hand held device.
Okay. Kind of in a summary, you have to develop a way of
distributing the output from ASR to multiple uses. So we
had to build that software.
We ran speech-to-text using two different computers. One
was an older, a little bit slower. The other was the new
Dell, yeah, like this.
The slower one -- and it was your computer. I don't know --
>> It was actually a Mac that was running parallel, so I was

virtually running PC because that's the only way the
software works, and it did okay in testing, but in the
actual physical test, it did poorly. So you need -- it's --
you're running process on top of process on top of process,
and it died. So we found out that you need a decent machine
to actually be able to pump this up out fast.
>>PETER LALLEY: Speed is very, very important in these
machines.
More training of the software than is advertised, and that
goes back to your statement.
The 15-minute training doesn't work. Not for this kind.
Pronunciation and pace is critical, but you don't have to
talk like this, boom, boom.
We are working to improve the output from Dragon so that it
doesn't come up as pars or come out in big chunks, but it
comes out quicker, and we're working to have a better
software to display one to many.
And clearly, a number of students really liked the PDA. And
the problem that we would face is that we don't have a way
of putting the non-text information on to that PDA so they
would still be looking up and down. If we can work out a
way with a slightly bigger screen, we might be able to put
it all on that PDA.
And these are the people that have been involved. Dan,
Cathy, you're here, Donna's here, I'm sorry I don't have
your -- she's supposed to be totally unbiassed.
Janet MacLeod does the evaluation for us. She's a staff
member in the department of research, and these are the
students that have worked with us. Chris Muench is now in

North Carolina at IBM. Where is Ed? Ed's working some place in a tech program. Allison is still a student here. And Leanne West and her research student, Ethan Adler from Georgia Institute of Technology.
So this is a brief overview of where we're at, what we encountered as we went through the process, where we're hoping we're headed. If you have any questions, we would all be happy to answer them. And Donna would be happy to give you a demonstration.
Yes?
>>AUDIENCE MEMBER: We have been working with Google for more than three years. And we have done research between Dragon and another type. And we have found that it's not easy to focus in the pronunciation and speech where you are working in a speech. So maybe you can use a third person to do that. It's not low cost, but maybe you can control your results. Just having one more person repeating what you are saying, you know, something like that.
The other thing I want to add is we are using free software, it's like a remote software to have access to different computers, and in this way you can split for more than one

computer, and the text for free.
It's not -- the delay is not as important as it could be because it's an internet-based, and it's good enough, and it's for free. So for -- you can use to share the information in more than one computer.
>>PETER LALLEY: That's good. Where are you from? I'd like to talk to you.
>>AUDIENCE MEMBER: I'm from Bogota Columbia.
>>PETER LALLEY: Okay. As far as the training, or getting faculty to use it, their speech patterns is critical, like you pointed out, and it may not be usable by some faculty members, like Dan.
>> I talk very fast.
>>PETER LALLEY: He's high-speed.
>> Go, go, go! Gotta go!
>>PETER LALLEY: So he would be eliminated immediately. Training Dragon, and I use it in my office, what I found was really important was when you first open up the package, you read a couple paragraphs, it gets used to your voice, and then it asks you to go into your files, and if you allow Dragon to go through your files, it picks up, you know, your E-mail or documents. It picks up all your vocabulary. It picks up your pattern of writing and speaking, and that really was very helpful in training.

I know a lot of people skip that step, but I think if you take that step, it works fairly well.
Now, distributing on a Web is a good idea. That's what you're working on? Have a Web-based system?
>>AUDIENCE MEMBER: Not exactly a Web. It's like remote access, so I can -- I can show you guys how it works.
>>PETER LALLEY: Any other questions? Mike?
>>AUDIENCE MEMBER: I just have a little bit more of a question about the comprehensive testing slide. If you could show that again, the comprehensive testing slide. Okay. That one. That one.
So just exactly what are those numbers there? I mean, they are studies of the words that are accurate?
>>PETER LALLEY: No. They're the scores on the tests that we gave them.
>>AUDIENCE MEMBER: So in other words, you gave students a test of the information.
>>PETER LALLEY: That was presented, yeah. And they -- they did the tests. There were five questions, multiple choice. Did the tests, and then we scored the test, and those are the mean scores from those tests from 67 students.
>>AUDIENCE MEMBER: Can you explain the results, you know, when the students who, you know, said that they didn't like the ASR in the beginning, and how did they end up getting this result, then? I mean, you said that scored last in their preference initially, but yet they have a higher

score. I mean, do you have any idea why that happened?
>>PETER LALLEY: Well, look at these two scores.
>>AUDIENCE MEMBER: Right. For the academic. Okay.
>>PETER LALLEY: The academic is when we improved the communication between ASR and the multiple computers, where we -- where Donna began to give the lectures. So it was much clearer. Now, Dan gave the lectures. And he was just too fast.
>>AUDIENCE MEMBER: So, I mean, the thing is, the students didn't like it in the beginning, and yet they have a high score. Why is that?
>>DANIEL BOGAARD: I think a lot of that was familiarity. They're very familiar with CART, and -- C-Print much more and somewhat with CART, and dragon was new. And some of -- we had a good smattering of different kinds of students, some of the really tech, geeky students like we have here at RIT, they kind of liked the new, but some of the others were, it's something new and I'm not that excited about it. So I think actually over time it would become more

"Likeable". It's comfort zone.
>>AUDIENCE MEMBER: Can I add just one thing, then? I think maybe a high score with that is because they had to work harder, possibly, than with either CART or C-Print. Because without the punctuation or anything you have to kind of work your way through it, more, and you have a class that's one hour or something, you know, for high school, you would get like totally overwhelmed, I mean, where it is now, I understand that it's going to be improving and everything. So, you know, if this was an all-day thing or a three-hour presentation then, you know, CART is probably better.
>>PETER LALLEY: That's a good point. That's something that we should look at. We didn't ask any of the questions that would get to your question. Meaning they had to focus more.
>>AUDIENCE MEMBER: Did you do any test statistical significance?
>>PETER LALLEY: Janet did do that. Janet MacLeod is doing that. We don't have it yet.
>>AUDIENCE MEMBER: ASR, does it exclude the sign language users?
>>PETER LALLEY: Yeah. There was no interpreting. It was all text in these tests. I think that would depend on how people want to use it. If you have a class with kids that are some really like interpreting, some like the text, I would say that you should provide both.
>> Actually, I have some responses, just some little things to elaborate. Would that be okay?
>>PETER LALLEY: Sure. But what you see up here is CART. Right? And it's very good. It's very good. And C-Print is -- the captionist we have here at RIT are very, very good. Well-trained. And Mike is the developer of C-Print.

>> Hello. My name is Donna Hager. I would like to respond to two of the questions that we had. One was in regards to training. It doesn't seem to be picking me up.
>>DANIEL BOGAARD: It does. It does. Let me explain some of the things you're seeing here. This is using the heads-up displays, which you all can come and play with if you'd like, so it's going to the PDA to a wireless network, and then to the heads-up display, and it's picking me up somewhat, too. What it's doing is figuring it out, figuring out what is being said, and then it's sending it.
So what happens is basically when she speaks, it's going out to the -- it's figuring out through Dragon and sending it out through the server, all right, and that's going out on a wireless signal to the PDA or to the heads-up displays.

>> So you could see everything I was saying?

>>DANIEL BOGAARD: Yes.

>> Okay. Regarding training the tool, I did spend an hour with the tool. One thing that really helped me was training the tool to not only recognize my typical speech and choices of terminology, but also to take my lectures and spend a few moments going over the atypical terminology, or those words that were larger, more complex, and getting the tool to recognize those terms prior to delivery.

Also, a thought came to mind regarding the comment about the students doing better academically and having to perhaps concentrate more. And yes, there is the chance that they would be slightly overwhelmed or stressed.

However, it may not be a bad thing that they need to concentrate on the content, because if they're scoring more highly, it is because they are less relaxed. If the terms are coming out clearly, they maybe sort of sit back and relax versus actually staying engaged in the information.

Just a comment.

Any questions? Yes?

>>AUDIENCE MEMBER: Have you looked at speed of presentation?

>> I have not.

>>PETER LALLEY: We know very fast speakers like Dan, it does not work well yet.

>> As an example, when I was at my home training the tool, at one point I was in my bedroom with my laptop and the tool training, and always being conscious that I could not ramble on rapidly like some.

My son came into the room and said, Mommy, I am going outside.

And so forgetting that I was training the tool, I started to yell -- not yell -- have you done your chores? Have you done your homework? Have you practiced your cello? I did all of that in very rapid succession, and it was garbled on the screen. So at some level you do need to be conscious.

Now, I don't feel I'm speaking at a snail's pace. I'm not a rapid speaker.

>>AUDIENCE MEMBER: I was just thinking that the higher scores may be a function of rate of presentation rather than the technology. If it's coming -- if a person is talking as they northerly do and it's coming across the print more quickly and you have to filter it more quickly and comprehend, that may lower the score, whereas the slower presentation may give you more think time, and actually

raise comprehension.
>>PETER LALLEY: Actually, when people speak very rapidly, the tool gets confused, and what's coming up is unintelligible sometimes.
>>AUDIENCE MEMBER: But a student comprehension, when this is working properly, a function of the technology or a function of the fact that they have more comprehension time to process the information.
>>PETER LALLEY: Could be. That's something that we should look at.
>>DANIEL BOGAARD: The thing I would like to reiterate is actually that with the heads-up displays on, and, yeah, they're a little weird to get used to after while, but once you get those actually on even for a little while, it seemed like the students would kind of flip it up and just watch the PDA much more.
All right. You know, I'm a complete tech geek, and even I -- I don't think it was worth the heads-up display, but that's what we're going to go solve -- try to solve some of these problems. Another thing we might try on our next testing is just well, what if you have a PDA, and if they have that.
That to me is relatively interesting because they can just carry that to any class and, you know, any class that you might have the professor and have the software running, it wouldn't be that, actually, difficult.
>>AUDIENCE MEMBER: Might be some of the disorientation caused by having to switch focus. I wonder whether you would consider putting a camera on the instructor and blending both modes into one display.
>>DANIEL BOGAARD: Actually, one of the things that -- besides testing the three systems, one of the original ideas we wanted was, well, why can't we make a Web site. I'm a big web guy. So why can't we make a web page that has the lecture or the text generation or an area for them to maybe take notes to chat or something else and record the whole thing, and then okay I missed part of the lecture. I can go back home and watch it again or do it distance and a lot of interesting things like that. So we've done some research in that area.
>>AUDIENCE MEMBER: But in real time?

>>DANIEL BOGAARD: Yes. Yeah.
>>PETER LALLEY: When Leanne West, her group, used the heads-up display, they were working with not academic people, not students. They were working with adults trying

to work our way that they could see the captions and movies.
So they could take that to a theatre, have that heads-up
display, the captions come right across, and they can be
watching the movie.
And as I recall her data, 84% of the people that used it
that way liked it. And they didn't feel like they were --
they looked odd or people would look at them and think
they're kind of strange people using something like that.
They never tested it in the classroom. So there may be some
differences in perception on how well the heads-up display
works as a captioning tool for movies versus how students
would look at it in the classroom.
And it does take some time to get used to, and the
disorientation is a problem.
>>AUDIENCE MEMBER: I was thinking that maybe the problem --
or the reason why the students wasn't wanting to use this
option, it was because of the delay. Maybe the delay is one
sentence where you stopped talking, so maybe it's better to
use CART or C-Print. That's because of the delay.
And the other thing I want to add is the third person
repeating what you are saying is in this case, instructor
has to be focused and the other thing is you can add
punctuation. So if you say "Point" it write a point. So if
you have one more person doing the job in -- I don't know,
inside or something like that, you can add punctuation, and
in this case get better results.
>>DANIEL BOGAARD: I agree. I mean, there's some of these
problems. The interesting thing about this is I think you
can reach a much wider audience because it's cheap. All
right. And if you can work out some of these bugs and, you
know, I -- I am, again, a big tech geek and, you know,
technology gets better all the time. Two years ago, Dragon
was horrid. It's getting better all the time.
So if -- you know, if it's possible that we could reach a
much larger audience because of the cost in the end, that
this might become very viable.
It will get smarter. Sometimes it does put punctuation in
because it figures out, okay, I have a noun, I have a verb,
I'm pretty much done on the sentence, right, and sometimes
it doesn't.
But taking actually out the human element, you're going
to --
>> Yes. I have one more comment relative to the tool. In
order to more readily organize the material, at the onset, I
think because we are out of our comfort zone, it appears
difficult to organize the material, to train the tool.

However, there are verbal cues or rules that we can use that will allow us to more accurately organize the material, almost it grammatically correct, things like "Scratch that" which causes the cool to go backwards and erase what we just said, as well as "New paragraph" which tells the tool to start a new paragraph so that the information is not run-on. Those types of tools or commands, if you will, make the tool more enhanced, and more easy to use.
At the onset it's more difficult. However, those things become rote and eventually the tool is not that difficult to use.
>>PETER LALLEY: I think we have to close. We can stay --
>>MODERATOR: One more. Last comment.
>>PETER LALLEY: Okay. One more.
>>AUDIENCE MEMBER: You said that it cost $200.
>>PETER LALLEY: This software cost 90. This software cost $90.
>>AUDIENCE MEMBER: Who -- is it two?
>>PETER LALLEY: Go ahead.
>>DANIEL BOGAARD: You need one copy of Dragon, all right, so one copy only, and then after that, actually for -- depends what your out put is. What we built -- if you're a member -- I will find the picture.
We built this. We built that, a different server, to output this to its -- to anything we want to output it to. So you get Dragon and then it actually translates it and it outputs it to -- you're not looking at the same thing I am. Sorry. That was kind of dumb.
So this, all right, and this is the visual we output to. And then when we started talking with Leanne West at Georgia tech, they have their own piece of software, but you only have one copy of Dragon.
>>AUDIENCE MEMBER: The version for college?
>>DANIEL BOGAARD: Yes. So one copy of Dragon for -- well, if you have five lectures going in five different rooms, you probably need five copies.
>>PETER LALLEY: It eats being distributed from that server to those students.
>>AUDIENCE MEMBER: When the student is using the text, can they save this information?
>>DANIEL BOGAARD: Yes. It's just -- it saves to the palm, and then you can put it up. Or we could -- you know, when Dragon is talking we can save it, we can put it up on a server so all the students could get it, whatever we want.
>>PETER LALLEY: I think we have to close.
>>DANIEL BOGAARD: We'll be around if you guys want to play with the heads-up.
>>PETER LALLEY: I thank Jim and Sarah and our wonderful CART operator in the back. I would be glancing at that. It was wonderful. Thank you.

>>MODERATOR: Okay. We do have to close, but before we do,
we just want to make sure that you do take the time to fill
out an evaluation, either on a hard copy here or online, and
please make sure to remember the number T10E.
>>PETER LALLEY: Thank you.