

## **Captions**

**(T230B)**

### **New Accomplishments Using Voice Recognition for Captioning of Chemistry Videotapes Made During Regular F2F Courses.**

**Robert H. Paine**

New Accomplishments Using Voice Recognition for Captioning of Chemistry Videotapes Made During Regular F2F Courses

Robert Paine

Rochester Institute of Technology

Presented at the Instructional Technology and Education of the Deaf Symposium, held at the National Technical Institute for the Deaf, Rochester, NY; June 25-29, 2001. <http://www.rit.edu/~techsym>

\*\*These captions were generated live during the above presentation and may contain spelling or grammatical errors common to real-time captioning. You are seeing exactly what the audience read as the presentation was given.

1 June 26, 2001.

2 Session T 230 B - "New Accomplishments Using

3 Voice Recognition For Captioning Of Chemistry

4 Videotapes Made During Regular F2F Courses".

5 Presenter: Robert Paine.

6 .

7 PAT BILLIES: I think we are ready to roll.

8 It is my pleasure to introduce you to Dr. Robert Paine,

9 a Professor of Chemistry at RIT, at the College of

10 Science. He is going to be talking to us this

11 afternoon about new accomplishments using voice  
12 recognition for captioning of chemistry videotapes made  
13 during regular face-to-face classes. Our interpreters  
14 this afternoon are Billy Ridout and Dwight Godwin and  
15 Woody Waga will be providing realtime services for us.  
16 Join me in welcoming Dr. Robert Paine.

17 >>ROBERT PAINE: Thank you. I'm not used to  
18 the mike. So if you can't hear me or not making sense,  
19 I will take questions at any time about anything. One  
20 of the most often asked questions is, how does a  
21 chemistry professor get into this situation. Well,  
22 with a lot of enthusiasm. One of the things that will  
23 allow me to get to what you are probably all interested  
24 in, voice recognition or captioning, requires a little  
25 preamble. So I've got some boilerplate slides up

2

1 here. This I pass on to the students because I like  
2 them to think outside the box. Many things we do today  
3 were ideas that were on star track. How many star  
4 track fans here? How many star wars fans? Chapter two  
5 is scheduled for May 22 next year. The little square  
6 for information is now what we now have as a computer  
7 and we used to call them floppies and things like

8 that. Eventually voice recognition will pick up where  
9 it was on star track and be able for me or you to be  
10 able to speak in this language and will printout in  
11 another language. So it will be a transmitter.  
12 RIT is well known and becoming more well  
13 known for its work with handicapped folks of all kinds,  
14 never registered in the sales pitches. Distance  
15 learning has been in existence at RIT since 1979. I  
16 have put my own definition up, and I will amplify a  
17 little bit. At anywhere, any time, and I don't see  
18 many people here that would qualify but it goes back to  
19 the old continental classroom and things that started  
20 on PBS, and is out grown and out texted almost always.  
21 It is the method of moving information, not people.  
22 This is kind of my definition and not  
23 everyone agrees with it. They like to think it can  
24 only be realtime or it can only be real separation.  
25 Distance learning is any time you can't get ahold of me

3

1 or any other professor, it ought to be feasible.

2 We have about 2,000 distance learning

3 students in all classrooms and combinations which

4 include the web. As I stand here now, I am responsible

5 for and am teaching five courses this summer. I have  
6 one student in Pakistan. I have a whole variety of  
7 students around the country, and we supply lecture  
8 material for them with videotapes which in the last 7  
9 years, everything that the university has is captioned  
10 now. I'd like to demonstration, chat session, lab  
11 courses and I'll leave that for another time. We try  
12 to think that distance learning can handle any students  
13 we handle in a face-to-face. Now I'd like to go over  
14 and see if I can switch this.

15 These are two terms creeping into the media,  
16 the printed media, F2F, means what a lot of folks used  
17 to call normal classes. I don't know how you could  
18 have a normal class. It is made up of individuals.  
19 Each individual is unique. So we created face-to-face  
20 for the lecture presentation like we are doing now, or  
21 recitation presentation. DL is getting sloppy. It can  
22 be called on line learning now as we fold in the web.  
23 West of the Rockies it is called distributed learning  
24 and that works for this.

25 The captioning I will describe in a few

4

1 moments has application to a variety of audiences. We

2 started out with hearing impaired students. We found,  
3 after our first experiment, that captioning works for  
4 everybody. Supposing you are sitting in a class and  
5 you are looking at a boring guy like myself and a word  
6 like enchalty, and when it is displayed, the software  
7 trade the student doesn't have to stop and figure out  
8 what it means. I apologize for doing that. Dyslexic  
9 students, don't have the same thinking process, slow  
10 readers, slow notetakers, students for whom English is  
11 a second language, persons who are not effective  
12 listening. I'm going to give you an example. I think  
13 you will understand this example. If you are  
14 listening, you might be doing it to me. The speaker  
15 says something and it starts a thought process. You  
16 get a question up here and all of a sudden you start  
17 going off side ways. In 15, 20 seconds you are  
18 thinking about something and you don't have any idea  
19 what the speaker is talking about. Now you will get  
20 back on course. That's what I mean by mountain goat  
21 thinking. If you have a videotape or a CD eventually  
22 that you can stop, then you can handle things like  
23 this. One of the best memory training, if you are a  
24 mountain goat thinker is to think to yourself, how did

25 I get here and work your way back.

5

1 This is where we made our videotapes, and I

2 don't have a pointer. But you see how the white bar is

3 outlined. That is just the shape of the video picture,

4 so now the professor is using the white bar -- thank

5 you. So now the professor would have it half out.

6 This display station is what the student sees. The

7 studio will hold 70 people, computer, and I have

8 another slide up here in a minute. A flat screen and

9 the chemistry is right there.

10 Here is a table and I won't go any further

11 than that. This is the viewpoint from my point of view

12 and we have control of these. We can handle 75

13 students of all kinds, including two wheelchairs and

14 still make the videotapes in this live presentation.

15 On some occasion we have two-way television feed and

16 other live two-way TV. We haven't used that with

17 these. Voice recognition, I mean. The Elmo is one

18 like I am using up here and is very useful. The one I

19 have, I have the movable camera so I take it off an

20 angle and show a chemical demonstration.

21 This is put in at my request so I can control

22 the production. More than one camera, and what we are  
23 trying to do is to help the professor see what the  
24 students see. This is well received, well accepted,  
25 da, da. Most of the materials in the lecture part of  
6

1 it are reusable.

2 This is the educational part and why I got  
3 into it. A student or anyone like yourself here is  
4 listening. So you have a listening range. You could  
5 be taking notes, writing notes and make sure it makes  
6 sense. So you try to comprehend. I don't get it.

7 Someone may have a question. Students have all those  
8 rates going on. They compete.

9 After I may get people asking questions.

10 After we think about what the student is doing, what's  
11 the professor doing up front? He is learning to talk,  
12 makes sure he is on schedule, what's coming next. He  
13 is alert to anything to might -- now, if you have  
14 questions, that's what I want. The idea is, can you  
15 predict if everything is going okay? There are rates  
16 on the professor. The rate the professor talks, thinks  
17 and looks at may be different than the student rate.  
18 So we have a conflict within the educational process

19 that I would suggest that videotape or CDs and those  
20 same devices, get rid of some of these rates.  
21 The learning process now -- sorry. Videotape  
22 is less intimidating. How many of you took chemistry?  
23 How many of you like chemistry? A lot of people,  
24 including my wife, were intimidated. The professor was  
25 unapproachable. By golly! My students can turn me off

7

1 if they have the videotape.

2 >>AUDIENCE: If you turn them off.

3 >>ROBERT PAINE: That's possible. Sometimes

4 my voice can be a little boring, put them to sleep, so

5 forth. When we started we had a class in this setting

6 of 25 students. Four of them were hearing

7 handicapped. Back -- he was over here, and VCR -- not

8 the VCR but the TV set was over here and I was in the

9 middle and students were trying to find out the best

10 place to put the interpreters and then they had to look

11 at four different places in the lecture.

12 As soon as we went live on a single screen

13 and had picture and words, they had one place to look.

14 The attention of the hearing impaired jumped

15 overnight. The apathy left. We still had the



16 interpreters there but now they were supporting rather  
17 than being unsupported.

18 I put this on because you might want to know  
19 how we handle exams. We register proctors around the  
20 world and we send exams to proctors, and I do all the  
21 grading. How do you find a proctor? A librarian, a  
22 nun, true. A local pastor. If you are working  
23 industrially, you have to be supervised to act as your  
24 proctor. I am pleased to tell you, we had a success  
25 rate for using supervisors at the work place at 100

8

1 percent. We have never been turned down for a  
2 supervisor to help us in this cause. It is exciting to  
3 know that people are supporting what we are trying to  
4 do. The student can watch the tape with the captioning  
5 any time, any place, and I suggest 10 o'clock at night,  
6 take the VCR in the bathroom. Nobody will bother you.  
7 That doesn't usually happen. But we have single  
8 parents, people all over the world who may be hearing  
9 impaired and have some form of captioning in some form  
10 or another.

11 This is about what our distribution is. A  
12 third within the county, a third outside the county

13 that are using the distance learning technique, and  
14 another third on campus. We have now made the system  
15 available to campus learners as well. We will have  
16 exams twice a week during the summer, and they could  
17 still use the tape in their dorm room or wherever they  
18 wish to, or some of them are working. Some co-op  
19 programs are working. The synergism I described and  
20 found that if you came, it was useful to everyone.  
21 Some other abbreviations, ASR, VRT, which is  
22 the one I like to use. I already talked about this  
23 one.

24 We embarked four or five years ago trying to  
25 use a variety of existing software to help work on  
9

1 putting live words up instantaneously, my word, my  
2 goal. My goal eventually is to have an interactive  
3 class so that the system will recognize the other  
4 voices in the room. And they could ask questions and  
5 it will appear captioned. We are not there yet, but  
6 what I don't for that is if a student asks a question,  
7 I will repeat the question and then put the answer  
8 afterward. When we get to the demo tape, it is  
9 unedited and we will have some fun with that. Some of

10 the other things you heard about, C-Print started here,  
11 shadowing. When a person talks into a hush mask, it  
12 supplies the punctuation, period, comma, things like  
13 that. We are pretty good at English but not good  
14 enough yet to use it for teaching English. But we have  
15 made some big strides. The shadowing is the person  
16 that uses the Stenomask is there to take notes.  
17 This system, as I will show you in sketch  
18 form in a moment, is software independent. We are  
19 using Dragon Nationally Speaking because it has the  
20 largest vocabulary. We have tried others, including  
21 via voice without much success. We had some silly  
22 successes and we are going to look at anything new that  
23 comes on the market. One of the things we ran into,  
24 which we expected, but it was much more prevalent, we  
25 had placed Mikes around the room, two up front, one on  
10

1 me, one at the ceiling, and all of that made the  
2 software psychotic. The ambient noise is what it was  
3 trying to work with and we cut out all the extraneous  
4 ones except the one for the speaker and that seemed to  
5 restore his personality.

6 The voice independent I just talked to you

7 about. One of the things that we have learned which I  
8 think you may want to get into if you are going to get  
9 into this science yourselves, the software recognizes  
10 the difference between my morning voice and my  
11 afternoon voice. Fatigue, the way I talk, I don't dare  
12 giggle because it tries to print that out. Some of the  
13 things it does print makes me giggle anyway. It knows  
14 my difference between my voice standing up and sitting  
15 down. A lot of the training sitting down is way  
16 normally speak standing up and we have to revise that.  
17 The software knows the difference between a  
18 boy's voice and a girl's voice, and that will come with  
19 training. On the microphones, the head sets that are  
20 used with the radio announcing or if you fly an  
21 airplane, female voice comes across much better. With  
22 Dragon we found that 3 out of 4 men turned out to be  
23 better than women, but I think it is training of the  
24 software. It was a surprise as well.

25 We are at the point where we have

11

1 successfully reached no prior art out there unless it  
2 is manufactured confidential, and it is our patent  
3 application that was accepted March 6 of this year. So

4 we are getting ready to release it. If you took the  
5 hand out, that information is on there. What we would  
6 like to show you very quickly is a chart drawing of how  
7 we did this and then show you the unedited practice  
8 tape that I brought.

9 I was trying to decide whether it was better  
10 in this mode or another one. Thank you. That's a good  
11 choice. We will bring it down and do it like this. I  
12 have another drawing where I put my words on it.  
13 The video office is where I split the screen,  
14 digitizing the audio and video simultaneously and then  
15 bring it through and split our deliver mode over here  
16 and audio and macro that we created within this  
17 computer and goes back into the captioning encoder and  
18 matches the screen so it comes out on a video monitor.  
19 While we are captioning the tape in the classroom for  
20 the students, they will see right across the bottom,  
21 the captioning. "The video storage device outputs  
22 split audio and the NTSD video feed and it goes to the  
23 right. The computer seeks software on the left and  
24 transcribes the audio and text and outputs are entered  
25 into the encoded captioned data. The recorder hardware

1 back to the right now accepts captioned data and  
2 marries it to the video feed."  
3 We had a split coming uncaptioned and we  
4 could go back and caption it if we weren't happy with  
5 it.

6 And any questions? I'm about to show a short  
7 videotape. In its glory we have not edited it and we  
8 are about halfway through the first trial and the  
9 second trial is what I want to focus on.

10 Thank you.

11 Last quarter I had one hearing impaired  
12 student in a class of 65, and that hearing impaired  
13 student was number one in the class. I'd like to give  
14 the videotapes credit for it but it might have been  
15 just part of it. Let us examine each of these  
16 categories.

17 (Captioned tape shown).

18 Try to watch the caption and listen to my  
19 voice. We were experimenting.

20 On the pause, I want you to look what's on  
21 the screen. This is what we call a two line close up  
22 and I will talk about that in a minute.

23 My tongue got in the way of CV and came out

24 as C. So I'm learning how to speak.

25 You can stop that. Now, whatever tapes we

13

1 send out, I never correct my mistakes. If I make

2 mistakes, it has two values: The students will often

3 see the mistake and wait to see if I corrected it. The

4 second value is, never make a mistake. If the

5 professor makes the mistake, okay and they will feel

6 better about it. We started here with a 3 line push

7 up, 3 lines instead of two you saw. It was cumbersome

8 and if I stopped talking the 3 line push up will stop.

9 Once I start talking it will pick up where I left even

10 though I was on the next page. I am very delighted to

11 say two young men, students here working with me on the

12 project, and I have to give them credit for it, and

13 they're on the patent, and I said, look, remember how

14 your printer works and you send something from the

15 computer and goes into memory and you do something

16 else, and why can't we use that technology to make this

17 thing work better so that it will printout. Okay. I

18 give them the idea. I come back in a week. How did

19 you make out? He said I did it in an hour. I did it

20 in an hour. It is delightful to work with those kind

21 of young folks. But what he did was, he went from a 3  
22 line push up to a two line push up, and built in a  
23 pause. So if I paused five seconds, the system will  
24 dump everything and we picked up 60 percent punctuation  
25 doing that because that gives periods, gave us the  
14

1 start of the new sentence, and things like that. And  
2 we did another trick in here that we think we can use  
3 but we are not sure and we hope we will use it if we  
4 have more than one speaker. If I say "green" the  
5 computer prints the words on the screen in green. If I  
6 say "red" it will come out in red. We didn't expect  
7 this. This is an artifact but now we are trying to see  
8 if we can use color coding of different speakers to  
9 carry this on.

10 Can I come back up with the slides?

11 >>AUDIENCE: Go ahead.

12 >>ROBERT PAINE: It says "slide projector."

13 That is the overhead? I want to get the credits up  
14 here before I got out of here. We worked with a grant  
15 from the Dodge Memorial Endowment Fund at RIT, and the  
16 Camielle and Henry Dreyfus Special Grant Program in the  
17 Chemical Sciences. We have 4 and-a-half minutes for



18 questions.

19 >>AUDIENCE: Please use the microphone in the  
20 back of the room. Not everybody at once.

21 >>ROBERT PAINE: Yes.

22 >>AUDIENCE: I think this goes back two  
23 years, this last comment. How do you deal with class  
24 participation and how successful or how accurate is the  
25 voice recognition with people who haven't been

15

1 trained?

2 >>ROBERT PAINE: At the moment when I have a  
3 class of, let's say, 35, there is quite a bit of  
4 ambient noise. So the only Mike that's the one I have  
5 up front. It does not pick up a voice from the  
6 audience well. So what I do in this case, and I will  
7 come back to Mikes in a minute, but I will repeat the  
8 question. So the question prints out my voice. I ask  
9 the questioner did I get it right, and then I go ahead  
10 and answer it. The reason I ask, did I get it right, I  
11 don't want to printout something they didn't intend.  
12 It seems to be, in looking at microphones, it messes my  
13 hair up, but the mike comes with a head set over like  
14 this, and it is best, I am told. That's good for up

15 front. What we are trying is this sort of thing where  
16 we would have in the room I showed you, four different  
17 Mikes that have them in the aisles angled inward. So  
18 it is awkward for a student to get out there if they  
19 could talk toward the microphone. That's our current  
20 experimentation. We are not happy with it but it is  
21 worth a try. What will happen, I'm afraid, is that you  
22 know how we are. There was a good example. What  
23 happens if a student coughs while another student is  
24 trying to ask a question? I have to tell you a couple  
25 of funny things. You saw it up here with chlorine.

16

1 That dumb software believed I had a girlfriend named  
2 Laurie and every time I said chlorine it printed out  
3 Laurie. You know what homonyms are, and for and four,  
4 and if you have a list of things, I said first, second,  
5 third and hope we don't get into fourth. I got into a  
6 funny one, technical word I think you know, zinc, how  
7 do you spell it? Sink, without fail, and we couldn't  
8 convince it it could not do that. So we are working on  
9 that. I have to get my teeth fixed. The other thing  
10 you saw up there, which we don't understand but is man  
11 made, sodium came out NA most of the time. We have

12 gone through 118 elements, and there were 92 when I  
13 started. There were 118 elements and it prints out the  
14 symbols for half of them and the other half it prints  
15 the whole word out. So I think when we made it  
16 psychotic, it is still half and half. We are still  
17 having fun with it because I cannot look at the  
18 captioning while I am practicing because if I look at  
19 the captioning while I'm practicing, it doesn't  
20 printout what I said, then I start to giggle and I lose  
21 everything. The thing I would like to leave on, if any  
22 of you are in this, I would love to talk with you, but  
23 we are now starting a program of accuracy evaluation.  
24 If you read the subject, it forces you to slow down  
25 faster than you thought normal. I think I am around  
17

1 110 words a minute. Reading will give you much better  
2 accuracy than if I am talking to you. I like to get  
3 the students involved. I like to say, what's next when  
4 it is up. The computer doesn't like it. It doesn't  
5 understand when it is reflected and that puts it back  
6 in the boring monotone. What do I do to keep the  
7 students away? I put a joke in or things like that and  
8 the computer doesn't have any trouble there. The

9 software can handle it except for the picture request.

10 Anything else? Thank you very much.

11 (applause).

12 PAT BILLIES: Thank you. I might enjoy

13 chemistry together if I had a teacher with such a great

14 dry sense of humor. Thank you. I remind you to fill

15 out your evaluations. This is session number T230B.

16 Of course, you can fill out the green hard copy or fill

17 them out on line. But we do want your feedback.

18 Secondly, in your green bag is a yellow

19 recommendation form. The recommendation meeting is

20 tomorrow afternoon. You are invited to join us in the

21 theatre at 1:30. The committee for recommendations,

22 though, is seeking your input before that meeting. So

23 if you could take some time today to take a look at

24 this and see if you can give us some of your ideas,

25 that would be terrific. You can leave them on the back

18

1 table in this room or on the registration table down by

2 theater.

3 >>AUDIENCE: Is that on line, too? There is

4 an overall evaluation on line.

5 PAT BILLIES: I don't know that, Susan.

6 Thank you for coming and enjoy the rest of the day.

7 (Session concluded at 3:15).

8 .