

**IBM Accessibility Facebook Expert Hour Transcript - 12/08/2011**  
**Automated Captioning for Lectures in Higher Education**  
**With: the Liberated Learning Consortium**

**IBM Accessibility:** Hello everyone. Welcome to the December IBM Accessibility Expert Hour, our last session of 2011, and thank you for joining us. Our topic today is Automated Captioning for Lectures in Higher Education.

[2 people](#) like this.

**Ccac Captioning:** automated? Better be good :-)

**IBM Accessibility:** Joining us will be experts from Liberated Learning Consortium:

\* Ginny Perelson, Director of Ross Center for Disability Services, University of Massachusetts, Boston

\* Keith Bain, International Manager, Liberated Learning Consortium; Adjunct Prof. Department of Finance, Information systems, Management Science, Sobey School of Business

\* Dr Mike Wald, Senior Lecturer, Web and Internet Science Research Group, ECS, University of Southampton

[2 people](#) like this.

**Anil Joshi:** Hi, Let me congratulate Liberated Learning to joining this forum and help in taking this important point forward.

**Keith Bain:** Hi Anil - thanks for your encouragement!! We're happy to be here.

[1 person](#) likes this.

**IBM Accessibility:** And joining us from the IBM Human Ability and Accessibility Center:

\* Peter Fay, Advocacy and Outreach Program Manager

**IBM Accessibility:** If this is your first Facebook Expert Hour, here is how it works: To participate in the real-time Q&A session, "Like" IBM Accessibility. Then to ask a question, type it into the box on your screen that says, "Write something..." and press Enter. Our experts will answer your question as a Comment, and it will show up underneath your question.

**IBM Accessibility:** If you use a screen reader, we suggest using the Facebook HTML-only version: <http://m.facebook.com/home.php>

**IBM Accessibility:** Please remember to refresh occasionally -- sometimes Facebook has trouble keeping up with the pace of our Expert Hours.

**IBM Accessibility:** Okay, let's open the wall for questions. :-)

**T.m. Weissenberger:** Will there be a transcript for this session?

[Anil Joshi](#) likes this.

**IBM Accessibility:** Yes, there will be one. It will be on <http://www-03.ibm.com/able/education/index.html#Education> in a few days. :-)

**IBM Accessibility:** Could you tell us what speech recognition-based tools are available for captioning/transcription?

**IBM Accessibility Facebook Expert Hour Transcript - 12/08/2011**  
**Automated Captioning for Lectures in Higher Education**  
**With: the Liberated Learning Consortium**

[Keith Bain](#): Liberated Learning recently designed a new system - taking lessons learned from ViaScribe and caption editing system.

[Cindy Poore-Pariseau](#): Keith, do you know if the speaker has to train the software to recognize their speech patterns for it work effectively (similar to dragon technology)

[Ccac Captioning](#): anyone here today using Liberated Learning or ViaScribe now?

[Keith Bain](#): Hi Cindy - we have both speaker independent and speaker dependent systems. Speaker independent systems do not require training and automatically adapt to the speaker. However, the speaker adaptation is not saved.

[Peter Fay](#): Check out IBM's Media Captioner and Editor on our website: <http://www-03.ibm.com/able/projects/captioneditor.html>



[Keith Bain](#): CCac - I am currently using Dragon NaturallySpeaking for this chat - it is a speaker dependent system, meaning I've trained it extensively

[Ccac Captioning](#): Dragon needs your voice training - how does this apply to classroom education?

[Ccac Captioning](#): Extensively is the key! We are most interested in how these systems will evolve for real time group conversations in education and in many other venues

[Keith Bain](#): We found that instructors were very reluctant to do much voice training in our early research - yet accuracy and training are highly correlated. We subsequently tried to develop training techniques that were transparent to the user including something we called batch enrollment a process that used actual lecture data to create a profile for the speaker. Some of our new systems are fully speaker independent - meaning they do not require voice training and rather adapt dynamically to the speaker. There are also hybrid systems that start off as speaker independent - no training and then as they are used, the resulting data is fed into the system to create a profile. We think these have the most promise, as generally speaking, speaker dependent systems are currently more accurate.

[1 person](#) likes this.

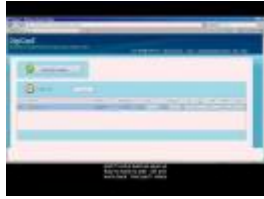
[T.m. Weissenberger](#): What are some good tools for producing synchronous captions for videos? Is there a preferred format?

[Peter Fay](#): IBM Research has a tool called Media Captioner and Editor that produces synchronized captions. It supports a wide variety of file formats.

[Ginny Perelson](#): It depends on what you are meaning by "good tools"

**IBM Accessibility Facebook Expert Hour Transcript - 12/08/2011**  
**Automated Captioning for Lectures in Higher Education**  
**With: the Liberated Learning Consortium**

Peter Fay: You can see a video demo on YouTube:  
[http://www.youtube.com/watch?v=tRQ6\\_q2r3EY](http://www.youtube.com/watch?v=tRQ6_q2r3EY)



Keith Bain: IBM ViaScribe was the world's first specially designed speech recognition captioning system

Keith Bain: IBM caption editing system was another version

Christine Banke: What are the latest stats with regards to the accuracy of the transcription, prior to editing of course?

Keith Bain: Our basic evaluation consists of a test set of about 50,000 words - 10 different speakers from 10 different lectures. Accuracy varies between different speech recognition systems and between different acoustic and language models. The best system we have benchmarked to date is about 85% accurate - This is a speaker independent system, mean it has no training of the speakers in the test set  
[1 person](#) likes this.

Christine Banke: Thank you, Keith.

Peter Fay: With Media Captioner and Editor we see results in 72-85 %, depending on speaker, audio quality etc. (These are results for generic language models)

Keith Bain: we have a new statistical analysis coming out in early January

Christos Kouroupetroglou: Are there any tests with those tools in a classroom environment?

Keith Bain: Hi Christos - there have been many projects and research studies on some of these tools - visit [http://liberatedlearning.com/?page\\_id=565](http://liberatedlearning.com/?page_id=565) for a list of publications

Mike Wald: I am using speech recognition in all my classes

Christos Kouroupetroglou: That's helpful and encouraging...

Larry Goldberg: Hi - where has this been implemented?

Keith Bain: Hi Larry - over 30 universities worldwide - see [http://liberatedlearning.com/?page\\_id=7](http://liberatedlearning.com/?page_id=7) for a list of institutions

Mike Wald: I am using it in Southampton University UK

**IBM Accessibility Facebook Expert Hour Transcript - 12/08/2011**  
**Automated Captioning for Lectures in Higher Education**  
**With: the Liberated Learning Consortium**

[Peter Fay](#): Hi Larry: We are just starting a couple of pilots with Media Captioner and Editor this month...we hope to have some live customers and partners soon.  
[1 person](#) likes this.

[Mike Wald](#): [www.synote.org](http://www.synote.org)

[Larry Goldberg](#): great - looking forward to helping out.

[Christos Kouroupetroglou](#): How possible is it in the future such technologies to be included in a classrooms basic infrastructure... (i.e. interactive whiteboards, smart desks...)?

[Christos Kouroupetroglou](#): Thinking for the 2020 horizon...

[Ginny Perelson](#): We are hopeful that Universal Design will be standard in higher education, which will of course mean that technologies are core to the classroom design as well as the curriculum, etc

[Mike Wald](#): Any class can use it in Southampton University

[Keith Bain](#): Some of our partners have already done this - see  
<http://webcast.trentu.ca/CourseCast/Viewer/Default.aspx?id=260b7c1c-0f7a-412d-a6fe-cb576e49fa7d>

[Keith Bain](#): this is an example of integration of SR with Panopto - the captions tab to see the SR generated text - Mike Wald at Southampton is also integrated this in Synote

[Peter Fay](#): Christos: IBM software group is actively looking at integrating our collaboration tools with interactive whiteboards, video conferencing systems, social networking and more....

[Christos Kouroupetroglou](#): It's nice to know it's implemented in some places... but how can this become widespread?

[Peter Fay](#): Christos: We believe cloud implementation will help, and we're also recruiting business partners to help broaden our reach.

[Ginny Perelson](#): Increase the membership in Liberated Learning consortium, conduct more research to show results, and continue to educate about the possibilities,

[Christos Kouroupetroglou](#): Peter and Ginny: Thanks. Do you think that promoting the value that transcripts add in a lecture (not just for people with hearing impairments) could also help in this mainstreaming. (i.e. they could be automatically uploaded in the university's intranet, etc.)

[1 person](#) likes this.

**IBM Accessibility Facebook Expert Hour Transcript - 12/08/2011**  
**Automated Captioning for Lectures in Higher Education**  
**With: the Liberated Learning Consortium**

[Ginny Perelson](#): Yes exactly-speech recognition is value-added for all learners, those whose native language is not English (in the USA), and the growing percentage of learners who have some sort of hearing impairment

[Peter Fay](#): Christos: Absolutely, agree w/ @Ginny's post

[Christine Banke](#): Where are the goals of LL over the next 2-5 years or so?

[Keith Bain](#): better systems - better accuracy - more usage

[Keith Bain](#): We also need more partners - a critical mass of like-minded organizations will increase innovation  
[1 person](#) likes this.

[Ginny Perelson](#): We would like to conduct specific educational research to determine effectiveness related to outcomes

[Peter Fay](#): From IBM we hope to leverage cloud-based real-time automatic speech recognition to reach a broader audience of students and faculty.

[Fran Hayden](#): Hi all! I'm wondering if you can talk about how you correct the errors in automatic speech recognition captioning.

[Mike Wald](#): It is easier to correct errors after the lecture as trying to correct live introduces a delay

[Mike Wald](#): The time taken depends on how many errors there are

[Fran Hayden](#)” That makes sense - Thanks!

[Keith Bain](#)” depending on accuracy and the editor's familiarity with the content, the amount of editing required varies significantly

[Peter Fay](#): Hi Fran: Our media captioner and editor tool includes an editing environment that allows you to listen to the audio track and edit the caption at the same time.

[Larry Goldberg](#): Are interactive transcripts part of the system? That is, random search via the text transcript, post-presentation?

[Keith Bain](#): yes - we refer to the concept of multimedia transcripts - we have different formats, but essentially they use an XML-based structure to synchronize media.

[Mike Wald](#): yes our transcripts are synchronized with the recording and can be searched and replayed from that point

[Keith Bain](#): This is a flash-based example  
[http://transcribeyourclass.ca/demos/WebSphere\\_Video HTS\\_Demo/SLWeb.html](http://transcribeyourclass.ca/demos/WebSphere_Video HTS_Demo/SLWeb.html) it is

**IBM Accessibility Facebook Expert Hour Transcript - 12/08/2011**  
**Automated Captioning for Lectures in Higher Education**  
**With: the Liberated Learning Consortium**

not fully accessible, but it does allow interactivity. We have other non-flash-based examples that allow you to customize the text, interact, Search for key terms etc.

[Larry Goldberg](#): Can these be integrated into learning management systems like Blackboard and hosted sites like iTunes U?

[Keith Bain](#): we also have special systems such as Synote, MIT lecture browser, NoteFinder, that work with the SR generated transcripts and provide tools that students can use to interact with the multimedia transcripts

[Peter Fay](#): Larry: This is part of our plan...we're currently looking at our analytics portfolio for searching transcripts, linking directly to that portion of the video...and more. Stay tuned!

[1 person](#) likes this.

[Jan Myland](#): We use captions transcribed by "transcribeyourclass" with the Panopto webcast system and the captions are fully searchable.

[4 people](#) like this.

[Larry Goldberg](#): Does Panopto have built-in transcription?

[Mike Wald](#): no it doesn't have it built in but you can upload transcripts

[IBM Accessibility](#): Another question for you, can speech recognition capture classroom discussions?.

[Ccac Captioning](#): what about if all students speak into the mic of their laptop?

[Mike Wald](#): only if everybody else in the room stays completely silent !!

[1 person](#) likes this.

[Christos Kouroupetroglou](#): What about different languages?

[Christos Kouroupetroglou](#): Where is the status with various languages?

[Keith Bain](#): It depends on the system - for example our IBM Hosted Transcription Service Supports U.S. English, Mandarin, and Spanish.

[Peter Fay](#): Currently IBM Media Captioner and Editor currently supports English and Mandarin...and we're looking at adding more in the future.

[Keith Bain](#): We are also building systems that can leverage available languages - building a language from scratch requires a large amount of resources, which means that minority languages are often neglected

[1 person](#) likes this.

**IBM Accessibility Facebook Expert Hour Transcript - 12/08/2011**  
**Automated Captioning for Lectures in Higher Education**  
**With: the Liberated Learning Consortium**

[Mike Wald](#): Speech recognition is available in a range of languages but there are not all the different languages for all different systems as it is very expensive to collect the speech and text and analyze it

[Keith Bain](#): in our Net4Voice project in the EU, our partners used Italian and German systems

[Ccac Captioning](#): available to try this in French for a few months - Net4Voice - will look that up

[Christos Kouroupetroglou](#): As a native speaker of Greek... I wouldn't like to see my language disappearing in the future due to technology that cannot understand it. But as I see, up to now, there is no easy and cheap way to implement speech recognition system for a new language.

[Christos Kouroupetroglou](#): Do you think that this cost will lower in the future? How?

[Mike Wald](#): it would be helpful if Europe funded such developments for all European languages

[1 person](#) likes this.

[Mike Wald](#): if enough people asked for this then it might happen

[Keith Bain](#): In our experience, creating a new language in a particular domain requires about 250 to 300 hours of domain data. Obviously, the more data you have for a particular domain, the better your results will be.

[Christos Kouroupetroglou](#): Could crowd sourcing techniques help in that?

[Larry Goldberg](#): Does that estimate include accurate transcripts for comparison/learning of the new language?

[Mike Wald](#): the companies providing free speech recognition on mobiles are using the data from the speakers to improve the accuracy

[Peter Fay](#): Christos: Our research team in Tokyo is doing some promising work in a crowd sourcing for captioning. cool stuff!

[Christos Kouroupetroglou](#): Peter: Nice to know

[Marnie Hoover](#): Hello, I'm curious, just how effective is speech recognition with various accents?

[Ginny Perelson](#): The accent is one variable-research is planned to determine the critical variables to distinguish the most accurate capture

**IBM Accessibility Facebook Expert Hour Transcript - 12/08/2011**  
**Automated Captioning for Lectures in Higher Education**  
**With: the Liberated Learning Consortium**

[Mike Wald](#): with speaker dependent systems (e.g. dragon) the speaker can train the system and improve recognition but the result will depend on how different it is from the voices the system was trained on

[Mike Wald](#): with speaker independent systems the speaker can't train them but they are designed to work with a wider range of accents

[Keith Bain](#): We are also developing a multilingual language model. It is built using a medley of different speakers with different accents to provide a statistical model that can recognize a wide variety of speakers

[Marnie Hoover](#): okay, I understand .... thank you!

[Keith Bain](#): Some other variables are equally important. Even if someone has an accent, if they speak relatively clearly and record good quality audio, they can still get good results

[IBM Accessibility](#): What does the research demonstrate regarding speech recognition and student achievement in higher education?

[Ginny Perelson](#): The most promising research is in the area of instructional design and cognitive load theory: [insitute.nsta.org/scipack\\_research/research.htm](http://insitute.nsta.org/scipack_research/research.htm).

[Ginny Perelson](#): We need more though

[Keith Bain](#): Our partners at Purdue University recently conducted a study showing that students who received speech recognition generated transcripts did better on exams - paper should be published in early 2012

[Mike Wald](#): my research suggests that it makes learning easier but it is difficult to measure grade improvements

[Larry Goldberg](#): That study will be of huge value in so many ways, even if grade improvement is a difficult measurement.

[Keith Bain](#): Thanks to everyone for your questions - if you have more, contact us anytime at [liberated.learning@SMU.ca](mailto:liberated.learning@SMU.ca)

[IBM Accessibility](#): We are now at the end of our time. What a great session -- thanks for joining us today. :-)

[IBM Accessibility](#): [www-03.ibm.com](http://www-03.ibm.com)

By enabling human capability through innovation, the IBM Human Ability and Accessibility Center strives to help all people maximize their potential, regardless of age or ability.

[IBM Accessibility](#): The session will remain on the Facebook wall, and we are now posting them on the [ibm.com/able](http://ibm.com/able).