

UNIVERSITY OF TEXAS AT DALLAS



JOHN H.L. HANSEN, PH.D.

CRSS: Center for Robust Speech Systems

Erik Jonsson School of Engineering & Computer Science

Department of Electrical Engineering

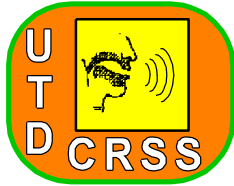
University of Texas at Dallas

Electrical Engineering, EC33 P.O. Box 830688

Richardson, Texas 75083-0688 U.S.A

PHONE: 972-883-2910 FAX: 972-883-2710

EMAIL: John.Hansen@utdallas.edu <http://crss.utdallas.edu/>



August 1, 2013

To: ISCA: International Speech Communications Association

RE: Summary of ISCA Distinguished Lecturer activities for John H.L. Hansen for 2012/13



This memo summarizes the ISCA DL (Distinguished Lecturer) activities for John Hansen during the 2012/13 year. Upon accepting the duties of ISCA DL, Hansen committed to making two trips to countries/regions which have historically seen less speech related activities. The goals were to:

- 1) Encourage students to pursue education and research/job opportunities in speech and language processing,
- 2) Provide direct feedback to individual existing MS/PhD students performing research and to encourage them to continue their work and pursue careers in the speech/language field,
- 3) Engage with local researchers in the field to help encourage and explore options for hosting workshops or conferences in speech (i.e., Interspeech) in the future

In the initial discussions with ISCA, Hansen noted three potential countries to visit which included (i) India, (ii) Turkey, and (iii) Chile/Argentina. These areas are different in their presence in the speech and language processing field. India has a large level of activity with many accomplished researchers, but have yet to host a major international conference in the field of speech processing. Turkey has hosted IEEE ICASSP-2000, and has also proposed to host Interspeech in the past, but their bid was not selected. Turkey also has a number of young speech researchers in the field who are increasing their activity and have the capacity to host a major conference. Chile/Argentina have several key active researchers with good presence in the field, and continued visits and collaboration would help strengthen and support their activity. All three potential locations are active and emerging in terms of ISCA and speech/language processing.

After a series of email communications with potential contacts, two trips were planned for the summer of 2012 to (i) India, and (ii) Argentina. This report summarizes the activities and contact information of the collaborations.

INDIA:

Dates of Travel: June 18 – 29, 2012

During June 18-29, 2012, John Hansen traveled to visit three Universities/Institutes in India. In order to incorporate an extensive hands-on laboratory experience, Dr. Abhijeet Sangwan, a staff researcher with the Center for Robust Speech Systems (CRSS) UTDallas traveled and participated in the laboratory instruction.

It should be noted that ISCA provided generous support for the base airline travel for Hansen to-and-from India. All other travel expenses, including domestic air travel within India for Hansen and Sangwan, international travel for Sangwan from Dallas to India, all hotel and travel costs within India, were supported by CRSS-UTDallas through Hansen's Distinguished University Endowed Chair (this totaled +\$6500; there were no direct charges for any of these schools for local accommodations).

1. Hosts: [CMRIT, Bangalore (INDIA)] CMR Institute of Technology
Prof. H. N. Shankar, Dean – Academics & Research, CMRIT, Bangalore
#132, AECS Layout, IT Park Road
Bangalore - 560 037 INDIA
Tel: +91 (80) 6579 1751 (Direct);
Fax : +91 (80) 2852 4477
Email: hnshankar@cmrit.ac.in



Prof. R. Muralishankar, Dept. of Electronics & Communication Engg., CMRIT, Bangalore
#132, AECS Layout, IT Park Road
Bangalore - 560 037, INDIA.
Tel : +91 (80) 2852 4630 & +91 (80) 2847 6965 x 555
Fax : +91 (80) 2852 4477
Email: muralishankar@cmrit.ac.in
Web : <http://it.cmr.ac.in/>



2. Workshop Abstract:

Monitoring Human Behavior: “Who Said What When and How?”



John H.L. Hansen, (ISCA Distinguished Lecturer)

Abhijeet Sangwan, [workshop support]

CRSS: Center for Robust Speech Systems; Dept. of Electrical Engineering
Erik Jonsson School of Engineering & Computer Science
University of Texas at Dallas, U.S.A.



<http://crss.utdallas.edu/>

John.Hansen@utdallas.edu; abhijeet.sangwan@utdallas.edu

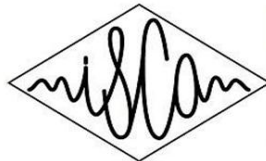
Seminar: Hansen

There is significant interest in the development of effective human-machine interactive systems for a wider range of personal services. Speech and Speaker Processing/Recognition research has advanced significantly in recent years, but performance in real environments remains a major challenge. In this talk, we consider several recent research efforts in the field of robust speech and speaker recognition. The talk is partitioned into three phases: (i) speech analysis/modeling, (ii) monitoring speaker characteristics, and (iii) applications – Prof-Life-Log and UTDrive In-Vehicle systems. For Analysis & Modeling, challenges on the ability to

identify those traits which distinguish an individual subject will be considered. Next, speaker monitoring including speaker stress/emotion, vocal effort (e.g., whisper, soft, neutral, loud, shout), Lombard effect (speech produced in noise), language/dialect, will be considered. Finally, applications of how to Monitor Human Behavior will be explored. This will consider: (a) speech recognition and keyword spotting based on Phonological Features, (b) Speaker ID & Dialect/Language ID, and (c) Spoken Document Retrieval (SDR) based on SpeechFind® (<http://SpeechFind.utdallas.edu>).

Small Vocabulary Speech Recognition (Hands On Workshop): Sangwan

Modern automatic speech recognition (ASR) systems are complex machines. While research and development in ASR requires extensive expertise, building practical applications with existing state of the art can be relatively simpler and a rewarding experience. In this workshop, we will build a small vocabulary speech recognition engine to explore components of an ASR engine, namely, (i) acoustic, (ii) language and (iii) pronunciation models. We will also point out a number of publicly available resources that can assist you in building practical ASR systems. The workshop can be immensely helpful for students who are working towards their final year projects and corporations that may be looking towards integrating speech recognition technology into their products and services. The workshop will use Sphinx-4, a state-of-the-art Java-based open source speaker independent continuous speech recognizer available under BSD style license. It provides a simple/powerful API (application programming interface) that allows you to rapidly build applications.



Dr. John H. L. Hansen
Univ. of Texas, Dallas, USA

ONE-DAY WORKSHOP
SPEECH PROCESSING & APPLICATIONS

Thursday, June 21, 2012

Organized by

Dept. of Electrical and Electronics Engg.
Dept. of Telecommunication Engg.

CMR Institute of Technology

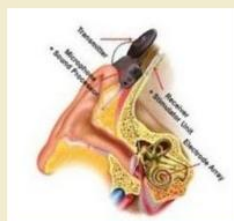
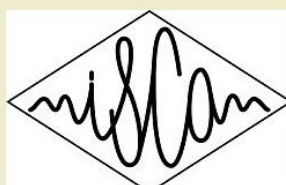


Dr. Abhijeet Sangwan
Univ. of Texas, Dallas, USA

Co-sponsor: International Speech Communication Association, France



<https://sites.google.com/a/cmrit.ac.in/muralishankar/events/speech-signal-processing-workshop-june-21-2012-cmrit>
One-day Workshop on Speech Processing and Applications



Thursday, June 21 2012

Organized by
Dept. of Electronics &
Communication Engg.,
Dept. of Telecommunication
Engg.
CMR Institute of Technology
132, AECS Layout, IT Park Road, Bangalore 560
037, INDIA



[DOWNLOAD printable brochure here](#)

Chairs

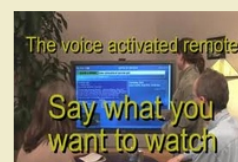
Prof. R. Muralishankar, Dept. of Electronics & Communication Engg., CMRIT, Bangalore
Prof. H. N. Shankar, Dean - Academics & Research, CMRIT, Bangalore

Co-Chairs

Prof. Indumathi Govind Rao, Head, Dept. of Electronics & Communication Engg., CMRIT, Bangalore
Prof. Fathima Jabeen, Head, Dept. of Telecommunication Engg., CMRIT, Bangalore

Sponsors

CMR Institute of Technology, Bangalore, India
International Speech Communication Association,
Baixas, France



Technical Sponsor

IEEE Signal Processing Society, Bangalore Chapter

Organization Committee

Prof. Mamtha Agiwal, Dept. of Electrical & Electronics Engg., CMRIT, Bangalore
Ms. A. Pichu, Dept. of Telecommunication Engg., CMRIT, Bangalore
Ms. Pratiksha Naik, Dept. of Electronics & Communication Engg., CMRIT, Bangalore
Ms. Pragya Pallavi, Dept. of Electronics & Communication Engg., CMRIT, Bangalore
Mr. Abhishek Javali, Dept. of Telecommunication Engg., CMRIT, Bangalore
Mr. G. B. Umesh, Dept. of Telecommunication Engg., CMRIT, Bangalore

Theme and Objectives

Speech and Speaker Processing/Recognition research has advanced significantly in recent years, but performance in real environments remains a major challenge. Effective human-machine interactive systems demand robust speech and speaker recognition. Applications to monitor Human Behavior are central to personalized services. The issues, challenges and applications will be brought out here.

Special Feature

- Modern Automatic Speech Recognition (ASR) systems for building practical applications will be demonstrated.
- Excellent opportunity for students: Prof. Hansen and Dr. Sangwan will be looking for good students to work with their team at Univ. of Texas, Dallas, USA.

Who should attend?

- For faculty, the workshop can be immensely helpful in gaining knowledge and in looking at vistas for research in Speech Processing and Applications.
- For students working towards their projects, this opens up new challenges.
- Researchers in labs and product developers in industries can gain insights into integrating speech recognition technology into their products and services.

Registration: Now OPEN! Register [here](#)

Last date for registration: June 19, 2012

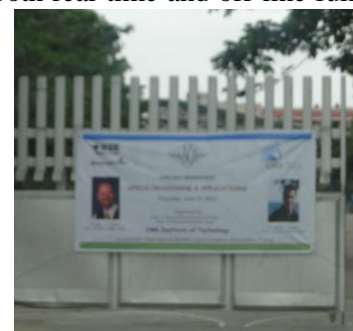
Limited spot registrations on June 21, 2012, may be accommodated.



3. Format:

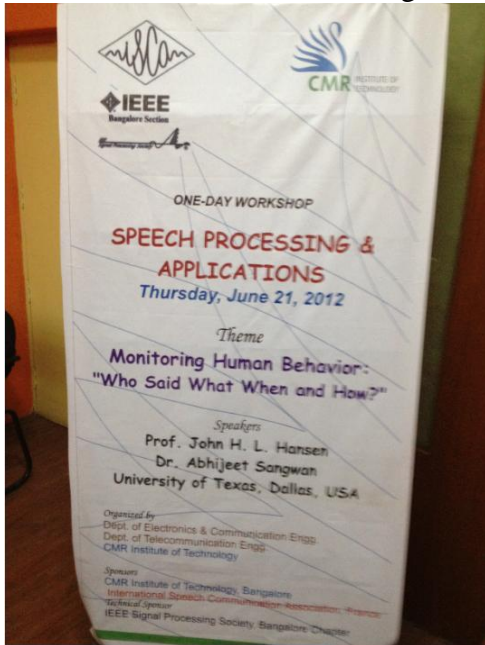
The one day workshop was organized with two lecture sessions (each about 1hr each), with a break between for lunch. Followed by a hands-on laboratory session (also about 1 hr) where both real-time and off-line full speech recognition systems were distributed to all participants. All slides were shared with the workshop participants. After the hands-on laboratory, all participants assembled in the main lecture hall for a Question & Answer session, which included feedback comments from the attending students.

The workshop was advertised in the local paper and a large poster produced which appeared at the entrance to the University (shown here).



4. Attendance: around 70 people participated in the short course.

This pictures show: (i) the Banner in the main lecture hall where the workshop took place; (ii) a student after completing the workshop {all attending participants received a certificate of completion which was signed by the Hosts and Hansen as DL instructor}; (iii) picture of the hands-on laboratory where Dr. Sangwan and Hansen presented fundamentals in building speech recognition systems, along with the distribution of a working JAVA system to each participant.



1. Hosts: IISc, Bangalore (INDIA)
Speech & Audio Group (Prof. T. V. Sreenivas and Harshavardhan.S).
CEDT (Prof. Gopakumar and Prof. Joy Kuri)



Prof. K.Gopakumar
Chairman, Department of Electronic
Systems Engineering-DESE
(Formerly CEDT)
Indian Institute of Science

Bangalore-560012, INDIA

<http://www.dese.iisc.ernet.in/ese/index.php/facultys/gopakumar-k>

Email: kgopa@cedt.iisc.ernet.in

Ph: +91 80 2293 3089



Prof. K V S Hari
<http://www.ece.iisc.ernet.in/~hari/>
Chair, IEEE SP Society Bangalore Chapter
Vice-Chair, IEEE Bangalore Section
<http://www.ece.iisc.ernet.in/~hari/>

Email: hari@ece.iisc.ernet.in

Department of ECE, Indian Institute of Science, Bangalore 560012, India

Ph: +91 80 22932745, Fax: +91 80 23600563



Prof. Joy Kuri [kuri@cedt.iisc.ernet.in]
Department of Electronic Systems Engineering (DESE)
Indian Institute of Science (IISc)
Bangalore 560012

<http://www.cedt.iisc.ernet.in/people/kuri/>

Email : kuri@cedt.iisc.ernet.in

Phone : +91 80 23600810 Ext 228 / +91 80 22933091

Fax : +91 80 22932290

2. Format:

Hansen presented a seminar which lasted about 1hour. Because of the structure and specific student attendance, we did not have a hands-on workshop with the entire group here. Instead, after the morning presentation, and lunch, we moved into a lecture room for round-table extended Question-Answer format. Individual students gave short presentations/overviews of their current research activities, and Hansen provided feedback regarding impact and technical trade-offs.

Attendance: around 65 people participated in the seminar; with about 20 in the afternoon Question-and-Answer session.

1. Hosts: IIT – Madras (Chennai, INDIA)



Prof. S. Umesh; IIT, Chennai (INDIA)
Room No. 303-C, ESB Bldg.
Department of Electrical Engineering
Indian Institute of Technology - Madras
Chennai 600 036

<http://www.ee.iitm.ac.in/~umeshs/>

E-mail: umeshs@iitm.ac.in

Tel: +91 44 2257-4461

Fax: +91 44 2257-4402



Prof. Hema A. Murthy; IIT, Chennai (INDIA)
Department of Computer Science and Engineering
Indian Institute of Technology Madras
Chennai - 600 036

Ph: +91-44-2257 4364

Fax: +91-44-2257 4352

E-mail: hema@cse.iitm.ac.in

<http://www.cse.iitm.ac.in/~hema/>



Prof. C.S. Ramalingam; IIT, Chennai (INDIA)
Department of Electrical Engineering
Indian Institute of Technology Madras
Chennai - 600 036

Ph: +91-44-2257 4434

Fax: +91-44-2257 4402


Email: csr@iitm.ac.in

Format:

Hansen gave a one-hour presentation on “Monitoring Human Behavior: Who said What When and How?”. After this session, the group moved to a computer lab room, where Dr. Sangwan presented an overview of the JAVA based software for speech recognition. Fully functioning speech recognition platform was distributed to all participants, followed by Question and Answer phase.


“Monitoring Human Behavior”
Who said What When and How?

John H.L. Hansen
Abhijeet Sangwan




UTD
CRSS

Center for Robust Speech Systems (CRSS)
Erik Jonsson School of Engineering & Computer Science
The University of Texas at Dallas
Richardson, Texas 75083-0688, U.S.A.
<http://crss.utdallas.edu>
John.Hansen@utdallas.edu Abhijeet.Sangwan@utdallas.edu



ISCA Distinguished Speaker Series
June 21, 2012



Email: John.Hansen@utdallas.edu CRSS: Monitoring Human Behavior- Who Said What, When and How: Slide 1 SLIDES © by John H.L. Hansen, 2012

Attendance: around 35 people participated in the seminar along with the Hands-On Laboratory experience and Question-and-Answer session.

While in Chennai, India, Hansen and Sangwan visited an historical region to see the local culture (pictures shown below of stone carvings).



ARGENTINA:

Dates of Travel: July 20 – 25, 2012

1. Host: Prof. Agustín Gravano [University of Buenos Aires, (Buenos Aires, Argentina)]

Email: gravano@dc.uba.ar

Mailing address: Departamento de Computación, FCEyN, Universidad de Buenos Aires. Pabellón I, Ciudad Universitaria. (C1428EGA) Buenos Aires, ARGENTINA

Phone: (+54) 11 4576-3390 - ext. 717

<https://www-1.dc.uba.ar/events/eci/2012/fotos/agustin-gravano-john-hanse-y-sergio-yovine/view>
<http://habla.dc.uba.ar/gravano/index.php?lang=eng>

Hansen traveled to University of Buenos Aires – ECI as an ISCA DL to participate in their Summer School program. ECI has been running a 1 week summer school in July (actually their winter) for approximately 17 years. Their program is designed to bring past alumni from their institution, as well as anyone interested in an intensive exposure to specific topics in the computer science/computer engineering/electrical engineering domains on to campus, to consider new and emerging areas. This is not a “traditional lecture” style talk, but instead consists of a series of intensive 3 hour lectures spread over 5 days plus at least 1 hands-on lab experience and a final exam, all to be administered by the visiting speaker. While this is significantly more than a traditional Distinguished Lecturer visit, it does off the opportunity to have much more in-depth coverage of materials. The preparation work is significantly more challenging, especially since a hands-on laboratory experience is needed, as well as a final exam which must be administered.

Below is the course abstract drafted by Hansen for the ECI Summer school. Because the audience is broad, the topics covered a wide range of fundamentals on speech diarization, including speech science, speech modeling, signal processing, speech recognition, speaker recognition, and applications.

2. Summer Course Abstract:

Advanced Speech Processing Techniques: “Who Said What When and How?”

John H.L. Hansen
(ISCA Distinguished Lecturer)



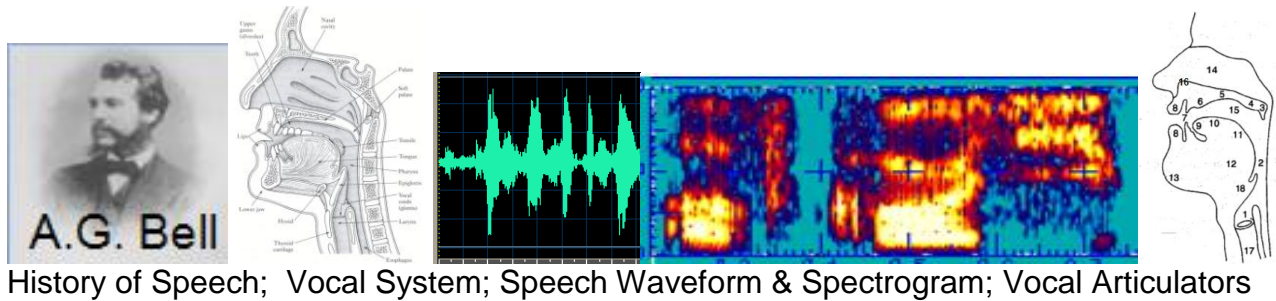
CRSS: Center for Robust Speech Systems; Dept. of Electrical Engineering
Erik Jonsson School of Engineering & Computer Science
University of Texas at Dallas, U.S.A.
<http://crss.utdallas.edu/> John.Hansen@utdallas.edu;



Course Summary

There is significant interest in the development of effective human-machine interactive systems for a wider range of personal services. Speech and Speaker Processing/Recognition research has advanced significantly in recent years, but performance in real environments remains a major challenge. In this talk, we consider several recent research efforts in the field of robust speech and speaker recognition. The talk is partitioned into three phases: (i) speech analysis/modeling, (ii) monitoring speaker characteristics, and (iii) applications – Prof-Life-Log and UTDrive In-Vehicle systems. For Analysis & Modeling, challenges on the ability to identify those traits which distinguish an individual subject will be considered. Next, speaker monitoring including speaker stress/emotion, vocal effort (e.g., whisper, soft, neutral, loud, shout), Lombard effect (speech produced in noise), language/dialect, will be considered. Finally, applications of how to Monitor

Human Behavior will be explored. This will consider: (a) speech recognition and keyword spotting based on Phonological Features, (b) Speaker ID & Dialect/Language ID, and (c) Spoken Document Retrieval (SDR) based on SpeechFind® (<http://SpeechFind.utdallas.edu>).

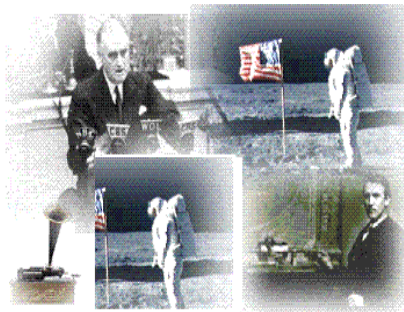


History of Speech; Vocal System; Speech Waveform & Spectrogram; Vocal Articulators

Topical Course Outline:

A brief overview of the topics covered in this course are listed below.

- (1) Speech Science: phonetics, linguistics, acoustic phonetics
- (2) Models for Speech Production
- (3) Features for Speech Recognition (MFCCs, PLP, RASTA, PMVDR, etc.)
- (4) Speech Recognition: Hidden Markov modeling (HMM), Lexicons, Language Modeling
- (5) Speaker Recognition – Features, Models, Performance Evaluation (EER, DET Curves)
- (6) Speech & Speaker Recognition Applications: Spoken Document Retrieval, Hands-Free Voice Capture for ASR, Speaker Tracking

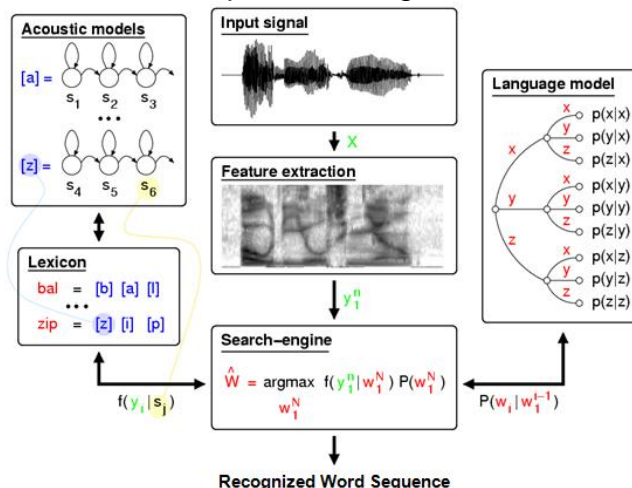


Searching Audio Archives



Vocal System Articulation

Overview of Speech Recognition Process



Small Vocabulary Speech Recognition (*Hands On Workshop*)

Modern automatic speech recognition (ASR) systems are complex machines. While research and development in ASR requires extensive expertise, building practical applications with existing state of the art tools can be relatively simpler and a rewarding experience. In this workshop, we will build a small vocabulary speech recognition engine to explore components of an ASR engine, namely, (i) acoustic, (ii) language and (iii) pronunciation models. We will also point out a number of publicly available resources that can assist you in building practical ASR systems. The workshop can be immensely helpful for students who are working towards their final year projects and companies that may be looking towards integrating speech recognition technology into their products and services. The workshop will use Sphinx-4, a state-of-the-art Java-based open source speaker independent continuous speech recognizer available under BSD style license. It provides a simple/powerful API (application programming interface) that allows you to rapidly build applications.

3. Course Logistics:



Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires



Short-Course Format: Hansen prepared three 3-hour lectures, which included:

- **PART A (3hr):** Speech Fundamentals, Language, Hearing Science; Speech Analysis & Speech Modeling
- **PART B (3hr):** Speech Recognition: (i) Acoustic Models, Lexicon, Language Models; (ii) Speaker Recognition: SID, Speaker Verification, In-Set/Out-of-Set ID; (iii) Language ID & Dialect ID (LID & DID); (iv) Acoustic Model Adaptation & Normalization
- **PART C (3hr):** Applications in Speech & Speaker Systems: (i) Monitoring Speaker Characteristics (Speaker Stress & Emotion; Vocal Effort/Whisper/Lombard Effect; Speaker Height; Driver Distraction/Stress & In-Vehicle Technology); (ii) SpeechFind® [<http://SpeechFind.utdallas.edu>]; (iii) ASR / Keyword Recognition / Prof-Life-Log
- **HANDS-ON Workshop (3hr):** complete JAVA based speech recognition system focused on USA “zip-code” number strings – code given to all attendees, installed on their own machines and tested during the lab experience.

4. Attendance: around 50 people participated in the short course (3x 3hr lectures plus 3hr lab; 4 sessions total)

This is the university website which summarizes the activities from this Summer Course.

<https://www-1.dc.uba.ar/events/eci/2012/fotos/agustin-gravano-john-hanse-y-sergio-yovine/view>

The screenshot shows the website for the Department of Computing at the University of Buenos Aires. The header includes the department name, faculty name, and navigation links like 'Mapa del Sitio', 'Accesibilidad', and 'Contacto'. A search bar is present. Below the header is a navigation menu with links for 'Inicio', 'Futuros Estudiantes', 'Estudiantes actuales', 'Docentes', and 'Graduados'. The main banner features the text 'Exactas EL lugar donde estudiar computación'. A breadcrumb trail indicates the current page is 'Inicio > Eventos > ECI > ECI 2012 > Fotos > Agustin Gravano, John Hanse y Sergio Yovine'. The main content area is divided into three columns: a left sidebar with a table of contents for the ECI 2012 event, a central image of three men standing in front of a registration banner, and a right sidebar with 'Últimas noticias' (Latest news) listing several articles with dates.

ECI 2012	
Inicio	
¿Qué es la ECI?	
Inscripción	<i>Nuevo</i>
Becas	
Cursos ECI 2012	
Programa	
Charlas y Eventos	
Calendario	
Charla Inaugural	
Charlas de Patrocinantes	
Experimento Empírico	
Posters	
Autoridades y Staff	
Auspiciantes	
Info Turística	
Cómo llegar	
Albergues	
Prensa y Difusión	

Agustin Gravano, John Hanse y Sergio Yovine

Haga clic aquí para ver la imagen a tamaño completo... — Tamaño: 287.6 kB

Últimas noticias

- Buen día compu, ¿qué hay de nuevo? 05/08/2013
- Estilos musicales y percepción humana 10/07/2013
- Entrevista a Victoria Bajar 28/06/2013
- Se abrió la inscripción a la ECI 2013 14/06/2013
- Lanzamiento del Desafío Dale Aceptar 2013 07/06/2013
- Más noticias...

Pictured above are: Agustin Gravano (Host), John Hansen, and Sergio Yovine (Host) in front of the summer school registration area (approximately +500 alumni, industry researchers, and students attended the summer school).





These pictures show: (i) dinner the night before for all Summer School instructors; (ii) Agustin Gravano and Sergio Yovine (Hosts) along with Hansen; (iii) attendees during the Hands-On computer lab session {all were allowed to receive full code and documentation for the DEMO setup; which would allow students to build speech recognition applications immediately}.

FINAL COMMENTS:

The potential to expand both speech and language processing activities in India and Argentina are excellent. There is great academic support, as well as motivated students. ISCA should continue to engage with these countries and encourage both participation in ISCA events, as well as representation on various committees and conferences/workshops.

Yours sincerely,

John H.L. Hansen, Associate Dean for Research, and Professor of Electrical Engineering
University Distinguished Chair in Telecommunications Engineering; IEEE Fellow; ISCA Fellow
Erik Jonsson School of Engineering and Computer Science
Professor, School of Brain and Behavioral Sciences (Speech and Hearing –Callier Center)
Center for Robust Speech Systems, Director
University of Texas at Dallas