

## Xavier Anguera Miró, Ph.D.

Ausias Marc 124, 5-1 – 08013 Barcelona  
+34 696954775  
[xanguera@gmail.com](mailto:xanguera@gmail.com)  
<http://www.xavieranguera.com>

---

### Motivation and goals

To help envision, research and develop the technologies of tomorrow.  
To hold a research/teaching position where I am able to apply my current technical skills while learning new ones.

### Research Interests

I am interested in the area of multimedia information retrieval using multimodal information.

### Education

- 2003 – 2006:           **Ph.D. Program in Telecommunications**  
Speech and language processing group (TALP)  
Department of theory of signal and communications (TSC).  
Technical University of Catalonia (UPC).  
**Thesis title:** “Robust Speaker Diarization for Meetings”.  
Thesis defended on Dec. 2006 and awarded Excellent with  
Summa Cum Laude.
- 1999 – 2001 :           **European Masters in Language and Speech (EMLS).**  
Sponsored by ISCA and EACL organizations.  
Technical University of Catalonia (UPC), University of Barcelona  
(UB) and University Pompeu Fabra (UPF).
- 1996 – 2001 :           **Telecommunications engineering (MS-EE)**  
Technical and superior School of Telecommunications  
engineering of Barcelona (ETSETB).  
**Masters Thesis title:** “Concatenation based Spanish Text-to-  
Speech Synthesis for embedded applications”.  
Technical University of Catalonia (UPC).
- 

### Selected publications

- “Speaker Diarization For Multiple-Distant-Microphone Meetings Using Several Sources of Information”, Jose M. Pardo, Xavier Anguera and Chuck Wooters, IEEE Transactions on Computers, September 2007, volume 56, number 9, pp. 1189-1224.
- “Acoustic beamforming for speaker diarization of meetings”, Xavier Anguera, Chuck Wooters and Javier Hernando, IEEE Transactions on Audio, Speech and Language Processing, September 2007, volume 15, number 7, pp.2011-2023.

- "Hybrid Speech/Non-Speech Detector Applied to Speaker Diarization of Meetings", Xavier Anguera, Mateu Aguilo, Chuck Wooters, Climent Nadeu and Javier Hernando, Speaker Odyssey 2006, San Juan de Puerto Rico, USA, June 2006.
  - "Partial Sequence Matching Using an Unbounded Dynamic Time Warping Algorithm", Xavier Anguera, Robert Macrae and Nuria Oliver, in Proc. ICASSP 2010
- 

## Professional experience

January 2009 – Present

**Associate professor** at Universitat Pompeu Fabra (about 9K students and 1k lecturers)

- Teaching Operating systems, C programming and concurrency to students from Computer Science and Electrical Engineering majors.

January 2008 – Present

**Research Scientist** at Telefónica I+D, Barcelona, Spain

- One of the founding members of the multimedia scientific group. Research on multimedia Information Retrieval technologies and in mobile computing.

January 2007 – December 2007:

**R&D Engineer** at Telefónica I+D, Barcelona, Spain

- Research on acoustic indexing within Spanish funded project I3Media (multimedia content analysis and generation in broadcasted media).

November 2006 – December 2006:

**Research stay** at LIMSI-CNRS, Paris, France

- Helped on the improvement of the Spanish ASR system within the European TC-Star project.

September 2005 – September 2006:

**Visiting researcher**, recipient of the ICSI fellowship between the International Computer Science Institute and the Spanish Education and Science Ministry (MEC).

- Research on robust Speaker Segmentation and Clustering for Meetings.
- Participation on the RT06s NIST Meetings evaluation.

September 2004 – September 2005:

**Visiting researcher** within the AMI Training Programme (AMI is funded by the 6<sup>th</sup> European Framework) at ICSI.

- Research on Robust Speaker Segmentation and clustering for Meetings.
- Speaker diarization in Broadcast News. Helped in the participation to the RT04F EARS Evaluation on speaker clustering and rich transcription.

- Participation on the RT06s NIST Meetings evaluation.

September 2003 – September 2004:

**Research position** in the TALP group (TSC department), UPC Barcelona.

- Speaker verification system and dialog implementation for PETRA project (Unified messaging system) and integration with the rest of the system. Project funded by the Spanish Government. Project finished in December 2003, worked on it from October to December 2003.
- Dialog system in VXML for retrieval of news from an online newspaper. Developed using Verbio dialog server from ATLAS Company.
- Research in speaker segmentation and tracking.
- Involved in the CHIL project (Computers in the Human Interaction Loop). Development of the speaker tracking system and integration of the demo room's audio technologies. Project funded by the European Union.
- Involved in the BIOSEC project (Biometrics and Security) on speaker recognition systems. Project funded by the European Union.

September 2001 – August 2003:

**Research Software Engineer** at Panasonic Speech Technology Laboratory (PSTL), Santa Barbara, California. Speech R&D Department.

- Improvement of Spanish Text-to-Speech system for unlimited text.
- Built a second robot prototype to perform research in human-machine interaction with control over the Internet and PDA by voice. Focus on dialog systems oriented to tele-operation and how robots can help humans.
- Assistance in the creation of an American Spanish Speech recognition Database within SPEECON project (creation of speech resources in multiple conditions in more than 20 languages). Project with participation of 12 Speech Companies.
- Conversion of French, Italian and German isolated words Text-To-Speech synthesis systems to unlimited synthesis.
- Development of software tools to automate Multilingual Speech labeling oriented to concatenation based speech synthesis.

February 2001 – September 2001:

**Full time Internship** at Panasonic Speech Technology Laboratory (PSTL), Santa Barbara, California. Speech R&D department.

- Participated in the creation of Spanish Text-To-Speech system for isolated words. Presented as my final degree thesis, with maximum qualification obtained.
- Research in Speaker Recognition based on NIST evaluations.
- Built a Robot prototype for multimodal dialog applications (human machine interaction). Focused towards dialog capabilities and to become a showcase of the company's speech technologies.

- Definition and design of a Speech Recognition Database for American Spanish (SPEECON). Phonetic description and creation of prompts.

May 2000 – December 2000:

**Part time Internship** at Hewlett Packard, Barcelona Site, Spain.

Large format printers R & D department.

- Research assistant in the writing system and optical sensors of large format printers.

---

## Skills

### - Computers:

Programming: Visual Basic, C, C++, Unix/Linux , Perl, Html, XML, HTML, javascript, Matlab.

Environments: Windows, Linux, Unix and Mac.

### - Languages:

Catalan: Fluent (Mother tongue).

Spanish: Fluent.

English: Fluent.

First certificate in English by Cambridge University.

Several years living in USA.

French: Knowledge of.

Level 1 in DELF exam from the official French institute.

---

## Other activities

- Toastmasters International (Public speaking and leadership organization). DTM (Distinguished Toastmasters) certification. Served in several leadership roles within the organization.
- Robotics enthusiast. Participated in multiple contests during university and built two robots for dialog interaction.
- Occasional ballroom dancer.

---

## Conference publications

### 2011

- “Fast Speaker Diarization Based on Binary Keys”, Xavier Anguera and Jean-François Bonastre, to appear in Proc. ICASSP 2011, Prague, Check Republic
- “Discriminant Binary Data Representation for Speaker Recognition”, Jean-François Bonastre, Xavier Anguera Miro, Pierre-Michel Bousquet, Driss Matrouf, to appear in Proc. ICASSP 2011, Prague, Check Republic.
- “Crosed-Form Expressions vs. BIC: a Comparison for Speaker Clustering”, Themis Stafylakis, Xavier Anguera, Vassilis Katsouros, George Carayannis, to appear in Proc. ICASSP 2011, Prague, Check Republic.

**2010**

- “Telefonica Research at TRECVID 2010 Content-Based Copy Detection”, Ehsan Younessian, Xavier Anguera, Tomasz Adamek, Nuria Oliver and David Marimon, NIST Trecvid Workshop notebook paper
- “Novel binary key representation for biometric speaker recognition”, Xavier Anguera and Jean-François Bonastre, in Proc. Interspeech 2010, Makuhari, Japan.
- “System output combination for improved speaker diarization”, Simon Bozonet, Nicholas Evans, Xavier Anguera, Oriol Vinyals, Gerald Friedland and Corinne Fredouille, in Proc. Interspeech 2010, Makuhari, Japan.
- “Improvements to the equal-parameter BIC for Speaker Diarization”, Themis Stafylakis, Xavier Anguera, in Proc. Interspeech 2010, Makuhari, Japan.
- “MuViSync: Realtime Music Video Alignment”, Robert Macrae, Xavier Anguera and Nuria Oliver, in Proc. ICME 2010
- “Enriching Music Mood Annotation by Semantic Association Reasoning”, Jun Wang, Xavier Anguera, Xiaou Chen and Deshun Yang, in Proc. AdMiRe Workshop, in ICME 2010
- “Partial Sequence Matching Using an Unbounded Dynamic Time Warping Algorithm”, Xavier Anguera, Robert Macrae and Nuria Oliver, in Proc. ICASSP 2010
- “Unrestricted Voice Annotations and Search of Personal Photographs in a Mobile Phone”, Xavier Anguera, Mauro Cherubini and Nuria Oliver, in Proc. Of Spoken Query 2010 Workshop on voice search at ICASSP 2010

**2009**

- “Telefónica Research Content-Based Copy Detection TRECVID Submission”, Xavier Anguera, Pere Obrador, Tomasz Adamek, David Marimon and Nuria Oliver, NIST Trecvid Workshop notebook paper
- “Revisiting the use of GMM-SVM for speaker verification”, Xavier Anguera, in Proc. Interspeech 2009
- “Multimodal video copy detection of social media”, Xavier Anguera, Pere Obrador and Nuria Oliver, in Proc. first SIGMM Workshop on Social Media (WSM2009) at ACM MM09
- “The role of tags and image aesthetics in social image search”, Pere Obrador, Xavier Anguera, Rodrigo de Oliveira and Nuria Oliver, in Proc. first SIGMM Workshop on Social Media (WSM2009) at ACM MM09
- “Text versus Speech: A Comparison of Tagging Input Modalities for Camera Phones”, M. Cherubini, X. Anguera, N. Oliver and R. de Oliveira, Proc. MobileHCI, Bonn, Germany, September 2009, (best paper award nominee)
- “Audio-Based Soccer Game Summarization”, Helenca Duxans, Xavier Anguera and David Conejero, in Proc. IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB09)
- “Audio-Based Automatic Management of Audio Commercials”, H. Duxans, D. Conejero and X. Anguera, in Proc. ICASSP 2009, Taipei, Taiwan. April 2009

**2008**

- "TV advertisements detection and clustering based on acoustic information", D. Conejero and X. Anguera, in proc. International Conference on Computational

Intelligence for Modelling, Control and Automation - CIMCA08, Viena, Austria, December 2008

- "MAMI: Multimodal Annotations on a Camera Phone", X. Anguera and N. Oliver, in Proc. MobileHCI, Amsterdam, September 2008
- "Sistema de Indexación Automática de Contenidos Multimedia", U.Urdapilleta, D. Conejero, X. Anguera, D. Cacenas and F.J. Caminero, in Proc. XVIII Jornadas Telecom I+D, Bilbao, Spain
- "Multimodal and Mobile Personal Image Retrieval: A User Study", X. Anguera, N. Oliver and M. Cherubini, in Proc. Workshop on Mobile Information Retrieval, MOBIR'08, Singapore
- "Multimodal Photo Annotation and Retrieval on a Mobile Phone", X. Anguera, J. Xu, N. Oliver, in Proc. ACM Intl. Conference on Multimedia Information Retrieval, Vancouver, Canada.

## **2007**

- "Model Complexity Selection and Cross-Validation EM Training for Robust Speaker Diarization", Xavier Anguera, Takahiro Shinozaki, Chuck Wooters and Javier Hernando, ICASSP, Hawaii, USA, April 2007.
- "Automatic Weighting for the Combination of TDOA and Acoustic Features in Speaker Diarization for Meetings", Xavier Anguera, Chuck Wooters, J.M Pardo and Javier Hernando, ICASSP, Hawaii, USA, April 2007.
- "Speaker Diarization for Conference Room: The UPC RT07s Evaluation System", Jordi Luque, Xavier Anguera, Andrey Temko, and Javier Hernando, RT07s Rich Transcription evaluation workshop, Washington, May 2007
- "The SRI-ICSI Spring 2007 Meeting and Lecture Recognition System", Andreas Stolcke, Xavier Anguera, Kofi Boakye, Ozgur Çetin, Adam Janin, Mathew Magimai-Doss, Chuck Wooters, and Jing Zheng, RT07s Rich Transcription evaluation workshop, Washington, May 2007

## **2006**

- "Purity Algorithms for Speaker Diarization of Meetings Data", Xavier Anguera, Chuck Wooters and Javier Hernando. ICASSP 2006, Toulouse, France, May 2006.
- "Speaker Diarization for Multi-Microphone Meetings Using only Between-Channel Differences", Jose M. Pardo, Xavier Anguera, Chuck Wooters, In S. Renals and S. Bengio, editors, Machine Learning for Multimodal Interaction: Third International Workshop (MLMI 2006), Lecture Notes in Computer Science. Springer
- "Automatic Cluster Complexity and Quantity Selection: Towards Robust Speaker Diarization", Xavier Anguera, Chuck Wooters, Javier Hernando, In S. Renals and S. Bengio, editors, Machine Learning for Multimodal Interaction: Third International Workshop (MLMI 2006), Lecture Notes in Computer Science. Springer
- "Robust Speaker Diarization for Meetings: ICSI RT06s Meetings Evaluation System", Xavier Anguera, Chuck Wooters and Jose M. Pardo, In S. Renals and S. Bengio, editors, Machine Learning for Multimodal Interaction: Third International Workshop (MLMI 2006), Lecture Notes in Computer Science. Springer
- "Purity Algorithms for Speaker Diarization of Meetings Data", Xavier Anguera, Chuck Wooters, Javier Hernando, MMUA 2006, Toulouse, France, May 2006.
- "Hybrid Speech/Non-Speech Detector Applied to Speaker Diarization of Meetings", Xavier Anguera, Mateu Aguilo, Chuck Wooters, Climent Nadeu and Javier Hernando, Speaker Odyssey 2006, San Juan de Puerto Rico, USA, June 2006.

- "Friends and Enemies: A Novel Initialization for Speaker Diarization", Xavier Anguera, Chuck Wooters and Javier Hernando, ICSLP'06, Pittsburgh, Pensilvania, USA, September 2006.
- "Robust Speaker Diarization for Meetings: ICSI RT06s evaluation system", Xavier Anguera, Chuck Wooters and Jose M. Pardo, ICSLP'06, Pittsburgh, Pensilvania, USA, September 2006.
- "The ICSI-SRI Spring 2006 Meeting Recognition System", A. Janin, A. Stolcke, X. Anguera, K. Boakye, O. Cetin, J. Frankel, and J. Zheng, In S. Renals and S. Bengio, editors, Machine Learning for Multimodal Interaction: Third International Workshop (MLMI 2006), Lecture Notes in Computer Science. Springer
- "Multi-Stream Speaker Diarization Systems for the Meetings Domain", Ascension Gallardo, Xavier Anguera and Chuck Wooters, ICSLP'06, Pittsburgh, Pensilvania, USA, September 2006.
- "Speaker Diarization for Multiple Distant Microphone Meetings: Mixing Acoustic Features And Inter-Channel Time Differences", Jose M. Pardo, Xavier Anguera and Chuck Wooters, ICSLP'06, Pittsburgh, Pensilvania, USA, September 2006.

## **2005**

- "XBIC: Real-Time Cross Probabilities Measure for Speaker Segmentation", Xavier Anguera. International Computer Science Institute Technical Report TR-05-008.
- "Robust Speaker Segmentation for Meetings: The ICSI-SRI Spring 2005 Diarization System", Xavier Anguera, Chuck Wooters, Barbara Peskin and Mateu Aguilo. In S. Renals and S. Bengio, editors, Machine Learning for Multimodal Interaction: Third International Workshop (MLMI 2005), Lecture Notes in Computer Science. Springer
- "Further Progress in Meeting Recognition: The ICSI-SRI Spring 2005 Speech-to-Text Evaluation System", Andreas Stolcke, Xavier Anguera, Kofy Boakye, Ozgur Cetin, Frantisek Grezl, Adam Janin, Arindam Mandal, Barbara Peskin, Chuck Wooters and Jing Zheng. In S. Renals and S. Bengio, editors, Machine Learning for Multimodal Interaction: Third International Workshop (MLMI 2005), Lecture Notes in Computer Science. Springer
- "Speaker Diarization for Multi-Party Meetings Using Acoustic Fusion", Xavier Anguera, Chuck Wooters and Javier Hernando. Automatic Speech Recognition and Understanding (ASRU). Puerto Rico, November 2005.
- "PETRA: Advanced Oral Interfaces for Unified Messaging Applications", David Hernando, Javier Hernando and Xavier Anguera. Buran magazine, IEEE Barcelona student branch. Number 22, September 2005.

## **2004**

- "Evolutive Speaker Segmentation using a Repository System", Xavier Anguera and Javier Hernando. ICSLP, Korea 2004.
- "Segmentació de locutor per a la indexació automàtica de bases de dades multimèdia en català", Xavier Anguera, Mireia Farrús, Javier Hernando and Alberto Abad. II Congrés d'enginyeria en llengua catalana, Andorra 2004.
- "Els sistemes de reconeixement de veu i traducció automàtica en català: present i futur", Mireia Farrús, Jan Anguita, Xavier Anguera, Josep M. Crego, Adrià de Gispert, Javier Hernando, Climent Nadeu. II Congrés d'enginyeria en llengua catalana, Andorra 2004.
- "XBIC: Nueva Medida para Segmentación de Locutor hacia el Indexado Automático de la Señal de Voz", Xavier Anguera, Javier Hernando and Jan Anguita. III Jornadas en Tecnología del Habla, Valencia, 17-10 Nov 2004.

- “Towards Robust Speaker Segmentation: The ICSI-SRI Fall 2004 Diarization System”, Chuck Wooters, James Fung, Barbara Peskin and Xavier Anguera. EARS Program RT-04 Workshop, nov 7-10 2004.
- 

### **Journal articles**

- “Speaker Diarization For Multiple-Distant-Microphone Meetings Using Several Sources of Information”, Jose M. Pardo, Xavier Anguera and Check Wooters, IEEE Transactions on Computers, September 2007, volume 56, number 9, pp. 1189-1224.
  - “Acoustic beamforming for speaker diarization of meetings”, Xavier Anguera, Chuck Wooters and Javier Hernando, IEEE Transactions on Audio, Speech and Language Processing, September 2007, volume 15, number 7, pp.2011-2023.
- 

### **Other publications**

- “Future Media Internet: Reseach challenges and the road ahead”, multiple authors, published by the European Commission, April 2010, ISBN: 978-92-79-15864-3
  - Final degree thesis for Telecommunications engineering: “Concatenation based Spanish Text-to-Speech Synthesis for embedded applications”. Presented on September 2001. Received maximum grade.
  - “Robust Speaker Diarization for Meetings”, Xavier Anguera, PhD Thesis, submitted in partial fulfillment for the PhD in the TSC department, UPC, Barcelona. Received Excellent qualification with Summa Cum Laude mention.
  - “Future Media Internet Research Challenges and the Road Ahead”, contribution as part of the “Future Media Internet” Task Force, European Commission, April 2010.
- 

### **Grants and prizes**

- Spanish fellowship “Programa Torres Quevedo” with a 3 years grant to work on multimedia Indexing at Telefonica Research. Granted in 2008.
  - ICSI fellowship for Spanish Scientists promoted by the Spanish Education and Science Ministry (MEC). This award is given once a year to 5 proficient Spanish scientists to pursue research stays at ICSI. Awarded for the period of September 2005 to September 2006.
  - AMI Training program for speech scientists. AMI project, March 2004. Awarded to European students in the speech area to perform research at ICSI from September 2004 to September 2005.
  - “Programa BE de mobilitat de personal investigador”. Generalitat de Catalunya, March 2004. Awarded to Catalan researchers to help in pursuing a stay in a foreign research institution. Awarded for a period of 6 months.
- 

### **Patents**

- “Multilingual Text-to-Speech system with limited resources”. Patent in progress

- “Interactive, Personalized Robot for Home use”. Patent in progress.
  - “Intelligent Nurse Robot”. Patent in progress.
  - 5 other patents in progress
- 

### **Community service**

- Conference reviews: Interspeech, ICASSP, ACM MM, ACM-MIR, ICME
  - Journal reviews: IEEE TASLP, ACM TOMCCAP, Journal of Computer Mathematics, Multimedia Tools and Applications (Springer)
  - TPCs: ACM MM applications track, IEEE-ICME, ACM MIR, ICMR
  - Conf. Organization: ACM-MIR 2009 (publicity chair), RECSYS 2010 (Local Chair), ICMI 2011 (Local Chair), ICME 2011 (Student Program Chair)
  - Event Organization:
    - ICME 2011 – Workshop on Multimodal Audio-based Multimedia Content Analysis (MAMCA)
- 

### **References**

Available upon request.