

Editorial Board

Simone Diniz Junqueira Barbosa

*Pontifical Catholic University of Rio de Janeiro (PUC-Rio),
Rio de Janeiro, Brazil*

Phoebe Chen

La Trobe University, Melbourne, Australia

Alfredo Cuzzocrea

ICAR-CNR and University of Calabria, Italy

Xiaoyong Du

Renmin University of China, Beijing, China

Joaquim Filipe

Polytechnic Institute of Setúbal, Portugal

Orhun Kara

TÜBİTAK BİLGEM and Middle East Technical University, Turkey

Tai-hoon Kim

Konkuk University, Chung-ju, Chungbuk, Korea

Igor Kotenko

*St. Petersburg Institute for Informatics and Automation
of the Russian Academy of Sciences, Russia*

Dominik Ślęzak

University of Warsaw and Infobright, Poland

Xiaokang Yang

Shanghai Jiao Tong University, China

Doroteo Torre Toledano Alfonso Ortega Giménez
António Teixeira Joaquín González Rodríguez
Luis Hernández Gómez Rubén San Segundo Hernández
Daniel Ramos Castro (Eds.)

Advances in Speech and Language Technologies for Iberian Languages

IberSPEECH 2012 Conference
Madrid, Spain, November 21-23, 2012
Proceedings

Volume Editors

Doroteo Torre Toledano
Universidad Autonoma de Madrid, Spain
E-mail: doroteo.torre@uam.es

Alfonso Ortega Giménez
Universidad de Zaragoza, Spain
E-mail: ortega@unizar.es

António Teixeira
Universidade de Aveiro, Portugal
E-mail: ajst@ua.pt

Joaquín González Rodríguez
Universidad Autonoma de Madrid, Spain
E-mail: joaquin.gonzalez@uam.es

Luis Hernández Gómez
Universidad Politécnica de Madrid, Spain
E-mail: luis@gaps.ssr.upm.es

Rubén San Segundo Hernández
Universidad Politécnica de Madrid, Spain
E-mail: lapiz@die.upm.es

Daniel Ramos Castro
Universidad Autonoma de Madrid, Spain
E-mail: daniel.ramos@uam.es

ISSN 1865-0929

e-ISSN 1865-0937

ISBN 978-3-642-35291-1

e-ISBN 978-3-642-35292-8

DOI 10.1007/978-3-642-35292-8

Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2012952611

CR Subject Classification (1998): I.2.7, I.2.6, H.5.1-2, H.5.5, I.5.2-4, I.7.5, I.2.1, J.5

© Springer-Verlag Berlin Heidelberg 2012

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

It was a pleasure and an honor to organize the *IberSPEECH 2012: Joint VII “Jornadas en Tecnología del Habla” and III Iberian SLTech Workshop*, that took place during November 21–23 in Madrid, Spain, hosted by the ATVS Biometric Research Group, Universidad Autónoma de Madrid.

This conference is the result of the merging of two conferences: the *Jornadas en Tecnología del Habla* (Spanish Speech Technology Workshop) and the Iberian SLTech Workshop. The first has been organized by the “*Red Temática en Tecnología del Habla*” (Spanish Speech Technology Thematic Network, <http://www.rthabla.es>) since 2000. This network was created in 1999 and currently includes over 200 researchers and 30 research groups in speech technology in Spain. The first Iberian SLTech Workshop was organized in Porto Salvo, Portugal, in 2009, by the Special Interest Group on Iberian Languages (SIG-IL, <http://www.il-sig.org/>) of the International Speech Communication Association (ISCA, <http://www.isca-speech.org>) and has been organized in conjunction with the “*Jornadas en Tecnología del Habla*” since 2010.

As a result, the *IberSPEECH: Joint “Jornadas en Tecnología del Habla” and Iberian SLTech Workshop* is one of the most important research meetings in the field of speech and language processing focusing on Iberian languages, attracting many researchers (about 120 in last edition), mainly from Spain and Portugal, and is also a natural meeting for researchers from Latin America. However, although the main focus is on Iberian languages and the Iberian region, the conference is not restricted to them. Proof of this are the ALBAYZIN Technology Competitive Evaluations, organized in conjunction with the conference, which in this edition attracted the interest of several research groups from all around the world, including the USA, UK, France, Japan, China, and Switzerland, among others.

The ALBAYZIN Technology Competitive Evaluations have been organized alongside with the conference since 2006, promoting the fair and transparent comparison of technology in different fields related to speech and language technology. In this edition we had five different evaluations: Language Recognition, Audio Segmentation, Speech Synthesis, Search on Speech, and Handwriting Recognition. The organization of each of these evaluations requires preparing development and test data, providing data along with a clear set of rules to the participants, and gathering and comparing results from participants. This organization was carried out by different groups of researchers and was crucial for the success of the evaluations. Although results from the evaluations cannot be included in this volume owing to timing restrictions, we would like to express our gratitude to the organizers and also to the participants in the evaluations.

In this edition we had over 80 articles submitted to the conference, and only 29 were selected for this publication. This selection was based on the scores and comments provided by our Scientific Review Committee, which includes over 75 researchers from different institutions mainly from Spain, Portugal, and Latin America, to which we also would like to express our deepest gratitude. Each article was reviewed by three different reviewers and the authors had some time to address the comments before submitting the camera-ready paper. The articles have been organized, following the oral sessions of the conference, into six different topics:

- Speaker Characterization and Recognition
- Audio and Speech Segmentation
- Pathology Detection and Speech Characterization
- Dialogue and Multimodal Systems
- Robustness in Automatic Speech Recognition
- Applications of Speech and Language Technologies

Besides the excellent research articles included in this volume, the conference had the pleasure of having three extraordinary keynote speakers: Jan “Honza” Cernocky (Brno University of Technology, BUT, Czech Republic), Philip Rose (Australian National University, Australia), and Pedro Moreno (Google Research, NY, USA).

The conference was mainly organized and supported by the Spanish Thematic Network on Speech Technology (“*Red Temática en Tecnología del Habla*”) and the ISCA Special Interest Group on Iberian Languages (SIG-IL). Besides this, we also received support from the Universidad Autónoma de Madrid (UAM) and the Campus Internacional Excelencia UAM+CSIC, which not only provided a fantastic venue for organizing the conference (the Escuela Politécnica Superior), but also financial support. Also, several companies provided financial support for the conference, including Google, Microsoft, and Telefónica (through Cátedra UAM-Telefónica). Last but not least, we had financial support from the MA2VICMR consortium. Without the financial support of all of them this conference would simply have not been possible.

We would also like to thank Springer, and in particular Alfred Hoffmann and Leonie Kunz, for the possibility of publishing this volume and their help and great work in preparing it. This will help increase the international impact of this conference.

Finally, we would like to thank all the people that have been putting their efforts in organizing this conference, and in particular the Organizing Committee and the local Organizing Committee, as well as all the authors that have presented their articles at the conference.

Doroteo Torre Toledano
Alfonso Ortega Giménez
António Teixeira

Organization

General Chairs

Doroteo Torre Toledano	Universidad Autónoma de Madrid, Spain
Alfonso Ortega Giménez	Universidad de Zaragoza, Spain
António Teixeira	Universidade de Aveiro, Portugal

Program Chairs

Joaquín González Rodríguez	Universidad Autónoma de Madrid
Rubén San Segundo Hernández	Universidad Politécnica de Madrid
Luis Hernández Gómez	Universidad Politécnica de Madrid

Publication Chair

Daniel Ramos Castro	Universidad Autónoma de Madrid
---------------------	--------------------------------

Demo Chairs

Daniela Braga	Microsoft (MLDC)
Alberto Abad Gareta	INESC

Award Chairs

Miguel Sales Dias	Microsoft (MLDC)
Climent Nadeu Camprubi	Universitat Politècnica de Catalunya

Evaluation Chairs

Javier Tejedor Noguerales	Universidad Autónoma de Madrid
Javier González Domínguez	Universidad Autónoma de Madrid

Scientific Review Committee

Alberto Abad Gareta	INESC
Olatz Arregi Uriarte	Euskal Herriko Unibertsitatea
José Miguel Benedí Ruiz	Universitat Politècnica de Valencia
M ^a Carmen Benítez Ortúzar	Universidad de Granada
Antonio Bonafonte Cávez	Universitat Politècnica de Catalunya
Daniela Braga	MLDC/Microsoft

VIII Organization

Francisco Campillo	Universidad de Vigo
Antonio Cardenal	Universidade de Vigo
Valentín Cardeñoso Payo	Universidad de Valladolid
Paula Carvalho	Universidade de Lisboa
Francisco Casacuberta Nolla	Universitat Politècnica de Valencia
María José Castro Bleda	Universitat Politècnica de Valencia
José Colás Pasamontes	Universidad Autónoma de Madrid
Ricardo de Córdoba Herralde	Universidad Politècnica de Madrid
Carmen de la Mota Gorriz	Universitat Autònoma de Barcelona
Ángel de la Torre Vega	Universidad de Granada
Laura Docío Fernández	Universidade de Vigo
Daniel Erro Eslava	Euskal Herriko Unibertsitatea
David Escudero Mancebo	Universidad de Valladolid
Rubén Fernández Pozo	Universidad Politècnica de Madrid
Javier Ferreiros López	Universidad Politècnica de Madrid
Ascensión Gallardo Antolín	Universidad Carlos III de Madrid
Fernando García Granada	Universitat Politècnica de Valencia
Carmen García Mateo	Universidade de Vigo
Marta Gatiús Vila	Universitat Politècnica de Catalunya
Juan Ignacio Godino Llorente	Universidad Politècnica de Madrid
Koldo Gojenola Gallettebeitia	Euskal Herriko Unibertsitatea
Pedro Gómez Vilda	Universidad Politècnica de Madrid
Javier González Domínguez	Universidad Autónoma de Madrid
Joaquín González Rodríguez	Universidad Autónoma de Madrid
Inma Hernaez Rioja	Euskal Herriko Unibertsitatea
Luis A. Hernández Gómez	Universidad Politècnica de Madrid
Fco. Javier Hernando Pericas	Universitat Politècnica de Catalunya
Lluís Felip Hurtado Oliver	Universitat Politècnica de Valencia
Eduardo Lleida Solano	Universidad de Zaragoza
Joaquín Llisterri Boix	Universitat Autònoma de Barcelona
Fernando Jesús López Colino	Universidad Autónoma de Madrid
Eduardo López Gonzalo	Universidad Politècnica de Madrid
María Teresa López Soto	Universidad de Sevilla
Ramón López-Cózar Delgado	Universidad de Granada
Nuno Mamede	INESC
José B. Mariño Acebal	Universitat Politècnica de Catalunya
Carlos David Martínez Hinarejos	Universitat Politècnica de Valencia
Helena Moniz	INESC
Juan Manuel Montero Martínez	Universidad Politècnica de Madrid
Nicolás Morales Mombiela	Nuance
Asunción Moreno Bilbao	Universitat Politècnica de Catalunya
Climent Nadeu Camprubi	Universitat Politècnica de Catalunya
Juan Luis Navarro Mesa	Universidad de Las Palmas de Gran Canaria
Eva Navas Córdón	Euskal Herriko Unibertsitatea
Juan Nolazco Flores	Tecnológico de Monterrey

Alfonso Ortega Giménez	Universidad de Zaragoza
Antonio Miguel Peinado Herreros	Universidad de Granada
Carmen Peláez Moreno	Universidad Carlos III de Madrid
José Luis Pérez Córdoba	Universidad de Granada
Paulo Quaresma	Universidade de Évora
Daniel Ramos Castro	Universidad Autónoma de Madrid
Andreia Rauber	Universidade Católica de Pelotas
José Adrián Rodríguez Fonollosa	Universitat Politècnica de Catalunya
Eduardo Rodríguez Banga	Universidade de Vigo
Luis Javier Rodríguez Fuentes	Euskal Herriko Unibertsitatea
Antonio José Rubio Ayuso	Universidad de Granada
Victoria Eugenia Sánchez Calle	Universidad de Granada
Joan Andreu Sánchez Peiró	Universitat Politècnica de Valencia
Emilio Sanchís Arnal	Universitat Politècnica de Valencia
Rubén San-Segundo Hernández	Universidad Politécnica de Madrid
Kepa Sarasola Gabiola	Euskal Herriko Unibertsitatea
Ibon Saratxaga Couceiro	Euskal Herriko Unibertsitatea
Encarnación Segarra Soriano	Universitat Politècnica de Valencia
Mário Silva	Universidade de Lisboa
Alberto Simões	ESEIG/IPP
Daniel Tapias Merino	SIGMA Technologies
António Teixeira	Universidade de Aveiro
Javier Tejedor Noguerales	Universidad Autónoma de Madrid
Doroteo Torre Toledano	Universidad Autónoma de Madrid
María Amparo Varona Fernández	Euskal Herriko Unibertsitatea
José Luis Vicedo González	Universitat d'Alacant

Local Organizing Committee

Doroteo Torre Toledano	Universidad Autónoma de Madrid
Joaquín González Rodríguez	Universidad Autónoma de Madrid
Daniel Ramos Castro	Universidad Autónoma de Madrid
Javier González Domínguez	Universidad Autónoma de Madrid
Javier Franco Pedroso	Universidad Autónoma de Madrid
Javier Tejedor Noguerales	Universidad Autónoma de Madrid
Fernando Jesús López Colino	Universidad Autónoma de Madrid
Javier Ortega García	Universidad Autónoma de Madrid
Julián Fierrez Aguilar	Universidad Autónoma de Madrid
Javier Galbally Herrero	Universidad Autónoma de Madrid
Rubén Vera Rodríguez	Universidad Autónoma de Madrid
Pedro Tomé González	Universidad Autónoma de Madrid
Leonardo Campillo Llanos	Universidad Autónoma de Madrid
Olga León Zurdo	Universidad Autónoma de Madrid
Antonio Moreno Sandoval	Universidad Autónoma de Madrid
Sara Antequera	Universidad Autónoma de Madrid

Table of Contents

Speaker Characterization and Recognition

Reliability Estimation of the Speaker Verification Decisions Using Bayesian Networks to Combine Information from Multiple Speech Quality Measures	1
<i>Jesús Villalba, Eduardo Lleida, Alfonso Ortega, and Antonio Miguel</i>	

On the use of Total Variability and Probabilistic Linear Discriminant Analysis for Speaker Verification on Short Utterances	11
<i>Javier González Domínguez, Rubén Zazo, and Joaquín González-Rodríguez</i>	

Cepstral Trajectories in Linguistic Units for Text-Independent Speaker Recognition	20
<i>Javier Franco-Pedroso, Fernando Espinoza-Cuadros, and Joaquín González-Rodríguez</i>	

Improving the Quality of Standard GMM-Based Voice Conversion Systems by Considering Physically Motivated Linear Transformations	30
<i>Tudor-Cătălin Zorilă, Daniel Erro, and Inma Hernaez</i>	

Evaluation of a New Beam-Search Formant Tracking Algorithm in Noisy Environments	40
<i>Dayana Ribas González, José Enrique García Laínez, Antonio Miguel, Alfonso Ortega Giménez, Eduardo Lleida, and José Ramón Calvo de Lara</i>	

Audio and Speech Segmentation

On the Influence of Automatic Segmentation and Clustering in Automatic Speech Recognition	49
<i>Paula Lopez-Otero, Laura Docio-Fernandez, Carmen Garcia-Mateo, and Antonio Cardenal-Lopez</i>	

Preliminary Results of Alignment of Text and Audio in News and Songs	59
<i>Darwin Patricio Córdova Lucero and Doroteo Torre Toledano</i>	

Aligning Very Long Speech Signals to Bilingual Transcriptions of Parliamentary Sessions	69
<i>Germán Bordel, Mikel Penagarikano, Luis Javier Rodríguez-Fuentes, and María Amparo Varona Fernández</i>	
Factor Analysis Segmentation and Classification in Broadcast News Domain	79
<i>Diego Castán, Alfonso Ortega Giménez, and Eduardo Lleida</i>	
Prosodic and Phonetic Features for Speaking Styles Classification and Detection	89
<i>Arlindo Veiga, Dirce Celorico, Jorge Proença, Sara Candeias, and Fernando Perdigão</i>	
 Pathology Detection and Speech Characterization	
Voice Pathology Detection on the Saarbrücken Voice Database with Calibration and Fusion of Scores Using MultiFocal Toolkit	99
<i>David Martínez, Eduardo Lleida, Alfonso Ortega, Antonio Miguel, and Jesús Villalba</i>	
Score Level versus Audio Level Fusion for Voice Pathology Detection on the Saarbrücken Voice Database	110
<i>David Martínez, Eduardo Lleida, Alfonso Ortega, and Antonio Miguel</i>	
Using HMM to Detect Speakers with Severe Obstructive Sleep Apnoea Syndrome	121
<i>Ana Montero Benavides, José Luis Blanco, Alejandra Fernández, Rubén Fernandez Pozo, Doroteo Torre Toledano, and Luis Hernández Gómez</i>	
Acoustic Analysis of European Portuguese Oral Vowels Produced by Children	129
<i>Catarina Oliveira, Maria Manuel Cunha, Samuel Silva, António Teixeira, and Pedro Sá-Couto</i>	
Impact of Age in ASR for the Elderly: Preliminary Experiments in European Portuguese	139
<i>Thomas Pellegrini, Isabel Trancoso, Annika Hämäläinen, António Calado, Miguel Sales Dias, and Daniela Braga</i>	

Dialogue and Multimodal Systems

Mutual Information and Perplexity Based Clustering of Dialogue Information for Dynamic Adaptation of Language Models	148
<i>Juan Manuel Lucas-Cuesta, Fernando Fernández-Martínez, Tirso Moreno, and Javier Ferreiros</i>	
A Multilingual SLU System Based on Semantic Decoding of Graphs of Words	158
<i>Marcos Calvo, Lluís Felip Hurtado, Fernando García, and Emilio Sanchís</i>	
Merging Intention and Emotion to Develop Adaptive Dialogue Systems	168
<i>Zoraida Callejas, David Griol, and Ramón López-Cózar Delgado</i>	
Language Technology for Handwritten Text Recognition	178
<i>Alejandro H. Toselli, Nicolás Serrano, Adrià Giménez-Pastor, Ihab Khoury, Alfons Juan, and Enrique Vidal</i>	
Character-Based Handwritten Text Recognition of Multilingual Documents	187
<i>Miguel A. del Agua, Nicolás Serrano, Jorge Civera, and Alfons Juan</i>	

Robustness in Automatic Speech Recognition

A Robust Pitch Extractor Based on DTW Lines and CASA with Application in Noisy Speech Recognition	197
<i>Juan A. Morales-Cordovilla, Pablo Cabañas-Molero, Antonio M. Peinado, and Victoria Sánchez</i>	
Speech Denoising Using Non-negative Matrix Factorization with Kullback-Leibler Divergence and Sparseness Constraints	207
<i>Jimmy Ludeña-Choez and Ascensión Gallardo-Antolín</i>	
MMSE Feature Reconstruction Based on an Occlusion Model for Robust ASR	217
<i>José A. González, Antonio M. Peinado, and Ángel M. Gómez</i>	
Automatic Speech Recognition Based on Ultrasonic Doppler Sensing for European Portuguese	227
<i>João Freitas, António Teixeira, Francisco Vaz, and Miguel Sales Dias</i>	

Applications of Speech and Language Technologies

Integrating a State-of-the-Art ASR System into the Opencast Matterhorn Platform	237
<i>Juan Daniel Valor Miró, Alejandro Pérez González de Martos, Jorge Civera, and Alfons Juan</i>	

Speech Reconstruction by Sparse Linear Prediction	247
<i>Ján Koloda, Antonio M. Peinado, and Victoria Sánchez</i>	
Steganographic Pulse-Based Recovery for Robust ACELP Transmission over Erasure Channels	257
<i>Domingo López-Oller, Ángel M. Gómez, José Luis Pérez Córdoba, Bernd Geiser, and Peter Vary</i>	
A Proposal for a Visual Speech Animation System for European Portuguese	267
<i>José Serra, Manuel Ribeiro, João Freitas, Verónica Orvalho, and Miguel Sales Dias</i>	
Online Learning of Log-Linear Weights in Interactive Machine Translation	277
<i>Francisco-Javier López-Salcedo, Germán Sanchis-Trilles, and Francisco Casacuberta Nolla</i>	
Author Index	287