

# Lecture Notes in Artificial Intelligence 8854

## Subseries of Lecture Notes in Computer Science

### LNAI Series Editors

Randy Goebel

*University of Alberta, Edmonton, Canada*

Yuzuru Tanaka

*Hokkaido University, Sapporo, Japan*

Wolfgang Wahlster

*DFKI and Saarland University, Saarbrücken, Germany*

### LNAI Founding Series Editor

Joerg Siekmann

*DFKI and Saarland University, Saarbrücken, Germany*

Juan Luis Navarro Mesa Alfonso Ortega  
António Teixeira Eduardo Hernández Pérez  
Pedro Quintana Morales Antonio Ravelo García  
Iván Guerra Moreno Doroteo T. Toledano (Eds.)

# Advances in Speech and Language Technologies for Iberian Languages

Second International Conference, IberSPEECH 2014  
Las Palmas de Gran Canaria, Spain, November 19-21, 2014  
Proceedings



Springer

## Volume Editors

Juan Luis Navarro Mesa  
Eduardo Hernández Pérez  
Pedro Quintana Morales  
Antonio Ravelo García  
Iván Guerra Moreno  
ETSIT, Las Palmas de Gran Canaria, Spain  
E-mail: {juanluis.navarro, eduardo.hernandez,  
pedro.quintana, antonio.ravelo}@ulpgc.es  
E-mail: iguerra@idetic.eu

Alfonso Ortega  
University of Zaragoza, Spain  
E-mail: ortega@unizar.es

António Teixeira  
University of Aveiro, Portugal  
E-mail: ajst@ua.pt

Doroteo T. Toledano  
Autonomous University of Madrid, Spain  
E-mail: doroteo.torre@uam.es

ISSN 0302-9743

e-ISSN 1611-3349

ISBN 978-3-319-13622-6

e-ISBN 978-3-319-13623-3

DOI 10.1007/978-3-319-13623-3

Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014955401

LNCS Sublibrary: SL 7 – Artificial Intelligence

© Springer International Publishing Switzerland 2014

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, 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.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

*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

The Spanish Thematic Network on Speech Technology (RTTH) and the ISCA-Special Interest Group on Iberian Languages (SIG-IL) are pleased to present the selected papers of IberSpeech 2014, Joint VIII Jornadas en Tecnologías del Habla and IV Iberian SLTech Workshop, held in Las Palmas de Gran Canaria, Spain, November 19–21. The Organizing Committee of IberSpeech believes and trusts that we have achieved the quality that the researchers in advances in speech and language technologies for Iberian languages value. To ensure this quality, each article was reviewed by at least three members of the Scientific Review Committee, who provided feedback to improve the final version of the articles in this book.

The conference has become mature as different editions have been organized, starting in Vigo 2010 with FALA and continuing in Madrid 2012 with the new denomination: IberSpeech. This new edition is a step forward the support of researchers in Iberian languages. IberSpeech is a joint event resulting from the merging of two conferences, the “Jornadas en Tecnología del Habla” and the Iberian SLTech Workshop. The first has been organized by the “Red Temática en Tecnologías 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 all over Spain. And the second, was organized by the Special Interest Group on Iberian Languages (SIG-IL, <http://www.il-sig.org/>) of the International Speech Communication Association (ISCA). The Iberian SLTech Workshop had its first edition in Porto Salvo, Portugal, in 2009.

As a result, IberSpeech is one of the most important research meetings in the field of speech and language processing focusing on Iberian languages, attracting many researchers (about 140 in the 2014 edition), mainly from Spain, Portugal, and from other Iberian-speaking countries in Latin America. We have also attracted the interest of several research groups from all around the world, including China, United Kingdom, France, Japan, Hungary, Israel, Norway, Czech Republic, and Germany.

Although the main focus is on Iberian languages and the Iberian region, the conference is not restricted to these topics. Proof of this are the ALBAYZIN Technology Competitive Evaluations, organized in conjunction with the conference, which in this edition have attracted the interest of several research groups. 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 have two different evaluations: Audio Segmentation and Search on Speech. The organization of each one of these evaluations requires the preparation of 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 is crucial for the success in participation that we are envisaging. Although results from the evaluations cannot be included in this volume due to timing restrictions, we would like to express our gratitude to the organizers and also to the participants in the evaluations.

We had 60 submitted papers and, after a strict peer-reviewing process, only 29 were selected for publication in this volume of Springer *Lecture Notes in Artificial Intelligence*. This selection was based on the scores and comments provided by our Scientific Review Committee, which includes over 79 researchers from different institutions mainly from Spain, Portugal, Latin America, USA, UK, Hungary, and Czech Republic, to whom we also would like to express our deepest gratitude. Each article was reviewed by at least three different reviewers and authors have had time to address the comments before submitting the camera-ready paper. The articles are organized into four different topics:

- Speech Production, Analysis, Coding and Synthesis
- Speaker and Language Characterization
- Automatic Speech Recognition
- Speech and Language Technologies in Different Application Fields

Besides the excellent research articles included in this volume, the conference had the pleasure of having two extraordinary keynote speakers: Dr. Pedro Gómez Vilda (Departamento de Arquitectura y Tecnología de Sistemas Informáticos de la Universidad Politécnica de Madrid, Spain) and Dr. Roger K. Moore (Department of Computer Science University of Sheffield, UK).

We would also like to thank Springer, and in particular Alfred Hoffmann, for the possibility of publishing this volume, his suggestions in order to increase the spread of the international scope of IberSpeech 2014, his help and great work in preparing it.

Finally, we would like to thank all those whose effort has made possible this conference, the members of the local Organizing Committee, the technical and program chairs, the reviewers and so many people who gave their best to achieve a successful conference.

November 2014

Juan Luis Navarro Mesa  
Alfonso Ortega Giménez  
António Teixeira

# Organization

## General Chairs

Juan Luis Navarro Mesa	Universidad de Las Palmas de Gran Canaria, Spain
Alfonso Ortega Giménez	Universidad de Zaragoza, Spain
António Teixeira	Universidade de Aveiro, Portugal

## Technical Chairs

Carmen García Mateo	Universidade de Vigo, Spain
Doroteo Torre Toledano	Universidad Autónoma de Madrid, Spain
Rubén San-Segundo Hernández	Universidad Politécnica de Madrid, Spain

## Program Chairs

Eduardo Hernández Pérez	Universidad de Las Palmas de Gran Canaria, Spain
Pedro Quintana Morales	Universidad de Las Palmas de Gran Canaria, Spain
Jesús Alonso Hernández	Universidad de Las Palmas de Gran Canaria, Spain

## Publication Chair

Sofía Martín González	Universidad de Las Palmas de Gran Canaria, Spain
-----------------------	---

## Demos Chairs

Javier Hernando Pericas	Universidad Politécnica de Cataluña, Spain
Rubén San-Segundo Hernández	Universidad Politécnica de Madrid, Spain
Daniela Braga	Voicebox Technologies, Portugal

## Awards Chairs

Javier Hernando Pericas	Universidad Politécnica de Cataluña, Spain
Inma Hernaez Rioja	Universidad del País Vasco, Spain
Daniela Braga	Voicebox Technologies, Portugal

## Evaluation Chairs

Antonio Ravelo García	Universidad de Las Palmas de Gran Canaria, Spain
Carlos Travieso González	Universidad de Las Palmas de Gran Canaria, Spain

## Local Organizing Committee

### Universidad de Las Palmas de Gran Canaria

Jesús B. Alonso Hernández  
Juan M. Caballero Suárez  
Fidel Cabrera Quintero  
Elena García Quevedo  
Iván D. Guerra Moreno  
Eduardo Hernández Pérez  
Sofía I. Martín González  
Manuel M. Medina Molina  
Juan Luis Navarro Mesa  
Pedro J. Quintana Morales  
Antonio G. Ravelo García  
Carlos Travieso González

## Evaluations and Scientific Review Committee

### Albayzin Evaluation Committee

Juan Luis Navarro Mesa	Universidad de Las Palmas de Gran Canaria, Spain
Alfonso Ortega Giménez	Universidad de Zaragoza, Spain
António Teixeira	Universidade de Aveiro, Portugal
Rubén San-Segundo Hernández	Universidad Politécnica de Madrid, Spain

### Scientific Review Committee

Alberto Abad	L2F/Spoken Language Systems Laboratory, Portugal
Jesús B. Alonso Hernández	Universidad de Las Palmas de Gran Canaria, Spain
Olatz Arregi Uriarte	Euskal Herriko Unibertsitatea, Spain
Plínio Barbosa	Universidade Estadual de Campinas, Portugal

Daniela Braga	Voicebox, Portugal
Juan M. Caballero Suárez	Universidad de Las Palmas de Gran Canaria, Spain
Fidel Cabrera Quintero	Universidad de Las Palmas de Gran Canaria, Spain
Valentín Cardeñoso Payo	Universidad de Valladolid, Spain
Paula Carvalho	Universidade de Lisboa, Portugal
María José Castro Bleda	Universitat Politècnica de Valencia, Spain
Jan Cernocky	BTU (Brno), Czech Republic
Chin Hui Lee	Georgia Tech, USA
Ricardo de Cordoba Herralde	Universidad Politècnica de Madrid, Spain
Carmen de la Mota Gorriz	Universitat Autònoma de Barcelona, Spain
Laura Docío Fernández	Universidade de Vigo, Spain
Daniel Erro Eslava	Euskal Herriko Unibertsitatea, Spain
David Escudero Mancebo	Universidad de Valladolid, Spain
Rubén Fernandez Pozo	Universidad Politècnica de Madrid, Spain
Javier Ferreiros López	Universidad Politècnica de Madrid, Spain
Julián Fierrez Aguilar	Universidad Autónoma de Madrid, Spain
Javier Franco Pedroso	Universidad Autónoma de Madrid, Spain
Ascensión Gallardo Antolín	Universidad Carlos III de Madrid, Spain
Carmen García Mateo	Universidade de Vigo, Spain
Elena García Quevedo	Universidad de Las Palmas de Gran Canaria, Spain
Juan Ignacio Godino Llorente	Universidad Politècnica de Madrid, Spain
Pedro Gómez Vilda	Universidad Politècnica de Madrid, Spain
Iván D. Guerra Moreno	Universidad de Las Palmas de Gran Canaria, Spain
Annika Hamalainen	Microsoft (MLDC), Portugal
Inma Hernaez Rioja	Euskal Herriko Unibertsitatea, Spain
Eduardo Hernández Pérez	Universidad de Las Palmas de Gran Canaria, Spain
Francisco Javier Hernando Pericas	Universitat Politècnica de Catalunya, Spain
Lluís Felip Hurtado Oliver	Universitat Politècnica de Valencia, Spain
Eduardo Lleida Solano	Universidad de Zaragoza, Spain
María Teresa López Soto	Universidad de Sevilla, Spain
Ramón López-Cózar Delgado	Universidad de Granada, Spain
Jordi Luque Serrano	Telefónica I+D, Spain
José B. Mariño Acebal	Universitat Politècnica de Catalunya, Spain
Sofía I. Martín González	Universidad de Las Palmas de Gran Canaria, Spain



Carlos David Martínez Hinarejos	Universitat Politècnica de Valencia, Spain
Manuel M. Medina Molina	Universidad de Las Palmas de Gran Canaria, Spain
Hugo Meinedo	Microsoft (MLDC), Portugal
Helena Moniz	INESC, Portugal
Juan Manuel Montero Martínez	Universidad Politècnica de Madrid, Spain
Nicolas Morales Mombiola	Nuance
Antonio Moreno Sandoval	Universidad Autónoma de Madrid, Spain
Climent Nadeu Camprubi	Universitat Politècnica de Catalunya, Spain
Juan Luis Navarro Mesa	Universidad de Las Palmas de Gran Canaria, Spain
Eva Navas Cordón	Euskal Herriko Unibertsitatea, Spain
Géza Nemeth	BME, Hungary
Juan Nolasco Flores	Tecnológico de Monterrey, Mexico
Alfonso Ortega Giménez	Universidad de Zaragoza, Spain
Antonio Miguel Peinado Herreros	Universidad de Granada, Spain
Carmen Peláez Moreno	Universidad Carlos III de Madrid, Spain
José Luis Pérez Córdoba	Universidad de Granada, Spain
Ferrán Pla Santamaría	Universitat Politècnica de Valencia, Spain
Paulo Quaresma	Universidade de Évora, Portugal
Pedro J. Quintana Morales	Universidad de Las Palmas de Gran Canaria, Spain
Daniel Ramos Castro	Universidad Autónoma de Madrid, Spain
Andreia Rauber	Universidade Católica de Pelotas, Portugal
Antonio G. Ravelo García	Universidad de Las Palmas de Gran Canaria, Spain
José Adrián Rodríguez Fonollosa	Universitat Politècnica de Catalunya, Spain
Eduardo Rodríguez Banga	Universidade de Vigo, Spain
Luis Javier Rodríguez Fuentes	Microsoft (MLDC), Portugal
Joan Andreu Sánchez Peiró	Universitat Politècnica de Valencia, Spain
Emilio Sanchís Arnal	Universitat Politècnica de Valencia, Spain
Rubén San-Segundo Hernández	Universidad Politècnica de Madrid, Spain
Diana Santos	University of Oslo, Norway
Kepa Sarasola Gabiola	Euskal Herriko Unibertsitatea, Spain
Encarnación Segarra Soriano	Universitat Politècnica de Valencia, Spain
Mário Silva	Universidade de Lisboa, Portugal
Alberto Simões	ESEIG/IPP, Portugal
Richard Stern	Carnegie Mellon University, USA

António Teixeira	Universidade de Aveiro, Portugal
Javier Tejedor Noguerales	Universidad de Alcalá de Henares, Spain
Doroteo Torre Toledano	Universidad Autónoma de Madrid, Spain
Isabel Trancoso	IST/University of Lisbon & INESC-ID, Portugal
Carlos Travieso González	Universidad de Las Palmas de Gran Canaria, Spain
María Amparo Varona Fernández	Euskal Herriko Unibertsitatea, Spain
Aline Villavicencio	Federal University of Rio Grande do Sul, Portugal

# Table of Contents

## Speech Production, Analysis, Coding and Synthesis

Analysis and Synthesis of Emotional Speech in Spanish for the Chat Domain: A Parametric Approach . . . . .	1
<i>Yesika Laplaza and Juan María Garrido</i>	
Developing a Basque TTS for the Navarro-Lapurdian Dialect . . . . .	11
<i>Eva Navas, Inma Hernaez, Daniel Erro, Jasone Salaberria, Beñat Oyharçabal, and Manuel Padilla</i>	
Fine Vocoder Tuning for HMM-Based Speech Synthesis: Effect of the Analysis Window Length . . . . .	21
<i>Agustín Alonso, Daniel Erro, Eva Navas, and Inma Hernaez</i>	
Quantitative Analysis of /l/ Production from RT-MRI: First Results . . .	30
<i>Samuel Silva, Paula Martins, Catarina Oliveira, and António Teixeira</i>	
Statistical Text-to-Speech Synthesis of Spanish Subtitles . . . . .	40
<i>S. Piqueras, M.A. del-Agua, A. Giménez, J. Civera, and A. Juan</i>	

## Speaker and Language Characterization

Unsupervised Accent Modeling for Language Identification . . . . .	49
<i>David Martínez González, Jesús Villalba, Eduardo Lleida, and Alfonso Ortega</i>	
Global Speaker Clustering towards Optimal Stopping Criterion in Binary Key Speaker Diarization . . . . .	59
<i>Héctor Delgado, Xavier Anguera, Corinne Fredouille, and Javier Serrano</i>	
Unsupervised Training of PLDA with Variational Bayes . . . . .	69
<i>Jesús Villalba and Eduardo Lleida</i>	
On the Use of Convolutional Neural Networks in Pairwise Language Recognition . . . . .	79
<i>Alicia Lozano-Diez, Javier Gonzalez-Dominguez, Ruben Zazo, Daniel Ramos, and Joaquín Gonzalez-Rodriguez</i>	
Global Impostor Selection for DBNs in Multi-session i-Vector Speaker Recognition . . . . .	89
<i>Omid Ghahabi and Javier Hernando</i>	

## Automatic Speech Recognition

Phoneme-Lattice to Phoneme-Sequence Matching Algorithm Based on Dynamic Programming . . . . .	99
<i>Ciro Gracia, Xavier Anquera, Jordi Luque, and Ittai Artzi</i>	
Deep Maxout Networks Applied to Noise-Robust Speech Recognition . . .	109
<i>F. de-la-Calle-Silos, A. Gallardo-Antolín, and C. Peláez-Moreno</i>	
A Deep Neural Network Approach for Missing-Data Mask Estimation on Dual-Microphone Smartphones: Application to Noise-Robust Speech Recognition . . . . .	119
<i>Iván López-Espejo, José A. González, Ángel M. Gómez, and Antonio M. Peinado</i>	
Language Model Adaptation for Lecture Transcription by Document Retrieval . . . . .	129
<i>A. Martínez-Villaronga, M.A. del-Agua, J.A. Silvestre-Cerdà, J. Andrés-Ferrer, and A. Juan</i>	
Articulatory Feature Extraction from Voice and Their Impact on Hybrid Acoustic Models . . . . .	138
<i>Jorge Llombart, Antonio Miguel, and Eduardo Lleida</i>	
CVX-Optimized Beamforming and Vector Taylor Series Compensation with German ASR Employing Star-Shaped Microphone Array . . . . .	148
<i>Juan A. Morales-Cordovilla, Hannes Pessentheiner, Martin Hagmüller, José A. González, and Gernot Kubin</i>	
Flexible Stand-Alone Keyword Recognition Application Using Dynamic Time Warping . . . . .	158
<i>Miquel Ferrarons, Xavier Anquera, and Jordi Luque</i>	
Confidence Measures in Automatic Speech Recognition Systems for Error Detection in Restricted Domains . . . . .	168
<i>Julia Olcoz, Alfonso Ortega, Antonio Miguel, and Eduardo Lleida</i>	
Recognition of Distant Voice Commands for Home Applications in Portuguese . . . . .	178
<i>Miguel Matos, Alberto Abad, Ramón Astudillo, and Isabel Trancoso</i>	

## Speech and Language Technologies in Different Application Fields

Assessing the Applicability of Surface EMG to Tongue Gesture Detection . . . . .	189
<i>João Freitas, Samuel Silva, António Teixeira, and Miguel Sales Dias</i>	

Towards Cross-Lingual Emotion Transplantation . . . . .	199
<i>Jaime Lorenzo-Trueba, Roberto Barra-Chicote, Junichi Yamagishi, and Juan M. Montero</i>	
A Preliminary Study of Acoustic Events Classification with Factor Analysis in Meeting Rooms . . . . .	209
<i>Diego Castán, Alfonso Ortega, Antonio Miguel, and Eduardo Lleida</i>	
A Spoken Language Database for Research on Moderate Cognitive Impairment: Design and Preliminary Analysis . . . . .	219
<i>Fernando Espinoza-Cuadros, Marlene A. Garcia-Zamora, Dania Torres-Boza, Carlos A. Ferrer-Riesgo, Ana Montero-Benavides, Eduardo Gonzalez-Moreira, and Luis A. Hernandez-Gómez</i>	
Towards Customized Automatic Segmentation of Subtitles . . . . .	229
<i>Aitor Álvarez, Haritz Arzelus, and Thierry Etchegoyhen</i>	
Bootstrapping a Portuguese WordNet from Galician, Spanish and English Wordnets . . . . .	239
<i>Alberto Simões and Xavier Gómez Guinovart</i>	
ATVS-CSLT-HCTLab System for NIST 2013 Open Keyword Search Evaluation . . . . .	249
<i>Javier Tejedor, Doroteo T. Toledano, and Dong Wang</i>	
Speech Watermarking Based on Coding of the Harmonic Phase . . . . .	259
<i>Inma Hernaez, Ibon Saratxaga, Jianpei Ye, Jon Sanchez, Daniel Erro, and Eva Navas</i>	
The Translectures-UPV Toolkit . . . . .	269
<i>M.A. del-Agua, A. Giménez, N. Serrano, J. Andrés-Ferrer, J. Civera, A. Sanchis, and A. Juan</i>	
The AhoSR Automatic Speech Recognition System . . . . .	279
<i>Igor Odriozola, Luis Serrano, Inma Hernaez, and Eva Navas</i>	
<b>Author Index . . . . .</b>	<b>289</b>