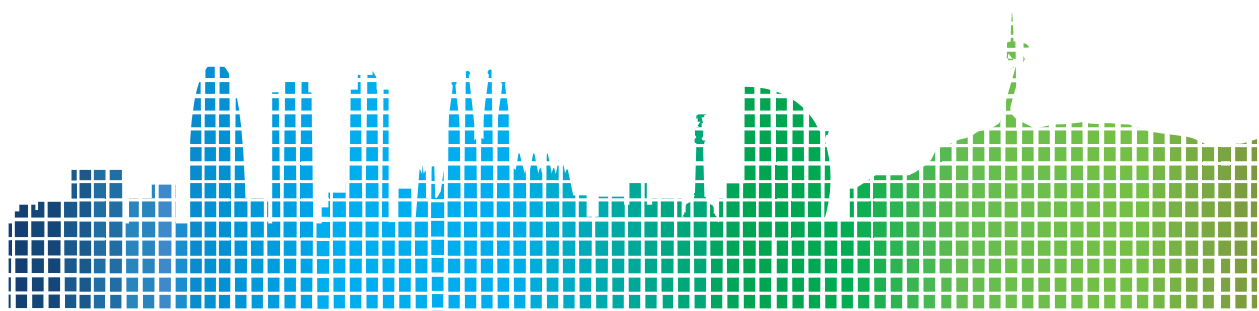


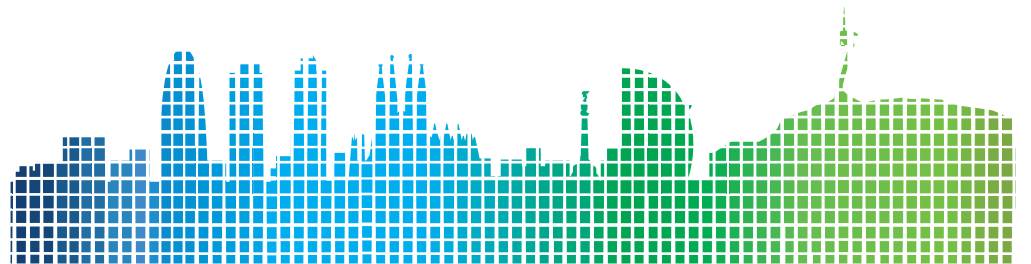
PROGRAM AND PROCEEDINGS

*Iber***SPEECH2018**
BARCELONA **NOVEMBER 21-23**



	Wed 21	Thu 22	Fri 23
8:00	Registration Starts at 08:00		
9:00	Opening 09:00-09:20	Oral 4: Synthesis, Production & Analysis 09:00-10:40	Oral 5: Text & NLP Applications 09:00 - 10:40
10:00	Oral 1: Speaker recognition 09:20-10:40		
10:40	Coffee break 10:40-11:00	Coffee break 10:40-11:00	Coffee break 10:40-11:00
11:00	Keynote: Tanja Schultz 11:00-12:00	Keynote: Rob Clark 11:00-12:00	Keynote: Lluís Marqués 11:00-12:00
12:00	Posters: Topics on Speech Technologies 12:00-13:30	Special Session: Projects, Demo and PhD thesis 12:00-13:30	Round Table 12:00-13:00
13:00			Closing 13:00-13:30
14:00	Lunch 13:30-15:00	Lunch 13:30-15:00	
15:00	Oral 2: ASR & Speech Applications 15:00-16:40	Albayzin Evaluations 15:00-16:40	
16:00		Coffee break 16:40-17:00	
17:00	Oral 3: Speech & Language Tech. Applied to Health 17:00-18:40	Albayzin Evaluations 17:00-19:00	
18:00		Coffee break 16:40-17:00	
19:00	RTTH Assembly 18:45-19:45		
20:00	Welcome Reception 20:00		
21:00		Gala Dinner 20:30	

*Iber***SPEECH2018**
BARCELONA **NOVEMBER 21-23**



Edited by Antonio Bonafonte, Jordi Luque and Jordi Pons

At the time of release, the proceedings can be downloaded from the website of IberSPEECH2018:
iberspeech2018.talp.cat

Barcelona, November 2018

Contents

Welcome Message	iii
Committees	vi
Organizing Committee	vi
Program Committee	viii
Organizing Institutions	xi
Awards	xii
Venue	xiii
Social Program	xvi
Invited Speakers	1
Tanja Schultz	1
Rob Clark	1
Lluís Màrquez	2
Technical Program	3
Abstracts	11
Author's Index	89

Welcome Message

Welcome to IberSPEECH2018 in Barcelona, November 21-23rd 2018, co-organized by Telefónica Research, the Center for Language and Speech Technologies and Applications (TALP) at the Universitat Politècnica de Catalunya (UPC), and the Research Group on Media Technologies (GTM) at La Salle, Universitat Ramon Llull. The current edition has received inestimable valued support by the Spanish Thematic Network on Speech Technology (RTTH), the Cátedra RTVE and the Voice Input Voice Output Laboratory (Vivolab) at Universidad de Zaragoza, and the ISCA Special Interest Group on Iberian Languages (SIG-IL). In addition, and for the very first time, IberSPEECH becomes an official ISCA Supported Event.

The IberSPEECH2018 event — the fourth of its kind using this name — brings together the 10th Jornadas en Tecnologías del Habla and the 6th Iberian SLTech Workshop events, aiming to promote interaction and discussion among junior and senior researchers in the field of speech and language processing for Iberian languages.

Barcelona is a modern capital of 1.7 million people and the sixth-most populous urban area in Europe. It is the home of many points of interest declared as World Heritage Sites by UNESCO such as Sagrada Família, Park Güell, and Palau de la Música Catalana, and also the birthplace for great minds like Antoni Gaudí, Joan Miró, Montserrat Caballé, Eduardo Mendoza, and many more. Barcelona offers a unique combination of landscapes and weather, coupled with exquisite gastronomical experiences that are the result of a blend of heritage, produce, terroir, tradition, creativity, and innovation.

The venue, the Telefónica Diagonal ZeroZero Tower by Spanish architects EMBA, is located at the very beginning of the famous Diagonal Avenue, which crosses Barcelona diagonally from the sea to the Llobregat river. The position of the Diagonal ZeroZero Tower is exceptional; it is very visible from the city and the coast, and it lays on the border between the consolidated city and the large expanses of public space in the Forum area. In fact, the address of the tower is Diagonal Avenue, number 0 and is just next to the Forum Building. The Forum area is next to the sea and the Besòs River end. It is a new business centre, located in a reformed area. Torre Telefónica is hosting several Telefónica Group companies as well as the Telefónica Research group, a leading industrial research lab following an open research model in collaboration with Universities and other research institutions. It promotes the dissemination of scientific results both through publications in top-tier peer-reviewed international journals, conferences, and technology transfer.

Following with the tradition of previous editions, IberSPEECH2018 will be a three-day event, bringing together the best researchers and practitioners in speech and language technologies in Iberian languages to promote interaction and discussion. The organizing committee has planned a wide variety of scientific and social activities, including technical paper presen-

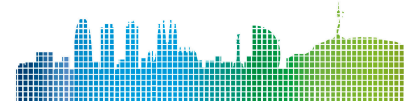
tations, keynote lectures, presentation of projects, laboratories activities, recent PhD thesis, discussion panels, a round table, and awards to the best thesis and papers.

The core of the scientific program of IberSPEECH2018 includes a total of 37 full regular paper contributions that will be presented distributed among 5 oral and 1 poster sessions. To ensure the quality of all the contributions, each submitted paper was reviewed by three members of the scientific review committee. All the papers in the conference will be accessible through the International Speech Communication Association (ISCA) Online Archive. Paper selection was based on the scores and comments provided by the scientific review committee, which includes over 86 researchers from different institutions (mainly from Spain and Portugal, but also from France, Germany, Brazil, Slovakia, Ireland, Greece, Hungary, Slovenia, Austria and United Kingdom).

Furthermore, it is confirmed to publish an extension of selected papers as a special issue of the Journal of Applied Sciences, "IberSPEECH 2018: Speech and Language Technologies for Iberian Languages", published by MDPI with fully open access. In addition to regular paper sessions, the IberSPEECH2018 scientific program features the following activities: the ALBAYZIN evaluation challenge session, a special session including the presentation of demos, research projects and recent PhD thesis, a round table and three keynote lectures.

Following the success of previous ALBAYZIN technology evaluations since 2006, this year ALBAYZIN evaluations have focused around multimedia analysis of TV broadcast content. Under the framework of a newly created Cátedra RTVE at Universidad de Zaragoza, we introduce and report on the results of the IberSPEECH-RTVE 2018 Challenge. The Corporación de Radiotelevisión Española (RTVE) has provided participants with an annotated TV broadcast database and the necessary tools for the evaluations, promoting the fair and transparent comparison of technology in different fields related to speech and language technology. It comprises four different challenge evaluations: Speech to Text Challenge (S2TC), Speaker Diarization Challenge (SDC) and Multimodal Diarization Challenge (MDC), organized by RTVE and Universidad de Zaragoza; and the Search on Speech Challenge (SoSC) jointly organized by Universidad San Pablo-CEU and AuDlaS from Universidad Autónoma de Madrid with the support of the ALBAYZIN Committee. Overall, 7 teams participated in the S2TC challenge, 8 teams in the SDC, 3 teams in the MDC, and 3 more teams in the SoSC challenge, which results in 21 system paper description contributions. Additionally, 11 special session papers are also included in the conference program. These were intended to describe either progress in current or recent research and development projects, demonstration systems, or PhD Thesis extended abstracts to compete in the PhD Award. Furthermore, IberSPEECH2018 features 3 remarkable keynote speakers: Prof. Tanja Schultz (University of Bremen, Germany and Institute of Carnegie Mellon, Pittsburgh, PA USA), Dr. Rob Clark (Google, London, UK) and Dr. Lluís Marquez (Amazon, Barcelona, Spain), to whom we would like to acknowledge for their extremely valuable participation.

Moreover, a round table with recognized experts brought discussion about the role of research and innovation from both academia and industry. Such a symbiosis creates market power by exploring and developing new categories which will eventually become the next blue oceans of our society. However, converting such activities into real businesses, making strong bones and figuring out new products that have a major impact in the real world is a non trivial task. A round table, we expected as an opportunity for both worlds on finding and exploring synergies and collaboration.



The social program of IberSPEECH2018 sets sail with the welcome reception at the Escola d'Enginyeria Barcelona Est (EEBE), at the recently created UPC Diagonal Besòs Campus, next to the Telefonica Tower. EEBE aims to become a top-quality academic centre in the field of engineering for the 21st-century industry that is capable of acting as an agent of transformation at a local and international level. The EEBE was born from the Barcelona College of Industrial Engineering (EUETIB) and from part of the teaching and research activity in chemical and materials engineering hitherto carried out at the Barcelona School of Industrial Engineering (ETSEIB). The gala dinner will be held at Restaurant Marítim, next to the legendary Barcelona Reial Club Marítim, designed by Lázaro Rosa-Violán from Contemporain Studio, to provide the aesthetics and flavours of different Mediterranean paradises.

Finally, we would like to thank all those whose effort made possible this conference, including the members of the organizing committee, the local organizing committee, the ALBAYZIN committee, the scientific reviewer committee, the authors, the conference attendees, the supporting institutions, and so many people who gave their best to achieve a successful conference.

Barcelona, November 2018
Jordi Luque, General Chair

Organizing Committee

General Chair:

Jordi Luque, Telefónica Research, Spain

General Co-Chairs:

Antonio Bonafonte, Universitat Politècnica de Catalunya, Spain

Francesc Alías Pujol, La Salle – Universitat Ramon LLull, Spain

António Teixeira, Universidade de Aveiro, Portugal

Technical Program Chair:

Javier Hernando Pericás, Universitat Politècnica de Catalunya, Spain

Technical Program Co-Chairs:

Alberto Abad, INESC ID Lisboa / IST Lisboa, Portugal

Xavier Anguera, ELSA Corp., United States

Carlos David Martínez Hinarejos, Universitat Politècnica de València, Spain

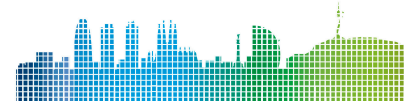
Eva Navas, University of the Basque Country, UPV- EHU, Spain

Carlos Segura, Telefónica Research, Spain

Publication Chairs:

Francesc Alías Pujol, La Salle – Universitat Ramon LLull, Spain

Antonio Bonafonte, Universitat Politècnica de Catalunya, Spain



Special Session and Awards Chair:

Ascensión Gallardo-Antolín, Universidad Carlos III, Spain

Plenary Talks & Round Table Chairs:

Joan Serrà, Telefónica Research, Spain

Marta Ruiz Costa-Jussà, Universitat Politècnica de Catalunya, Spain

Evaluations Chairs:

Alfonso Ortega, Universidad de Zaragoza, Spain

Eduardo Lleida, Universidad de Zaragoza, Spain

Luis Javier Rodríguez Fuentes, Universidad del País Vasco, Spain

Local Committee:

Francesc Alías Pujol, La Salle – Universitat Ramon LLull, Spain

Antonio Bonafonte, Universitat Politècnica de Catalunya, Spain

Jordi Luque, Telefónica Research, Spain

Jordi Pons, Universitat Pompeu Fabra, Spain

Bardia Rafieian, Universitat Politècnica de Catalunya, Spain

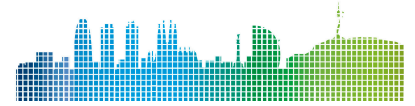
Marta Ruiz Costa-Jussà, Universitat Politècnica de Catalunya, Spain

Carlos Segura, Telefónica Research, Spain

Joan Serrà, Telefónica Research, Spain

Scientific Review Committee

Alberto Abad, INESC-IST, Portugal
Francesc Alías, La Salle – Universitat Ramon Llull, Spain
Aitor Alvarez, Vicomtech-IK4, Spain
Xavier Anguera, Miro ELSA Corp., Portugal
Jorge Baptista, INESC-ID Lisboa, Portugal
Plinio Barbosa, University of Campinas, Brazil
Fernando Batista, INESC-ID & ISCTE-IUL, Portugal
José Miguel Benedí, Universitat Politècnica de València, Spain
Antonio Bonafonte, Universitat Politècnica de Catalunya, Spain
Joao Cabral, Trinity College Dublin, Ireland
Francisco Casacuberta, Universitat Politècnica de València, Spain
María José Castro-Bleda, Universitat Politècnica de València, Spain
Ricardo Córdoba, Universidad Politécnica de Madrid, Spain
Conceicao Cunha, IPS Munich, Germany
Carme de-la-Mota, Universitat Autònoma de Barcelona, Spain
Arantza del Pozo, Vicomtech, Spain
Laura Docío-Fernández, University of Vigo, Spain
Daniel Erro, Cirrus Logic, Madrid, Spain
David Escudero, University of Valladolid, Spain
Nicholas Evans, EURECOM, France
Mireia Farrús, Universitat Pompeu Fabra, Spain
Rubén Fernández, Universidad Politécnica de Madrid, Spain
Javier Ferreiros, Universidad Politécnica de Madrid, Spain
Julian Fierrez, Universidad Autonoma de Madrid, Spain
Ascensión Gallardo, Universidad Carlos III de Madrid, Spain
Fernando García Granada, Universitat Politècnica de València, Spain
Carmen García Mateo, University of Vigo, Spain
Kafentzis George, University of Crete, Greece
Omid Ghahabi, EML European Media Laboratory GmbH, Germany
Juna Ignacio Godino, Llorente Universidad Politécnica de Madrid, Spain



Jon Ander Gómez, Universitat Politècnica de València, Spain
Emilio Granell, Universitat Politècnica de València, Spain
Inma Hernaez, University of the Basque Country (UPV/EHU), Spain
Javier Hernando, Universitat Politècnica de Catalunya, Spain
Lluís-F. Hurtado, Universitat Politècnica de València, Spain
Oliver Jokisch, Leipzig University of Telecommunications (HfTL), Germany
Oscar Koller, Microsoft Germany GmbH, Germany
Eduardo Lleida, University of Zaragoza, Spain
José David Lopes, Heriot Watt University UK
Paula López Otero, Universidade da Coruña, Spain
Jordi Luque, Telefónica Research, Spain
Carlos David Martínez Hinarejos, Universitat Politècnica de València, Spain
Helena Moniz, INESC/FLUL, Portugal
Juan Montero, Universidad Politécnica de Madrid, Spain
Nicolás Morales, Nuance Communications GmbH, Germany
Climent Nadeu, Universitat Politècnica de Catalunya, Spain
Juan L. Navarro-Mesa, Universidad de Las Palmas de Gran Canaria, Spain
Eva Navas, University of the Basque Country, Spain
Géza Németh, Budapest University of Technology & Economics, Hungary
Nelson Neto, Universidade Federal do Pará, Brazil
Hermann Ney, RWTH Aachen University, Germany
Alfonso Ortega, University of Zaragoza, Spain
Yannis Pantazis, Foundations for Research and Technology – Hellas, Spain
Carmen Peláez-Moreno University Carlos III Madrid, Spain
Thomas Pellegrini, Université de Toulouse; IRIT, France
Mikel Penagarikano, University of the Basque Country, Spain
Fernando Perdigao, Institute of Telecommunications (IT), Lisbon, Portugal
José L. Pérez-Córdoba, University of Granada, Spain
Ferran Pla, Universitat Politècnica de València, Spain
Jiri Pribil, Slovak Academy of Sciences Slovakia
Jorge Proenca, IT – Coimbra, Portugal
Michael Pucher, Acoustics Research Institute Austria
Paulo Quaresma, Universidade de Evora, Portugal
Ganna Raboshchuk, ELSA Corp., Portugal
Sam Ribeiro, The University of Edinburgh, UK
Eduardo Rodriguez Banga, University of Vigo, Spain
Marta Ruiz Costa-Jussà, Universitat Politècnica de Catalunya, Spain
Luis Javier Rodríguez-Fuentes, Univ. of the Basque Country UPV/EHU, Spain
Rubén San-Segundo, Universidad Politécnica de Madrid, Spain

Jon Sánchez, Aholab – EHU/UPV, Spain
Joan Andreu, Sanchez Universitat Politècnica de València, Spain
Emilio Sanchis, Universitat Politècnica de València, Spain
Diana Santos, University of Oslo, Norway
Ibon Saratxaga, University of the Basque Country, Spain
Encarna Segarra, Universitat Politècnica de València, Spain
Carlos Segura Perales, Telefónica Research, Spain
Joan Serrà, Telefónica Research, Spain
Alberto Simões, 2Ai Lab – IPCA, Portugal
Rubén Solera-Ureña, INESC-ID Lisboa, Portugal
António Teixeira, University of Aveiro, Portugal
Javier Tejedor, Universidad CEU San Pablo, Spain
Doroteo Toledano, Universidad Autónoma de Madrid, Spain
Isabel Trancoso, INESC ID Lisboa / IST, Portugal
Cassia Valentini-Botinhao, The University of Edinburgh, UK
Amparo Varona, University of the Basque Country, Spain
Andrej Zgank, University of Maribor, Slovenia
Catalin Zorila, Toshiba Cambridge Research Laboratory UK

Organizing Institutions

This conference has been organized by:



UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

laSalle

UNIVERSITAT RAMON LLULL

with the collaboration of:



Universidad
Zaragoza

rtve

IberSPEECH 2018 has been partially funded by the project Red Temática en Tecnologías del Habla 2017 (TEC2017-90829-REDT) founded by Ministerio de Ciencia, Innovación y Universidades.

Awards

Best Paper Award

All regular papers are candidates for this award. The award, given based on the review reports and the presentation at the conference, grants the authors the publication of an extended version of their work within the Special Issue of Applied Sciences journal (MDPI) entitled "IberSPEECH 2018: Speech and Language Technologies for Iberian Languages".

Best Albayzin evaluation system

Papers submitted to Albayzin evaluation tasks are candidates for these awards. The awards will be given to the winners of the Albayzin evaluation challenges, in accordance with the evaluation plan and rules defined for each task.

Best PhD Thesis award

Papers submitted to the PhD Thesis special session are candidates for this award. The award is given based on the decision of the committee formed by the members of the General chair, Technical Program chair and Special Session and Awards Chair. The award is given based on different criteria, including the quality of the document, impact of the thesis and clearness of the presentation at the conference.

Professional Career Prize in Speech Technologies

This is an honorary prize awarded by the Spanish Thematic Network on Speech Technology (RTTH) that recognizes experienced individuals who have made outstanding contributions related to speech technology research in Spain.

IberSPEECH2018 Edition

- José B. Mariño Acebal, Universitat Politècnica de Catalunya
- Antonio Rubio Ayuso, Universidad de Granada.

Main conference venue

Torre Telefónica

Address:

Torre Telefónica – Diagonal 00
Plaza de Ernest Lluch i Martín, 5
08019 Barcelona – Spain

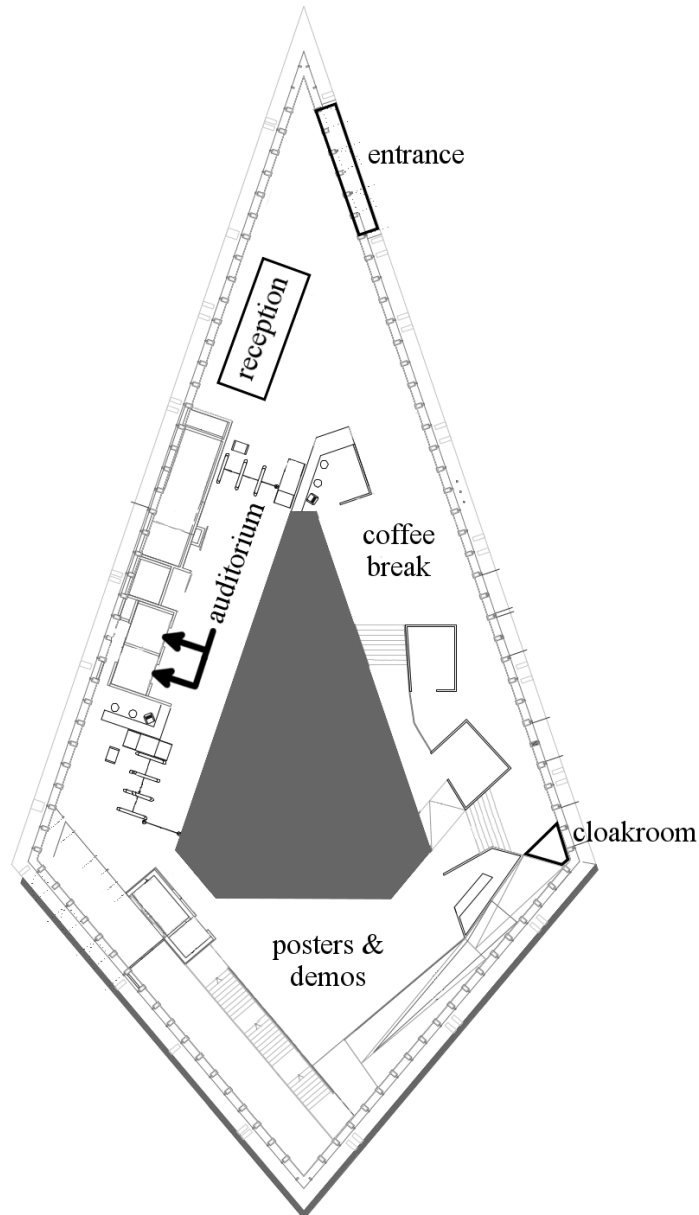


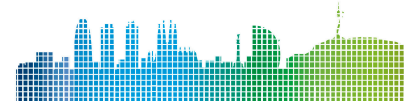
WI-FI access instructions

Connect to **IberSPEECH 2018**'s network with the following password:

- Network: Iberspeech 2018
- Password: *****

The main body of the conference will be held in Torre Telefónica (floor 0, see figure) and the Auditorium in floor 2 by accessing the elevators depicted in the figure. The following diagram outlines the main conference areas and services:





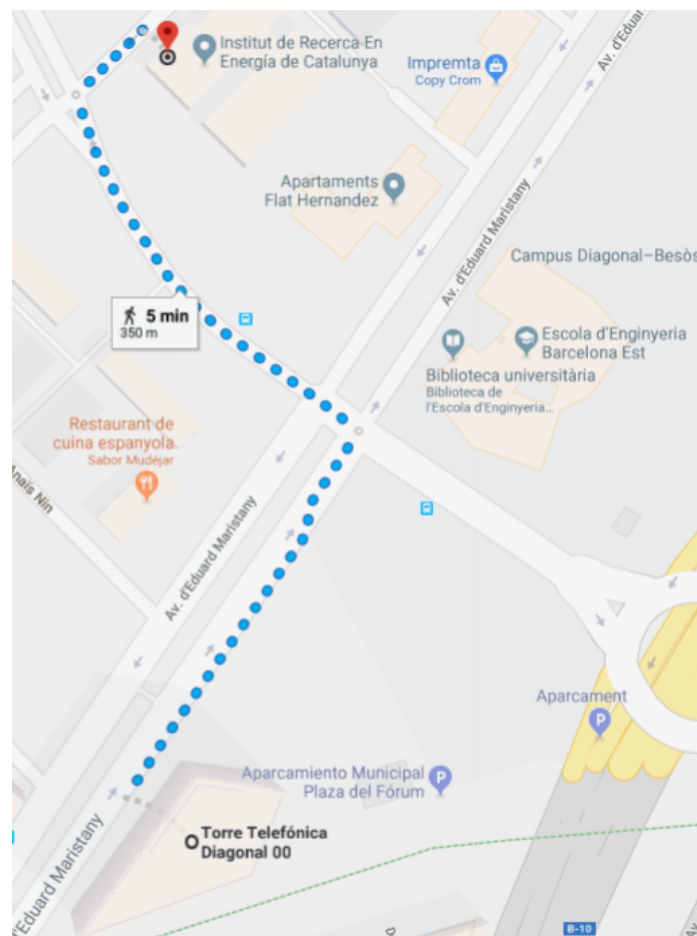
Lunch: 21 and 22 November, 13:30.

D'Ins Escola, Restaurant i Càtering

A gastronomic offer adapted to the occasion and a catering service with an added value that will enrich it: THE SOCIAL VALUE of the PEOPLE that work in this service. People who participate in a training and job placement program developed by the Fundació Formació i Treball.

Address: Carrer de Ramon Llull, 240, 08930, Sant Adrià del Besòs.

Wednesday 21st and Thursday 22nd lunches' will be given very close to Torre Telefónica (350 m).



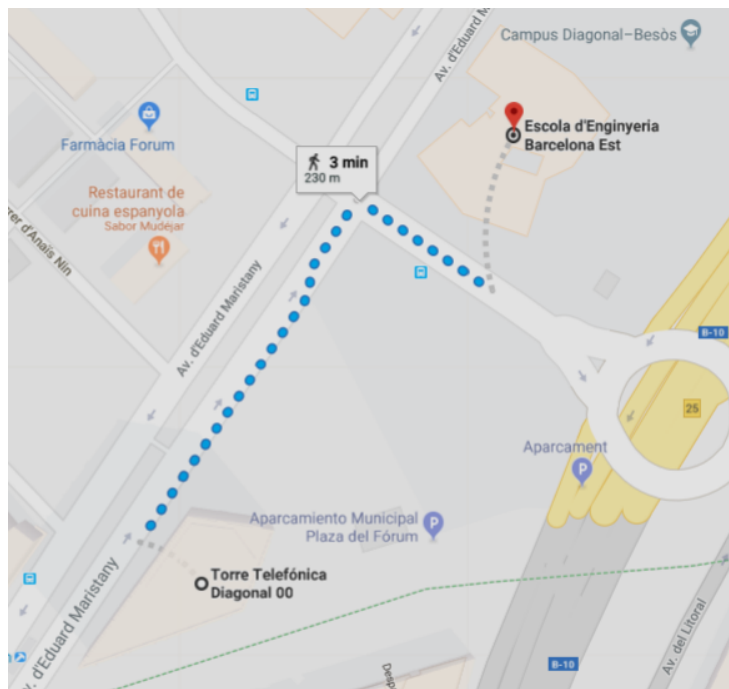
Social Program

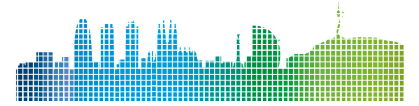
Welcome Reception: Wednesday 21 November, 20:00.

Escola d'Enginyeria Barcelona Est (UPC Diagonal Besòs)

Address: Av. Eduard Maristany, 16, 08019, Barcelona

The welcome reception will be given at the **Rambla de Colors** space in UPC Diagonal Besòs (230 m from Torre Telefónica). The space is placed between buildings C and I being possible to access it from both main entrances at each building. Use the number 37400 at the entrance phone to connect with reception and granting access to the building and go down to the **underground level**.





Gala dinner: Thursday 22 November, 20:30.

Restaurant Marítim

Address: Moll d’Espanya, 08039, Barcelona

Telephone: +34 93 221 17 75

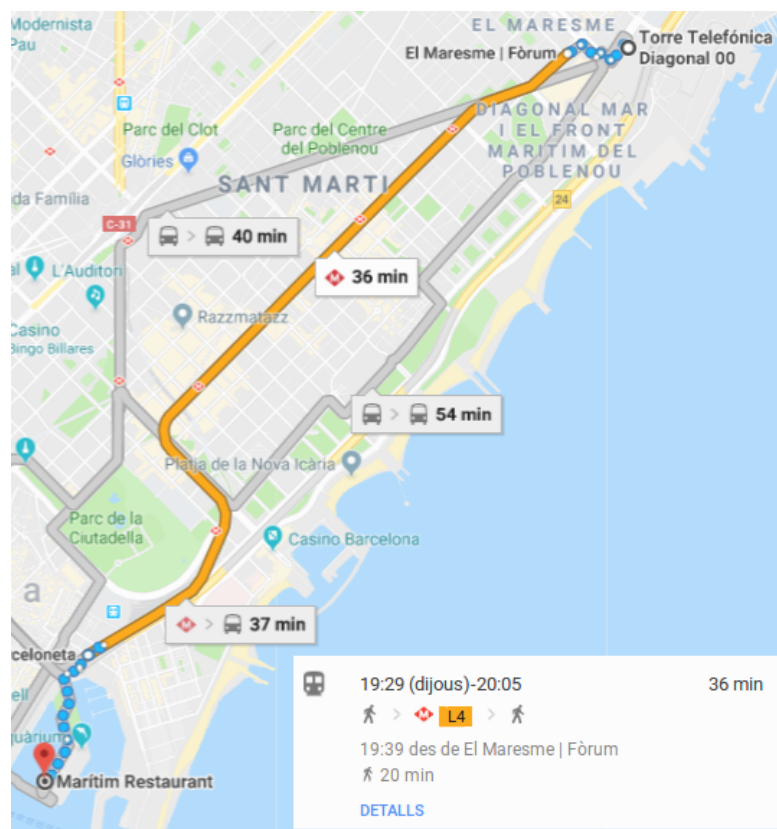
Web: www.maritimrestaurant.es

The gala dinner will be held downtown, next to the sea (close to Cristobal Colon statue).

How do I get there?

Take L4 from El Maresme – Fòrum metro station to Barceloneta.

You can also reach the restaurant by walking the seaside (1h 15 min).



Invited Speakers



Tanja Schultz received her diploma and doctoral degree in Informatics from University of Karlsruhe, Germany, in 1995 and 2000. Prior to these degrees she completed the state exam in Mathematics, Sports, Physical and Educational Science from Heidelberg University, Germany in 1989. She is currently the Professor for Cognitive Systems at the University of Bremen, Germany and adjunct Research Professor at the Language Technologies Institute of Carnegie Mellon, PA USA. Since 2007, she directs the Cognitive Systems Lab, where her research activities include multilingual speech recognition and the processing, recognition, and interpretation of biosignals for human-centered technologies and applications. Prior to joining University of Bremen, she was a Research Scientist at Carnegie Mellon (2000-2007) and a Full Professor at Karlsruhe Institute of Technology in Germany (2007-2015). Dr. Schultz is an Associate Editor of ACM Transactions on Asian Language Information Processing (since 2010), serves on the Editorial Board of Speech Communication (since 2004), and was Associate Editor of IEEE Transactions on Speech and Audio Processing (2002-2004). She was President (2014-2015) and elected Board Member (2006-2013) of ISCA, and a General Co-Chair of Interspeech 2006. She was elevated to Fellow of ISCA (2016) and to member of the European Academy of Sciences and Arts (2017). Dr. Schultz was the recipient of the Otto Haxel Award in 2013, the Alcatel Lucent Award for Technical Communication in 2012, the PLUX Wireless Biosignals Award in 2011, and the Allen Newell Medal for Research Excellence in 2002, and received the ISCA / EURASIP Speech Communication Best paper awards in 2001 and 2015.



Rob Clark received his PhD from the University of Edinburgh in 2003. His primary interest is in producing engaging synthetic speech. Before joining Google Rob was at the University of Edinburgh for many years involved in both teaching and research relating to text-to-speech synthesis. Rob was one of the primary developers and maintainers of the open source Festival text-to-speech synthesis system. In 2015 he joined Google where he is working on text-to-speech synthesis and prosody.



Lluís Màrquez is a Principal Applied Scientist at Amazon Research in Barcelona. From 2013 to 2017 he had a Principal Scientist role at the Arabic Language Technologies group from the Qatar Computing Research Institute (QCRI), and previously, he was Associate Professor at the Technical University of Catalonia (UPC, 2000-2013). He holds a university award-winning PhD in Computer Science from UPC (1999). His research focuses on natural language understanding by using statistical machine learning models. He has 150+ papers in Natural Language Processing and Machine Learning journals and conferences. He has been General and Program Co-chair of major conferences in the area (ACL, EMNLP, EACL, CoNLL, *SEM, EAMT, etc.), and held several organizational roles in ACL and EMNLP too. He was co-organizer of various international evaluation tasks at Senseval/SemEval (2004, 2007, 2010, 2015-2017) and CoNLL shared tasks (2004-2005, 2008-2009). He was Secretary and President of the ACL special interest group on Natural Language Learning (SIGNLL) in the period 2007-2011. More recently, he was President-elect and President of the European Chapter of the ACL (EACL; 2013-2016) and member of the ACL Executive Committee (2015-2016). Lluís Màrquez has been Guest Editor of special issues at Computational Linguistics, LRE, JNLE, and JAIR in the period (2007-2015). He has participated in 16 national and EU research projects, and 2 projects with technology transfer to the industry, acting as the principal site researcher in 10 of them, helping companies embed AI in their business.

Speaker Recognition

Wednesday, 21 November 2018, 09:20 – 10:40

Chair: Xavier Anguera, ELSA Corp.

- O1.1** Differentiable Supervector Extraction for Encoding Speaker and Phrase Information in Text Dependent Speaker Verification
09:20 - 09:40 Victoria Mingote, Antonio Miguel, Alfonso Ortega, Eduardo Lleida
- O1.2** Phonetic Variability Influence on Short Utterances in Speaker Verification
09:40 - 10:00 Ignacio Viñals, Alfonso Ortega, Antonio Miguel, Eduardo Lleida
- O1.3** Restricted Boltzmann Machine Vectors for Speaker Clustering
10:00 - 10:20 Umair Khan, Pooyan Safari, Javier Hernando
- O1.4** Speaker Recognition under Stress Conditions
10:20 - 10:40 Esther Rituerto-González, Ascensión Gallardo-Antolín, Carmen Peláez-Moreno

Keynote 1

Wednesday, 21 November 2018, 11:00 – 12:00

Chair: Joan Serrà, Telefónica Research

- KN1** Bio signal-based Spoken Communication
11:00 - 12:00 Tanja Schultz

Topics on Speech Technologies

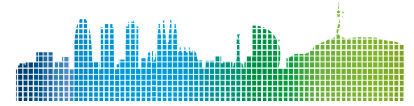
Wednesday, 21 November 2018

12:00 – 13:30

Chair: Alberto Abad, INESC-ID/IST

- P1.1** Bilingual Prosodic Dataset Compilation for Spoken Language Translation
12:00 - 13:30 Alp Öktem, Mireia Farrús, Antonio Bonafonte
- P1.2** Building an Open Source Automatic Speech Recognition System for Catalan
12:00 - 13:30 Baybars Külebi, Alp Öktem

- P1.3** Multi-Speaker Neural Vocoder
12:00 - 13:30 Oriol Barbany, Antonio Bonafonte, Santiago Pascual
- P1.4** Improving the Automatic Speech Recognition through the improvement of Language Models
12:00 - 13:30 Andrés Piñeiro-Martín, Carmen García-Mateo, Laura Docío-Fernández
- P1.5** Towards expressive prosody generation in TTS for reading aloud applications
12:00 - 13:30 Monica Dominguez, Alicia Burga, Mireia Farrús, Leo Wanner
- P1.6** Performance evaluation of front- and back-end techniques for ASV spoofing detection systems based on deep features
12:00 - 13:30 Alejandro Gomez-Alanis, Antonio M. Peinado, José Andrés González López, Angel M. Gomez
- P1.7** The observation likelihood of silence: analysis and prospects for VAD applications
12:00 - 13:30 Igor Odriozola, Inma Hernaez, Eva Navas, Luis Serrano, Jon Sanchez
- P1.8** On the use of Phone-based Embeddings for Language Recognition
12:00 - 13:30 Christian Salamea, Ricardo de Córdoba, Luis Fernando D'Haro, Rubén San-Segundo, Javier Ferreiros
- P1.9** End-to-End Speech Translation with the Transformer
12:00 - 13:30 Laura Cross Vila, Carlos Escolano, José A. R. Fonollosa, Marta R. Costa-Jussà
- P1.10** Audio event detection on Google's Audio Set database: Preliminary results using different types of DNNs
12:00 - 13:30 Javier Darna-Sequeiros, Doroteo T. Toledano
- P1.11** Emotion Detection from Speech and Text
12:00 - 13:30 Mikel de Velasco, Raquel Justo, Josu Antón, Mikel Carrilero, M. Inés Torres
- P1.12** Experimental Framework Design for Sign Language Automatic Recognition
12:00 - 13:30 Darío Tilves Santiago, Ian Benderitter, Carmen García-Mateo
- P1.13** Baseline Acoustic Models for Brazilian Portuguese Using Kaldi Tools
12:00 - 13:30 Cassio Batista, Ana Larissa Dias, Nelson Sampaio Neto



ASR & Speech Applications

Wednesday, 21 November 2018, 15:00 – 16:40

Chair: Carmen García Mateo, University of Vigo

- O2.1** Converted Mel-Cepstral Coefficients for Gender Variability Reduction in
15:00 - 15:20 Query-by-Example Spoken Document Retrieval
 Paula López Otero, Laura Docío-Fernández
- O2.2** A Recurrent Neural Network Approach to Audio Segmentation for Broadcast
15:20 - 15:40 Domain Data
 Pablo Gimeno, Ignacio Viñals, Alfonso Ortega, Antonio Miguel, Eduardo
 Lleida
- O2.3** Improving Transcription of Manuscripts with Multimodality and Interaction
15:40 - 16:00 Emilio Granell, Carlos David Martínez Hinarejos, Verónica Romero
- O2.4** Improving Pronunciation of Spanish as a Foreign Language for L1 Japanese
16:00 - 16:20 Speakers with Japañol CAPT Tool
 Cristian Tejedor-García, Valentín Cardeñoso-Payo, María J. Machuca, David
 Escudero-Mancebo, Antonio Ríos, Takuya Kimura
- O2.5** Exploring E2E speech recognition systems for new languages
16:20 - 16:40 Conrad Bernath, Aitor Alvarez, Haritz Arzelus, Carlos David Martínez

Speech & Language Technologies Applied to Health

Wednesday, 21 November 2018, 17:00 – 18:40

Chair: Mireia Farrús, Universitat Pompeu Fabra

- O3.1** Listening to Laryngectomees: A study of Intelligibility and Self-reported Lis-
17:00 - 17:20 tening Effort of Spanish Oesophageal Speech
 Sneha Raman, Inma Hernaez, Eva Navas, Luis Serrano
- O3.2** Towards an automatic evaluation of the prosody of people with Down syn-
17:20 - 17:40 drome
 Mario Corrales-Astorgano, Pastora Martínez-Castilla, David Escudero-
 Mancebo, Lourdes Aguilar, César González-Ferreras, Valentín Cardeñoso-
 Payo
- O3.3** Whispered-to-voiced Alaryngeal Speech Conversion with Generative Adver-
17:40 - 18:00 sarial Networks
 Santiago Pascual, Antonio Bonafonte, Joan Serrà, José Andrés González
 López
- O3.4** LSTM based voice conversion for laryngectomees
18:00 - 18:20 Luis Serrano, David Tavarez, Xabier Sarasola, Sneha Raman, Ibon Saratxaga,
 Eva Navas, Inma Hernaez
- O3.5** Sign Language Gesture Classification using Neural Networks
18:20 - 18:40 Zuzanna Parcheta, Carlos David Martínez Hinarejos

Synthesis, Production & Analysis

Thursday, 22 November 2018, 09:00 – 10:40

Chair: Francesc Alías Pujol, La Salle - Universitat Ramon Llull

- O4.1** Influence of tense, modal and lax phonation on the three-dimensional finite element synthesis of vowel [A]
09:00 - 09:20 Marc Freixes, Marc Arnela, Joan Claudi Socoró, Francesc Alías Pujol, Oriol Guasch
- O4.2** Exploring Advances in Real-time MRI for Speech Production Studies of European Portuguese
09:20 - 09:40 Conceicao Cunha, Samuel Silva, António Teixeira, Catarina Oliveira, Paula Martins, Arun Joseph, Jens Frahm
- O4.3** A postfiltering approach for dual-microphone smartphones
09:40 - 10:00 Juan M. Martín-Doñas, Iván López-Espejo, Angel M. Gomez, Antonio M. Peinado
- O4.4** Speech and monophonic singing segmentation using pitch parameters
10:00 - 10:20 Xabier Sarasola, Eva Navas, David Tavares, Luis Serrano, Ibon Saratxaga
- O4.5** Self-Attention Linguistic-Acoustic Decoder
10:20 - 10:40 Santiago Pascual, Antonio Bonafonte, Joan Serrà

Keynote 2

Thursday, 22 November 2018, 11:00 – 12:00

Chair: Antonio Bonafonte, Universitat Politècnica de Catalunya

- KN2** Synthesizing variation in prosody for Text-to-Speech
11:00 - 12:00 Rob Clark

Special Session

Thursday, 22 November 2018

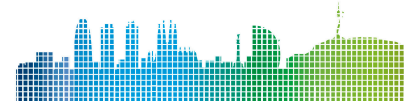
12:00 – 13:30

Chair: Ricardo de Córdoba, Universidad Politécnica de Madrid

- PhD Award** Thesis in 4 Minutes competition. Papers SP.8, SP.9, SP.10 and SP.11
12:00 - 12:20 Emilio Granell, Alicia Lozano-Diez, Omid Ghahabi, Igor Jauk

Show & Tell

- SP.1** Japañol: a mobile application to help improving Spanish pronunciation by Japanese native speakers
12:20 - 13:30 Cristian Tejedor-García, Valentín Cardeñoso-Payo, David Escudero-Mancebo



Ongoing Research Projects

- SP.2** 12:20 - 13:30 Towards the Application of Global Quality-of-Service Metrics in Biometric Systems
Juan Manuel Espín, Roberto Font, Juan Francisco Inglés-Romero, Cristina Vicente-Chicote
- SP.3** 12:20 - 13:30 Incorporation of a Module for Automatic Prediction of Oral Productions Quality in a Learning Video Game
David Escudero-Mancebo, Valentín Cardeñoso-Payo
- SP.4** 12:20 - 13:30 Silent Speech: Restoring the Power of Speech to People whose Larynx has been Removed
José Andrés González López, Phil D. Green, Damian Murphy, Amelia Gully, James M. Gilbert
- SP.5** 12:20 - 13:30 RESTORE Project: REpair, STorage and REhabilitation of speech
Inma Hernaez, Eva Navas, Jose Antonio Municio Martín, Javier Gomez Suárez
- SP.6** 12:20 - 13:30 Corpus for Cyberbullying Prevention
Asuncion Moreno, Antonio Bonafonte, Igor Jauk, Laia Tarrés, Victor Pereira
- SP.7** 12:20 - 13:30 EMPATHIC, Expressive, Advanced Virtual Coach to Improve Independent Healthy-Life-Years of the Elderly
M. Inés Torres, Gérard Chollet, César Montenegro, Jofre Tenorio-Laranga, Olga Gordeeva, Anna Esposito, Cornelius Glackin, Stephan Schlögl, Olivier Deroo, Begoña Fernández-Ruanova, Riberto Santana, Maria S. Kornes, Fred Lindner, Daria Kyslitska, Miriam Reiner, Gennaro Cordasco, Mari Aksnes, Raquel Justo

PhD Thesis

- SP.8** 12:20 - 13:30 Advances on the Transcription of Historical Manuscripts based on Multi-modality, Interactivity and Crowdsourcing
Emilio Granell, Carlos David Martinez Hinarejos, Verónica Romero
- SP.9** 12:20 - 13:30 Bottleneck and Embedding Representation of Speech for DNN-based Language and Speaker Recognition
Alicia Lozano-Diez, Joaquin Gonzalez-Rodriguez, Javier Gonzalez-Dominguez
- SP.10** 12:20 - 13:30 Deep Learning for i-Vector Speaker and Language Recognition: A Ph.D. Thesis Overview
Omid Ghahabi
- SP.11** 12:20 - 13:30 Unsupervised Learning for Expressive Speech Synthesis
Igor Jauk

Albayzin Evaluation

Thursday, 22 November 2018

15:00 – 16:40

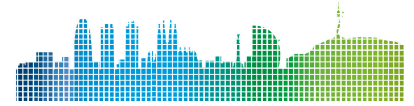
Chair: Alfonso Ortega & Eduardo Lleida, Universidad de Zaragoza

Multimodal Diarization Challenge

- AE.1** ODESSA/PLUMCOT at Albayzin Multimodal Diarization Challenge 2018
15:00 - 16:40 Benjamin Maurice, Hervé Bredin, Ruiqing Yin, Jose Patino, Héctor Delgado, Claude Barras, Nicholas Evans, Camille Guinaudeau
- AE.2** UPC Multimodal Speaker Diarization System for the 2018 Albayzin Challenge
15:00 - 16:40 Miquel Angel India Massana, Itziar Sagastiberri, Ponç Palau, Elisa Sayrol, Josep Ramon Morros, Javier Hernando
- AE.3** The GTM-UVIGO System for Audiovisual Diarization
15:00 - 16:40 Eduardo Ramos-Muguerza, Laura Docío-Fernández, José Luis Alba-Castro

Speaker Diarization Challenge

- AE.4** The SRI International STAR-LAB System Description for IberSPEECH-RTVE 2018 Speaker Diarization Challenge
15:00 - 16:40 Diego Castan, Mitchell McLaren, Mahesh Kumar Nandwana
- AE.5** ODESSA at Albayzin Speaker Diarization Challenge 2018
15:00 - 16:40 Jose Patino, Héctor Delgado, Ruiqing Yin, Hervé Bredin, Claude Barras, Nicholas Evans
- AE.6** EML Submission to Albayzin 2018 Speaker Diarization Challenge
15:00 - 16:40 Omid Ghahabi, Volker Fischer
- AE.7** In-domain Adaptation Solutions for the RTVE 2018 Diarization Challenge
15:00 - 16:40 Ignacio Viñals, Pablo Gimeno, Alfonso Ortega, Antonio Miguel, Eduardo Lleida
- AE.8** DNN-based Embeddings for Speaker Diarization in the AuDIA-S-UAM System for the Albayzin 2018 IberSPEECH-RTVE Evaluation
15:00 - 16:40 Alicia Lozano-Diez, Beltran Labrador, Diego de Benito, Pablo Ramirez, Doro-teo T. Toledano
- AE.9** CENATAV Voice-Group Systems for Albayzin 2018 Speaker Diarization Evaluation Campaign
15:00 - 16:40 Edward L. Campbell, Gabriel Hernandez, José R. Calvo de Lara
- AE.10** The Intelligent Voice System for the IberSPEECH-RTVE 2018 Speaker Diarization Challenge
15:00 - 16:40 Abbas Khosravani, Cornelius Glackin, Nazim Dugan, Gérard Chollet, Nigel Cannings
- AE.11** JHU Diarization System Description
15:00 - 16:40 Zili Huang, L. Paola García-Perera, Jesús Villalba, Daniel Povey, Najim Dehak



Search on Speech Challenge

- AE.12** GTM-IRLab Systems for Albayzin 2018 Search on Speech Evaluation
15:00 - 16:40 Paula López Otero, Laura Docío-Fernández
- AE.13** AUDIAS-CEU: A Language-independent approach for the Query-by-Example Spoken Term Detection task of the Search on Speech ALBAYZIN 2018 evaluation
15:00 - 16:40 Maria Cabello, Doroteo T. Toledano, Javier Tejedor
- AE.14** GTTS-EHU Systems for the Albayzin 2018 Search on Speech Evaluation
15:00 - 16:40 Luis J. Rodríguez-Fuentes, Mikel Peñagarikano, Amparo Varona, Germán Bordel
- AE.15** Cenatav Voice Group System for Albayzin 2018 Search on Speech Evaluation
15:00 - 16:40 Ana R. Montalvo, Jose M. Ramirez, Alejandro Roble, Jose R. Calvo

Speech to Text Challenge

- AE.16** MLLP-UPV and RWTH Aachen Spanish ASR Systems for the IberSpeech-RTVE 2018 Speech-to-Text Transcription Challenge
15:00 - 16:40 Javier Jorge, Adrià Martínez-Villaronga, Pavel Golik, Adrià Giménez, Joan Albert Silvestre-Cerdà, Patrick Doetsch, Vicent Andreu Císcar, Hermann Ney, Alfons Juan, Albert Sanchis
- AE.17** Limecraft Flow - workflows for story editing, subtitling and archiving
15:00 - 16:40 Victor Garcia, Nuria Sanchez, Angus Knights, Maarten Verwaest
- AE.18** Exploring Open-Source Deep Learning ASR for Speech-to-Text TV program transcription
15:00 - 16:40 Juan M. Perero-Codosero, Javier Antón-Martín, Daniel Tapias Merino, Eduardo López-Gonzalo, Luis A. Hernández-Gómez
- AE.19** The Vicomtech-PRHLT Speech Transcription Systems for the IberSPEECH-RTVE 2018 Speech to Text Transcription Challenge
15:00 - 16:40 Haritz Arzelus, Aitor Alvarez, Conrad Bernath, Eneritz García, Emilio Granell, Carlos David Martinez Hinarejos
- AE.20** Intelligent Voice ASR system for Iberspeech 2018 Speech to Text Transcription Challenge
15:00 - 16:40 Nazim Dugan, Cornelius Glackin, Gérard Chollet, Nigel Cannings
- AE.21** The GTM-UVIGO System for Albayzin 2018 Speech-to-Text Evaluation
15:00 - 16:40 Laura Docío-Fernández, Carmen García-Mateo
- AE.22** University of the Basque Country (GTTS@EHU) System for IberSPEECH-RTVE 2018 Speech to Text Transcription Challenge
15:00 - 16:40 Mikel Penagarikano, Amparo Varona, Luis J. Rodríguez-Fuentes, Germán Bordel

Text & NLP Applications

Friday, 23 November 2018, 09:00 – 10:40

Chair: José F. Quesada, Universidad de Sevilla

- O5.1** Topic coherence analysis for the classification of Alzheimer's disease
09:00 - 09:20 Anna Pompili, Alberto Abad, David Martins de Matos, Isabel Pavão Martins
- O5.2** Building a global dictionary for semantic technologies
09:20 - 09:40 Iklódi Eszter, Gábor Recski, Gábor Borbély, Maria Jose Castro-Bleda
- O5.3** TransDic, a public domain tool for the generation of phonetic dictionaries
09:40 - 10:00 in standard and dialectal Spanish and Catalan
Juan-María Garrido, Marta Codina, Kimber Fodge
- O5.4** Wide Residual Networks 1D for Automatic Text Punctuation
10:00 - 10:20 Jorge Llombart, Antonio Miguel, Alfonso Ortega, Eduardo Lleida
- O5.5** End-to-End Multi-Level Dialog Act Recognition
10:20 - 10:40 Eugénio Ribeiro, Ricardo Ribeiro, David Martins de Matos

Keynote 3

Friday, 23 November 2018, 11:00 – 12:00

Chair: Carlos Segura, Telefónica Research

- KN3** Automatic Question Answering: Problem Solved?
11:00 - 12:00 Lluís Màrquez

Round Table

Friday, 23 November 2018, 12:00 – 13:00

Chair: Marta R. Costa-Jussà, Universitat Politècnica de Catalunya

- RT** Panel discussion on Speech technologies: Industry and Academy
12:00 - 13:00 Marta R. Costa-Jussà