

Guest Editorial

Special Issue on Iberian Languages

Iberian languages (henceforth IL) are amongst the most widely spoken languages in the world. Nowadays, 628 million people on virtually all continents have Spanish, Portuguese, Catalan, Basque, Galician, etc. as their official language. This widespread usage is also accompanied by a growing technical use of these languages. Spanish and Portuguese rank third and eighth in terms of the number of web users (122 and 58 million users, respectively), whereas Portuguese ranks second in terms of the fastest growing web usage.¹

Consequently, important speech research centers and companies, both public and private, are focusing their interest on those languages. This effort has resulted in novel and generic approaches applicable to any language, as well as in the optimization of existing techniques or systems. It is worth highlighting that the community working on speech science and technology in IL speaking countries has already reached world-class level in many areas and has continuously increased in size in the last 15 years.

Speech technology proposed in the context of a non-Iberian language (e.g., English) may not be directly applicable to IL. All linguistic and paralinguistic dimensions, from phonetics to pragmatics, are amongst the features that certainly distinguish IL from others considered in speech science and technology research. As a result, original work and optimization of existing techniques and systems may be necessary in many areas of Iberian spoken language research.

The purpose of this Special Issue is to present recent progress and significant advances in all areas of speech science and technology research in the context of IL. We invited submissions addressing topics specific to IL and/or issues raised by analyses of spoken data that shed light on speech science and linguistic theories regarding these languages. The target was not to have submissions describing research which deals with the application of standard techniques to IL data, but rather research presenting relevant optimization of current technology and systems, and work exploring specific features of IL spoken corpora.

This call for papers originated a fairly significant number of submissions (26) from Spain, Portugal, Brazil, Chile, Cuba, and other non-IL countries. This issue includes only

12 papers. The range of topics of the current set of manuscripts spans all over speech science (prosody, production), speech technology (synthesis, recognition, language/accents and speaker verification), and spoken language systems (understanding, dialogue, translation, spoken term detection). This special issue does not concern only Spanish and Portuguese. In fact, research in Basque and Galician is also covered, as shown for instance by our first two papers.

González et al. address specific features of Galician that may have implications in the development of speech technology applications, namely for text-to-speech systems. The paper discusses phonetic features such as the handling of vocal contact and the determination of mid-vowel openness. It places special emphasis on the handling of clitics and verbs, noting the high interrelation between phonetics and grammatical information, and also addressing the task of morphosyntactic disambiguation.

Navas et al. describe a prosodic study for Basque, also in the framework of text-to-speech. Basque is an agglutinative and inflected language and POS features, widely used for other languages, are not enough to accurately predict the insertion of breaks in the text. Other morpho-syntactic features, like grammatical case and information about syntactic phrases have also been taken into account.

Prosody is also the topic of the paper by Martínez-Castilla and Peppé who propose a Spanish prosody assessment procedure adapted from an English one (Profiling Elements of Prosodic Systems – Children: PEPS-C). The paper describes the scope, principles and methods of the test and the modifications (other than lexical translation) that were required to produce a Spanish procedure.

A third paper on prosody by Meireles and Barbosa presents a prosody-related articulatory study (EMMA) showing the speech rate influence on the variation from antepenultimate stress words into penultimate stress words in Brazilian Portuguese. Their results benefit areas in speech technology directly concerned with speech variability at the word level. Moreover, once the variability of Brazilian Portuguese speech gestures due to speech rate change is well known, the results presented in the paper will contribute to realistic articulatory synthesis of Brazilian Portuguese.

The paper by Martins et al. has also potential implications for articulatory synthesis of European Portuguese.

¹ <http://www.internetworldstats.com/stats.htm> – visited in May 2008.

The paper describes a detailed analysis of the majority of the sounds of this language based on Magnetic Resonance Images. Some European Portuguese distinctive characteristics, such as nasality, have been addressed in more detail. Also coarticulation in stops and fricatives was investigated.

The fields of speaker and language verification are represented by two papers. The paper by Yoma et al. proposes an unsupervised, intra-speaker variability compensation (ISVC) method based on the Gestalt theory of perception to address the problem of limited enrolling data and noise robustness in text-dependent speaker verification (SV) tasks. Gestalt theory has recently been applied to other pattern recognition problems, but it has not been exhaustively applied to speech technology yet.

Rouas et al. describe a language/accent verification system for Portuguese, that explores different type of properties: acoustic, phonotactic and prosodic. The language verification system was trained for 10 languages. The variety identification task covers both European and Brazilian Portuguese, as well as Portuguese spoken in five African countries.

The paper by Tejedor et al. proposes the direct use of graphemes for acoustic modelling in a Spanish keyword spotting or spoken term detection system. This proposal is expected to work particularly well for languages such as Spanish, where despite the letter-to-sound mapping being very regular, the correspondence is not one-to-one, and there will be benefits from avoiding hard decisions at early stages of processing. The authors compare three approaches for Spanish keyword spotting or spoken term detection, and within each of these, they compare acoustic modelling based on phone and grapheme units.

Spoken dialogue systems are represented by a single paper by Martínez Hinarejos et al. The paper describes a statistical framework for a Spanish spoken dialogue corpus. Two statistical models based on the maximum likelihood assumption are presented, and two main applications of these models on a Spanish dialogue corpus are shown: labelling and decoding. The labelling application is useful for annotating new dialogue corpora. The decoding application is useful for implementing dialogue strategies in dialogue systems.

The final group of papers addresses machine translation, either from speech to sign language (for Spanish), from speech to speech (Spanish-to-Basque), or from text to text (English-to-Spanish). San-Segundo et al. describe the design, implementation and first user evaluation in a real application of a spoken language to sign language translation system in Spanish. The translation system is composed of a speech recognizer (for decoding the spoken utterance into a word sequence), a natural language translator (for converting a word sequence into a sequence of signs belonging to the sign language for deaf people), and a 3D avatar animation module (for playing back the hand movements). Two proposals for natural language translation have been evaluated: a rule-based translation module and a phrase-based statistical translation module.

Perez et al. describes the development of a text and speech translation system from Spanish to Basque using

finite-state transducers, paying special attention to the addition of linguistic knowledge. Two methods to cope with both linguistics and statistics are proposed. The first one performs a morphological analysis in an attempt to benefit from atomic meaningful units when it comes to rendering the meaning from one language to the other. The second approach aims at clustering words according to their syntactic role and uses such phrases as translation unit.

Our last paper, by de Gispert and Marino, presents an analysis of the impact of morphology derivation on *N*-gram-based Statistical Machine Translation (SMT) models. This analysis is carried out over a English to Spanish (a morphology-rich language) translation system. The authors show that verb form morphological richness greatly weakens the standard statistical models, proposing a posterior morphology classification by defining a simple set of features and applying machine learning techniques. In addition, the authors propose a simple technique to deal with Spanish enclitic pronouns.

This special issue is one of the first initiatives proposed by the recently created SIG-IL (ISCA Special Interest Group on Iberian Languages),² whose goal is precisely to promote research activities on IL, to sponsor and/or organize meetings, workshops and other events on related topics, and to make speech corpora publicly available by promoting joint evaluation efforts. Furthermore, the SIG-IL is also strongly committed to encouraging world-class research within its community in order to contribute with new ideas to the field of speech science and technology.

Many people contributed to make this special issue possible. Our first words of thanks are for the Editors in Chief of *Speech Communication*, who believed in our goals and motivation: Julia Hirschberg (former EiC), and Jean-Luc Gauvain. We would also like to thank Mary Lynn van Dijk (Journal Manager), for her invaluable assistance.

Our final words of thanks go to the more than 70 reviewers from all over the world who helped improving our work through many criticisms and suggestions, and to all the authors for their contributions and for helping promoting speech science and technology for Iberian Languages.

Isabel Trancoso
INESC-ID/IST, Portugal

Nestor Becerra-Yoma
Universidad de Chile, Chile

Plínio Barbosa
Univ. of Campinas, Brazil

Rubén San-Segundo
Universidad Politécnica de Madrid, Spain

Kuldip Paliwal
Griffith University, Australia

² www.il-sig.org.