

Note from the Editor

Special issue on Arabic Natural Language Processing and Speech Recognition: a study of algorithms, resources, tools, techniques, and commercial applications

Amy Neustein¹

© Springer Science+Business Media New York 2016

This special issue of the Journal consists of nearly two dozen articles that address both methodological approaches to Arabic Natural Language Processing and Automatic Speech Recognition as well as pilot-tested applications that are of commercial value. In keeping with prior special issues of this Journal, published under the auspices of the current Editor-in-Chief, this issue will consider both text-based (written) natural language communications as well as spoken language understanding as two sides of the same coin when studying Natural Language Processing. In addition, as reflected in the title, some of the key components of automatic speech recognition as they pertain to Arabic speech, such as speaker identification, speech synthesis, identifying regional accents, and evaluating natural language dialogue systems, are addressed using highly novel techniques and approaches. Given that the integrated study of both Natural Language Processing and Automatic Speech Recognition has proven to be the up and coming trend in the speech technology world today, as demonstrated at the recent Mobile Voice Conference, sponsored by AVIOS and William Meisel, President of TMA Associates, this issue well captures the zeitgeist by showing the leitmotif of the speech technology sciences as the integration of speech recognition with natural language understanding and generation.

Volume 19, No. 2 brings together the work of approximately sixty (60) contributing authors from leading universities in the U.S., Europe, Asia, Africa, and the Middle East. The guest editors are eminent researchers in Arabic Automatic Speech Recognition and Arabic Natural

Language Processing. But, even more important, their espousal of a weltanschauung of making software tools and resources free, open and available to developers and users is especially commendable, because it is this kind of commitment to both scholarship and open sourcing of language tools that will bring Arabic Natural Language Processing and Arabic Automatic Speech Recognition to the fore. The anthology contained in this volume traverses many areas. The authors have gone through a rigorous peer review process, with an acceptance rate of not more than 50 % of total submissions, wherein most articles published herein have been subjected to at least two phases of review and revision.

Considered among the collection are some very intriguing and novel ways of addressing the language processing components from stemming to morphological analysis to Part of Speech (POS) tagging, to named entity recognition, to acoustic modeling, to continuous speech recognition, etc., as they pertain to the Arabic language and dialects, with particular attention paid to the specific requirements of Arabic language that rely heavily on diacritics to determine the meaning of a word. But most important, the papers contained in this special issue show how new tools in Arabic Natural Language Processing and in Arabic Automatic Speech Recognition efficaciously bridge the gap between Arabic and other languages so that Arabic can enjoy the full advantages of Natural Language Processing techniques and Automatic Speech Recognition tools (and machine learning algorithms) as other languages do. For example, the fine paper on machine learning algorithms for data mining tasks shows how these advances can aid non-English speakers in making use of tools and services that are required in order to work with the massive amount of knowledge available on the Internet. Or, the paper on acoustic echo cancellation using adaptive filtering algorithms for Quranic accents

✉ Amy Neustein
amy.neustein@verizon.net

¹ Linguistic Technology Systems, Inc., Fort Lee, USA

shows how to better understand Arabic speech in noisy environments. Or, the paper which studies a very large diacritized corpus of training data demonstrates how morpho-syntactic analysis can be done correctly in spite of the absence of short vowels in Arabic text. These papers, like many others in this collection, serve as beacons of light for those who aspire, as we all should, to marshal the tools and resources of speech technology to benefit all languages in the universe as equitably and fairly as possible.

Lastly, this special issue resonates some very exciting, forward-looking technologies, bringing into the discussion a study of novel HPSG-driven open platform/open sourced

runtimes, and the world of machine-learned chatbots that are explored, dissected, and analyzed with great attention to detail. All in all, the papers in this collection show the assiduity, meticulousness, and commitment of its authors. But even more commendable, is the unflagging determination of its guest editors, Dr. Mohammad A. M. Abushariah and Dr. Bassam H. Hammo, to shepherd the authors through nearly a two-year long process to bring this special issue to fruition. Every Editor-in-Chief should have guest editors as competent and committed as those who participated in this special issue of IJST.