

EDITORS' INTRODUCTORY NOTE

In 1978, “Revue Roumaine de Linguistique” and “Cahiers de Linguistique Théorique et Appliquée” devoted all their issues, under the editorship of Alexandru Rosetti and Sanda Golopenția-Eretescu and the authorship of several linguists, to the presentation of the state of art in Romanian linguistics. Solomon Marcus contributed there with the article “Mathematical and Computational Linguistics and Poetics”, where the bibliography counted hundreds of entries belonging to Romanian authors, while the final list of authors who quoted and used their works is very long. After 1978, Romanian contributions continued to be in the attention of many authors, be they linguists, computer scientists, mathematicians or from other fields.

In the last 30 years, the diversification of the journals including research papers in the field made difficult to observe and evaluate the Romanian contribution. It is enough however to recall that Romanian authors published in the most important journals such as *Computational Linguistics*¹ (USA) and *Linguistics and Philosophy*² (USA) and there are Romanian names in the *Handbook of Computational Linguistics* published a few years ago by Oxford University Press³.

This issue brings into attention the results obtained by some researchers working at the University of Bucharest and some other Romanian universities, as well as some authors working in some research institutes. There is no claim to cover the totality of Romanian research in the field. We hope that many of the articles included in this volume will convince the reader of the need to bridge the computational and the linguistic approach; we hope to improve in this way the interaction between linguistics and computational linguistics.

There are in this issue some articles proposing a contemporary reading of some old Romanian contributions by Erika Nistor and Eliza Roman. Other authors who published in the sixties and the seventies of the past century, such as Minerva Bocșa from Timișoara University and Paul Schweiger, from Cluj-Napoca University, are not considered, but they deserve to be recalled. We hope to pay

¹ S. Marcus, C. Martin-Vide, G. Păun, “Contextual grammars as generative models of natural languages”, *Computational Linguistics*, XXIV, 2, 1998, 245–274.

² M. Kudlek, C. Martin-Vide, A. Mateescu, V. Mitrana, “Contexts and concepts of mildly context sensitivity”, *Linguistics and Philosophy*, XXVI, 6, 2003, 703–725.

³ R. Mitkov (ed.), *Handbook of Computational Linguistics*, Oxford, Oxford University Press, 2002.

attention to them in a next occasion. This activity of a contemporary reading of some old papers should continue.

A special mention deserves Professor Grigore C. Moisil, born hundred years ago, who is the father of Romanian Computer Science (born in 1957, with the first Romanian electronic computer) and who by his “mechanical grammar of Romanian” started the Romanian activity in the field and was the guide of many of us, particularly of the first Romanian activity concerning the automatic translation from English into Romanian; see, in this respect, the algorithm of translation English-Romanian proposed by Erika Nistor.

This issue presents a bridge between the first steps in Romanian computational linguistics made by researchers like Erika Nistor and Eliza Roman and the contemporary approaches in the field.

Almost half a century ago, Erika Nistor and Eliza Roman approached topics in the field called now information retrieval, by studies regarding the library organization. Some aspects in this respect are presented in E. Ionescu’s paper “Research in Information Retrieval in Romania, in the Period ’60-’80”. Another domain in which the two authors contributed was the text summarization. As one can see in the paper “On some Romanian Text Summarization Contributions”, F. Hristea and M. Popescu highlight the first attempts of these authors in applying Luhn’s algorithm for automatic creation of abstracts from Romanian texts.

Since Nistor and Roman’s time, computational linguistics has very much diversified and gone deeply, but Romanian contributions did not cease to exist. Nowadays there are a lot of Romanian scholars working in this field. Studies of some of them, few ones, are presented in this volume, which covers linguistic sub-fields such as lexicology (dictionaries and lexical ontologies), syntax, grammar, quantitative aspects of syllabification, phonology, prosody, and especially semantics.

One paper approaches methods regarding Wordnet ontologies. “Towards Building a WordNet Noun Ontology”, paper signed by G.A. Miller and F. Hristea, treats how to make automatically the distinction between words usually understood as referring to classes and words usually understood as referring to particular individuals and places.

A large number of papers are devoted to semantics. The Word Sense Disambiguation (WSD) problem, which is of large interest in the scientific community at the moment, finds room here in two contributions. D. Tătar proposes an algorithm which combines Yarowsky’s principle and the Naive Bayes Classifier method in order to determine which of the senses of an ambiguous word is preferred in a particular use of the word. How WSD methods are used in text classification one can see in the paper “Représentation de textes à l’aide d’étiquettes sémantiques dans le cadre de la classification automatique”. Its authors, C. Ignat and F. Rousselot, use them for handling polisemy in an algorithm for document representation.

L. Dinu's paper "On the Quantitative and Formal Aspects of the Romanian Syllables" is a sample of mathematical linguistics applied to the investigation of quantitative aspects of the syllable. The paper proposes a model of the written syllable, using Marcus contextual grammars.

Another domain well represented in this volume is the phonology, mainly connected to speech researches. M. Giurgiu's paper, "Experimental Results on Prosody for Romanian Text to Speech Synthesis", presents several experimental results on the study of the factors that affect the prosody in Romanian. The aim of the paper is to create the basis of a computational model for prosody description in a system for Text-To-Speech synthesis. L. Peev and F. Şerban, in their turn, present in the paper "Linguistics Premises for Designing a Text-to-Speech Synthesizer" the results of a project which made possible the study and elaboration of the formalized model of the Romanian phonology. Finally, M. Dinu in his paper "Un modèle tridimensionnel du système consonantique du roumain contemporain" proposes a method, borrowed from the codes theory, to the purpose of evaluating the similarities and differences between phonological units of the actual Romanian.

Various aspects of Romanian and foreign computational linguistics are reviewed in a separate section of the issue by M. Popescu, V. Barbu Mititelu, S. Marcus, E. Ionescu and L. Dinu.

*Ana-Maria Barbu
Solomon Marcus*