Towards a Computational Grammar of Spanish

Julio Sergio Santana Sepulveda,
Nora La Serna,
Luis Alberto Pineda Cortés

In this paper a research progress report on the construction of a computational grammar for Spanish is presented. This grammar is currently developed in the context of the DIME program (Diálogos Inteligentes Multimodales en Español) for modeling multimodal dialogues in which the system plays the role of an assistant for kitchen design. A basic grammar for common linguistic phenomena based on a version of GPSG is presented. An introductory discussion about grammatical features specific for Spanish, which are usually thought of as problematic for unification formalisms, like word order and clitics, is also presented. The strategy to enrich the grammar based on a corpus collected for the application domain is also sketched. Finally, an overview of the current implementation is presented.

Julio Sergio Santana Sepúlveda is a Research Associate of the Department of Computer Science at IIMAS, UNAM. He can be reached at checo@verde.iimas.unam.mx.

Nora La Serna is a Research Associate of the Department of Computer Science at IIMAS, UNAM. She can be reached at nora@leibniz.unam.mx.

Luis Pineda is a Titular Investigator and Head of Computer Science at the Institute for Applied Mathematics and Systems (IIMAS) of the National University of México (UNAM), AP. 20-726 ADMON. No. 20. Del. Alvaro Obregon 01000 Mexico, D.F. He is the author of more than 30 papers in computer graphics, intelligent CAD systems, knowledge representation, diagrammatic reasoning and computational linguistics. National Investigator (SNI) since 1993. He can be reached at luis@leibniz.iimas.unam.mx, see http://leibniz.iimas.unam.mx/~luis.

This work is been developed under partial support of Conacyt/NSF bilateral program for the development of computer science in a collaboration with the University of Rochester and ITESM, Campus Morelos, Conacyt grant 400316-5-C092A. We express special thanks to Prof. James Allen for valuable discussions.