

A Computational Model of the Spanish Clitic System

Luis A. Pineda, Ivan V. Meza

Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas (IIMAS)
Universidad Nacional Autónoma de México (UNAM)
{luis,ivanvladimir}@leibniz.iimas.unam.mx

Abstract. In this paper a computational model of the Spanish clitic system is presented. In this model clitic pronouns receive a dual analysis in which enclitics are considered inflexions while most proclitics are considered independent lexical units, hence proper clitics. The model covers the analysis of simple periphrases that, in addition to auxiliary and modals, have a single content verb (e.g. *puede comérselo, se lo ha querido comer*) and also the analysis of complex periphrases with more than one content verb (e.g. *le hubiera visto comérsela, se la hubiera visto comer*). The model introduces three operations on clitics: cancellation, composition and subsumption, and is formalized in Head-driven Phrase Structured Grammar; the standard machinery of this theory is extended with one combination scheme, the head-proclitic rule, and one principle, the clitic principle, that is satisfied by Spanish clitic sentences. A computational implementation of the theory with the Linguistic Knowledge Building (LKB) tool is also reported.

1 Introduction

Intuitively, a clitic is an unstressed particle that is attracted to a stressed word, its phonological host, and the resulting object is perceived as lexical unit¹; unlike inflexions and derivations, that are assembled with their stems at the morpho-lexical level of linguistic representation, clitics are combined with their host at the syntactic level. According to Zwicky and Pullum (1983, pp- 503):

“...word-clitic combinability is largely governed by SYNTACTIC considerations. The conditions governing the combinability of stems with affixes are of quite a different sort: they are MORPHOLOGICAL and/or LEXICAL in character, being concerned with the substructure of a finite set of words”

However, it is not always clear what is the linguistic level of representation for a given particle; in order to make this distinction Zwicky and Pullum (*ibid.*) advanced a number of criteria that we summarize as follows: (1) inflexions attach to words of specific syntactic categories while clitics do not exhibit this restriction, so clitics can attach to words of different categories and they often do so, (2) the combination host-clitic is very regular while inflexions show exceptions, (3) the meaning of clitic-host

¹ See, for instance, the introduction of Nevis (1991).