The syntactic evolution of French as seen in the MCVF corpus

Anthony Kroch and Beatrice Santorini
University of Pennsylvania

The MCVF corpus (Modéliser le changement: Les voies du français (MCVF) / Modeling Change: The Paths of French) was constructed between 2005 and 2010 with funding from the CRSH/SSHRC (Conseil de recherches en sciences humans du Canada / Social Sciences and Humanities Research Council of Canada) under the direction of France Martineau of the University of Ottawa, with extensive participation by our team at the University of Pennsylvania. The project produced a syntactically annotated (parsed) corpus of 1 million words of running text from 23 different sources. Since the end of that project, we have ourselves produced a supplementary corpus of 200,000 words from 14 additional sources at Penn. These corpora were annotated using a modified version of the annotation scheme developed for the series of Penn Parsed Corpora of Historical English, a scheme which was designed specifically to facilitate the diachronic investigation of word order patterns and changes.

Since its construction, the MCVF corpus has been used for research by a number of diachronic syntacticians, including ourselves. In our own work we have tended, as specialists in Germanic, to emphasize the comparative syntax of historical French and English, hoping thereby to stimulate a new field of corpus-based comparative diachronic syntax. In examining the evolution of medieval French and English, we have thus far discovered striking similarities and interesting differences in two domains: the well-known increasingly post-verbal position of objects and the declining frequency of so-called "verb-second" word order in the two languages. In our presentation we will show how these two diachronic processes manifest themselves as gradual changes in relative frequencies in the MCVF and Penn supplement corpora of French and make suggestions regarding the grammatical reanalyses that underlie the quantitative patterns that we have found. A brief comparison with the history of English will then be in order.