# Complex Predicates are Multi-word Expressions

**Martha Palmer**
Department of Linguistics
University of Colorado at Boulder
295 UCB
Boulder, Colorado 80309-029, USA
`Martha.Palmer@colorado.edu`

## Abstract

Practitioners of English Natural Language Processing often feel fortunate because their tokens are clearly marked by spaces on either side. However, the spaces can be quite deceptive, since they ignore the boundaries of multi-word expressions, such as noun-noun compounds, verb particle constructions, light verb constructions and constructions from Construction Grammar, e.g., caused-motion constructions and resultatives. Correctly identifying and handling these types of expressions can be quite challenging, even from the viewpoint of manual annotation. This talk will review the pervasive nature of these constructions, touching on Arabic and Hindi as well as English. Using several illustrative examples from newswire and medical informatics, current best practices for annotation and automatic identification will be described, with an emphasis on contributions from predicate argument structures.

## About the Speaker

Martha Palmer is a Professor of Linguistics and Computer Science, and a Fellow of the Institute of Cognitive Science at the University of Colorado. Her current research is aimed at building domain-independent and language independent techniques for semantic interpretation based on linguistically annotated data, such as Proposition Banks. She has been the PI on NSF, NIH and DARPA projects for linguistic annotation (syntax, semantics and pragmatics) of English, Chinese, Korean, Arabic and Hindi. She has been a member of the Advisory Committee for the DARPA TIDES program, Chair of SIGLEX, Chair of SIGHAN, a past President of the Association for Computational Linguistics, and is a Co-Editor of JNLE and of LiLT and is on the CL Editorial Board. She received her Ph.D. in Artificial Intelligence from the University of Edinburgh in 1985.