

Corpus Linguistics as a Tool in Legal Interpretation

Lawrence M. Solan - Don Forchelli Professor of Law, Brooklyn Law School.

[with Tammy gales, Associate Professor of Linguistics, Hofstra University].

In this paper we set out to explore conditions in which the use of large linguistic corpora can be optimally employed by judges and others tasked with construing authoritative legal documents. Linguistic corpora, sometimes containing billions of words, are a source of information about the distribution of language usage. Thus, corpora and the tools for using them are most likely to assist in addressing legal issues just when the law considers the distribution of language usage to be legally relevant. As Thomas Lee and Stephen Mouritsen have so ably demonstrated in earlier work, corpus analysis is especially helpful when the legal standard for construction is the ordinary meaning of the document's terms. We argue here that four issues should be addressed before determining that corpus analysis is likely to be maximally convincing. First, the legal issue before the court must be about the distribution of linguistic facts. Second, the court must decide what makes an interpretation "ordinary." Third, if one wishes to search a corpus to glean the ordinary meaning of a term, one must decide in advance what to search. Fourth, there different issues why a particular meaning might present a weak showing in a corpus search and these need to be understood. Each of these issues are described and discussed.