

Linguistics in an Age of Engineering

Christopher Manning

Christopher Manning, Linguistics F12
University of Sydney NSW 2006, AUSTRALIA

Abstract

In the early days of engineering models of language, a popular intuition was that the easiest way to improve performance was to remove linguists. This was not an unreasonable, or even necessarily wrong intuition, and reflected the fixation by linguists of the day on linguistic structures to the exclusion of consideration of statistical dependencies between lexical or phrasal units or the non-categorical nature of real language use. But while the linguistics-blind approach showed promising initial beginnings, and the sheer brute-force nature of the approach led to the development and popularization of important statistical methods for dealing with sparse data, nevertheless, looking to the future, I'll argue that now the right way to improve engineering models is through the incorporation of more linguistics. But what kind of linguistics? I want to sketch how formal approaches to linguistics should reshape themselves from this brush with engineering: in a quantitative direction which is actually rather more compatible with much of the 'fuzzier' linguistic work in such topics as grammaticalization and sociolinguistics, though realized with more mathematical precision.

