Does a Computational Linguist have to be a Linguist?

Martin Kay
Stanford University and Universität des Saarlandes
Stanford, CA, USA and Saarbrücken, Germany
`mjkay@stanford.edu` and `kay@coli.uni-saarland.de`

Invited Speaker Abstract

Early computational linguists supplied much of theoretical basis that the ALPAC report said was needed for research on the practical problem of machine translation. The result of their efforts turned out to be more fundamental in that it provided a general theoretical basis for the study of language use as a process, giving rise eventually to constraint-based grammatical formalisms for syntax, finite-state approaches to morphology and phonology, and a host of models how speakers might assemble sentences, and hearers take them apart. Recently, an entirely new enterprise, based on machine learning and big data, has sprung on the scene and challenged the ALPAC committee's finding that linguistic processing must have a firm basis in linguistic theory. In this talk, I will show that the long-term development of linguistic processing requires linguistic theory, sophisticated statistical manipulation of big data, and a third component which is not linguistic at all.