Keynote Speech: How do you spell that? A journey through word representations

Greg Kondrak
Department of Computing Science
University of Alberta
gkondrak@ualberta.ca

Abstract

Languages are made up of words, which in turn consist of smaller units such as letters, phonemes, morphemes and syllables. Words exist independently of writing, as abstract entities shared among the speakers of a language. Those abstract entities have various representations, which in turn may have different realizations. Orthographic forms, phonetic transcriptions, alternative transliterations, and even sound-wave spectrograms are all related by referring to the same abstract word and they all convey information about its pronunciation. In this talk, I will discuss the lessons learned and insights gained from a number of research projects related to the transliteration task in which I participated.

Bio.

Greg Kondrak is an Associate Professor at the Department of Computing Science, University of Alberta. He has served four times as an Area Chair for the ACL conference. He is currently the Secretary of the Special Interest Group on Computational Morphology and Phonology (SIGMORPHON). His research is focused on natural language processing at the sub-word level, including transliteration, grapheme-to-phoneme conversion, and phonetic similarity. Since 2007, he has co-authored five conference papers devoted to transliteration, and participated in all five editions of the NEWS shared task on transliteration. Altogether, he has published over 70 papers, most of which can be found in the ACL Anthology.