Abstract
There is growing interest in the automatic extraction of opinions, emotions, and sentiments in text (subjectivity), to provide tools and support for various natural language processing applications. Most of the research to date has focused on English, which is mainly explained by the availability of resources for subjectivity analysis, such as lexicons and manually labeled corpora. In this talk, I will describe methods to automatically generate resources for subjectivity analysis for a new target language by leveraging on the resources and tools available for English, which in many cases took years of work to complete. Specifically, I will try to provide answers to the following questions. First, can we derive a subjectivity lexicon for a new language using an existing English lexicon and a bilingual dictionary? Second, can we derive subjectivity-annotated corpora in a new language using existing subjectivity analysis tools for English and parallel corpora? Finally, third, can we build tools for subjectivity analysis for a new target language by relying on these automatically generated resources?

Bio:
Rada Mihalcea is an Associate Professor in the Department of Computer Science and Engineering. Her research interests are in computational linguistics, with a focus on lexical semantics, graph-based algorithms for natural language processing, and multilingual natural language processing. She is currently involved in a number of research projects, including word sense disambiguation, monolingual and cross-lingual semantic similarity, automatic keyword extraction and text summarization, emotion and sentiment analysis, and computational humor. Her research has been funded by the National Science Foundation, National Endowment for the Humanities, Google, and the State of Texas. She is the recipient of a National Science Foundation CAREER award (2008) and a Presidential Early Career Award for Scientists and Engineers (2009).