Junior Faculty Research Award in Science and Engineering
Research Statement
Andrew Rosenberg

The *way* people produce language conveys information about what a person is communicating as much as *what* they are saying. In some situations, a person's tone of voice is the only way to know they are being sarcastic, or they are angry. In other contexts, speech production is used to indicate qualities like the contrastive focus as in "I like **Italian** food" perhaps implying "not French food" or "I like Italian **food**" as opposed to "Italian cars". The speech production qualities that lead to these different interpretations, changes in pitch and loudness, rhythmic choices, emphasis and phrasing, vocal stress, are collectively referred to as "prosody".

My research is concerned with computational approaches to representing, understanding, analyzing and synthesizing prosody in speech. The goal of this work is two fold: it helps us understand how humans communicate and it enables machines to communicate in more human-like ways. Improving analysis of prosody allows for speech interfaces to better understand human communication. Improving synthesis of prosody allows for more natural and intelligible sounding machine generated speech. This work is fundamentally interdisciplinary, drawing from and contributing to Computer Science and Linguistics.

This work has let to a number of innovations in computational approaches in speech prosody. I distribute, update and maintain an open source tool the performs prosodic analysis on speech data. This tool, AuToBI, is used by academic and industrial research groups around the world and includes a number of advances to the state-of-the-art. With collaborators at Columbia and MIT, I have started a web-based data repository for prosodically annotated speech. This has brought the research community together to encourage the sharing of data, resources and tools related to our work. Additionally, I worked with the speech team at IBM to improve speech synthesis quality on Watson for the IBM Jeopardy! Challenge. While I started working with IBM for this event, my relationship with this team has continued, and led to continued work on speech synthesis.

I have received a number of external grants over the last few years. These projects involve the application of prosodic analysis. It has been particularly satisfying to explore the use of this information in a diverse set of contexts including Keyword Search, Information Extraction, Recognizing Deception and Charisma, Detecting Sarcasm, Conflict and Laughter.

I was recently awarded an NSF CAREER award titled "More than Words: Advancing Prosodic Analysis". This proposal has a particularly important place in my research. The work that we will do in project is directly concerns improving the core representations and understanding of prosodic information in speech spoken in English and other languages. This will significantly advance the state-of-the-art in understanding how people use prosody to communicate.

Computers effectively acting as conversational agents represents one of the great promises of artificial intelligence. This is an element of science fiction that we are slowly assimilating into our lives, if fitfully. While there is a long way to go, understanding prosody is fundamental to understanding human communication.