Su-Youn Yoon
NLP and Speech Group, Educational Testing Service

Title: Development of an Automated Speech Scoring System Based on Linguistics

Abstract:

In this presentation, I will provide a brief overview of SpeechRaterSM, an automated scoring system for non-native speakers' speech developed at Educational Testing Service. This system automatically assesses a non-native speaker's proficiency level using non-constrained spontaneous speech.

The overall architecture of SpeechRaterSM is as follows: for a given spoken response, SpeechRaterSM performs speech recognition and yields a hypothesized sentence. Next, it computes a wide range of features which cover the fluency, pronunciation and prosody, grammar and vocabulary, and content and topic development aspects of speech. Finally, a proficiency score is generated using a regression model; the model is trained using data scored by experts before testing.

The performance of the system is heavily influenced by the features, and both linguistic and applied linguistic research has significantly contributed the development of the features. First, in the conceptualization of features, previous linguistic research is reviewed, and significant features which can estimate the stage of the acquisition are selected. The actual implementation of features is also based on linguistic research. In this presentation, an actual example will be provided, along with a discussion of how Linguistics contributes to the realization of an automated scoring system.

Automated scoring research can also be beneficial to linguistic research. I will discuss how researches in the automated scoring can contribute linguistic research, and the important points which should be considered for this purpose.