For many ESL (English as a Second Language) and EFL (English as a Foreign Language) students, interacting with computerized applications is an integral part of their learning experience. NLP-based language models can be a valuable tool in assisting teachers and students alike by providing prompt feedback on certain aspects of language. Recent research in this area has been towards the development of grammar correction applications specifically targeting learners of English.

In this talk, I will first present work on determiner correction using a statistical model trained on well-formed texts written by native English speakers. However, such an approach is limited in that constructing a large enough error-annotated corpus to support a statistical approach is time-consuming and labor intensive. I will therefore describe a newer study focusing on preposition errors in which error correction models are trained exclusively on an error-annotated corpus produced by ESL learners. We address the design issues and the logistical problems that arise from the partially annotated nature of our data set.

Na-Rae Han is currently Lecturer in the Linguistics department and Director of Robert Henderson Language Media Center, which promotes use of technology in language instruction, at the University of Pittsburgh. She received her Ph.D. in Linguistics from the University of Pennsylvania in 2006 after completing her M.S.E. in Computer and Information Science there. She has previously worked as a researcher in the Automated Scoring and Natural Language Processing Group of Educational Testing Service (ETS) in Princeton, NJ, and as a research professor at Korea University in Seoul, Korea.