In this study, a probabilistic computational model is used to identify shifts in conversational style. “Conversational style” in particular, as distinct from “style” more broadly, is defined as the aspects of discourse that speakers use to mark involvement with their fellow participants in talk, following Tannen’s (2005) classification of “high involvement” versus “high considerateness,” that is, higher and lower levels of explicit marking of involvement in discourse. The computational model allows us to consider style shifting as something that is emergent from and co-constructed in discourse, rather than as individual speakers’ statistical preference or dis-preference of a given linguistic variant.

Style is typically considered as intraspeaker variation, that is, the linguistic choices that an individual social actor makes in a given context to express group affiliation or otherwise conduct identity work (Bell 1984, Bell & Johnson 1997, Coupland 2007). In contrast, this study takes an approach inspired by Fred Erickson’s observation that “talking with another person ... is like climbing a tree that climbs back” (cited in Tannen 1989: 13); in other words, participants in discourse are constantly orienting and reorienting themselves to one another. Therefore, no attempt will be made to categorize one or another speaker’s individual stylistic choices. Instead, style is viewed as a property of discourse rather than of speakers, a moving target that is jointly constructed and subject to continuous change.

A hidden Markov model was chosen as the computational representation of this styleshifting hypothesis. To implement a Markov model, a number of discrete states are identified — in this case, the high-considerateness and high-involvement conversational styles — together with a transition probability distribution that specifies how likely the model is to move from any state into any other state, or to stay in the same state, as the process moves forward. In a hidden Markov model, the state is not directly observable; rather, each state could give a number of different possible outputs, which are predicted according to an emission probability distribution. In this study, the length, turn construction type, and speaker identity are modeled as outputs of the Markov process. Based on observations of the outputs, the model allows us to infer the underlying state that is most likely to have generated each turn in the transcript. These results can then be interpreted using both quantitative (linear regression) and qualitative (discourse analysis) methodologies.

This model creates an opportunity for a quantitative study of conversational style that cannot be achieved using traditional quantitative sociolinguistic methods. Rather than taking either turn length or turn construction type as a proxy for conversational style, it represents style through a complex probabilistic model that accounts for the ways speakers construct turns at talk throughout the discourse. Preliminary results suggest that the computational model can identify coherent stylistic episodes in discourse that feature the expected linguistic markers of high and low involvement. More broadly, this result suggests that computational methods can be successfully applied to the study of talk in interaction.