Exploiting a Swedish multimodal corpus of spontaneous speech: gesture movement profiles and question typology

David House

Department of Speech, Music and Hearing, School of Computer Science and Communication, KTH, Stockholm, Sweden.

Abstract

In this seminar, I will report on the collection of a multimodal corpus of spontaneous dialogues and on two different areas of exploitation involving motion capture data (gesture movement dialogue profiles) and audio recordings (question typology). More than 120 half-hour sessions of spontaneous dialogue have been recorded within the Swedish database project, Spontal. The point of departure for the project was the fact that both vocal signals and gesture involving the face and body are important in everyday, face-to-face communicative interaction and that there is a great need for data in which we can more precisely measure these. The participants were all native speakers of Swedish and balanced (1) for gender, (2) as to whether the interlocutors were of opposing gender and (3) as to whether they knew each other or not. The dialogue topic was free, but after 20 minutes the participants were asked to open a box which contained objects whose identity or function was not immediately obvious. All recordings are comprised of high-quality audio, high-definition video and a motion capture system using infra-red cameras and reflective markers to capture body and head gestures.

One area of exploitation involves using motion capture data to automatically produce gesture movement profiles for the dialogues. Profiles of selected dialogues have been created by plotting average marker movement per speaker over time for the entire duration of the dialogues. The profiles can be correlated to important properties of the individual dialogues such as speaker activity, dialogue dominance alternation, and movement synchrony between the participants. The graphic displays of the profiles can illustrate these properties and provide interesting and useful measurements which can contribute to new ways of analyzing dialogue behavior.

A second area of exploitation involves using the corpus to create a question typology and analyze variation in question intonation in spontaneous speech. I will review current work aimed at extracting a database of 600 questions from the corpus complete with categorization and prosodic descriptions. I will report on coding and annotation of question typology and present results concerning prosodic correlates related to question type in the 600 questions. Current results indicate a prosodically salient distinction between the two categories termed, in our typology, forward and backward looking questions.