Social scenes are common in many video games and movies. To realistically recreate such scenes in computer graphics, it is critical to animate the conversing human characters. A social scene typically contains foreground characters and background characters. While foreground characters are the focus of the scene and thus are carefully created by artists, the sole purpose of background characters is to render the atmosphere and to add realism to the environment. However, with traditional content creation tools, the artists often need to spend as much time animating the background characters as the foreground characters, despite that the exact behavior of the background characters is not critical to the experience -- they only need behave naturally so that they do not detract the audience. In this talk, I will introduce a system I have developed that helps artists to rapidly generate the animations for the talking characters in the background. The system automatically generates the body motions for two talking characters from an audio recording of a conversation. In order to produce natural looking animations, the system must ensure that the character's body motions are smooth and are synchronized with the rhythm of the audio. For example, a speaker often uses hand gestures as they are stating an important point, and a listener will nod to acknowledge what the speaker is saying.

My talk will start with how we captured/recorded a dataset of real conversations and how we studied the statistics of the synchrony between body motion and audio signals from the database, then focus on the key algorithm of our system that generates novel motion sequences from input audios based on the captured data. I will conclude my talk with results from a user study that demonstrates the effectiveness of our system.