In this paper, we take the position that some high level human behaviors can be discovered in an unsupervised manner from low level gestures. We first locate "interesting" user behavior by analyzing data collected from accelerometers and microphones placed on the user's wrists and elbows. Using this data, we apply a simple intensity analysis and differencing algorithm to segment low level actions. We then propose to discover behavioral scripts from successive clustering of the time-ordered data through analysis of both the audio and accelerometer signals.