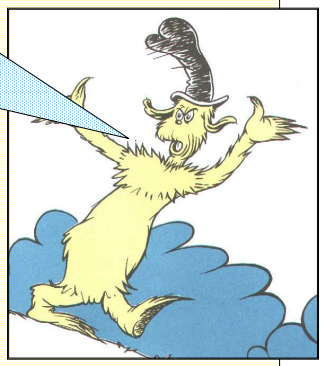
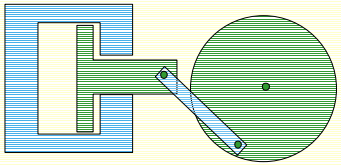
Natural Gesture in Descriptive Monologues

Jacob Eisenstein and Randall Davis
MIT Artificial Intelligence Laboratory

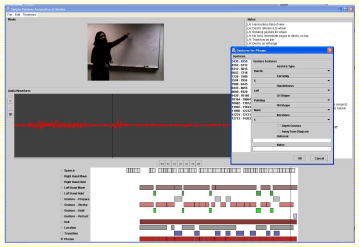


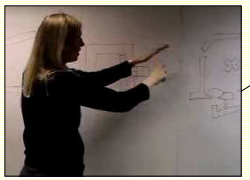
Experimental Design

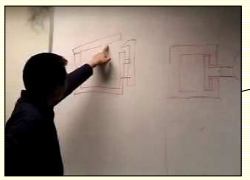
Nine participants were shown simulations of mechanical devices and asked to describe them. Each participant described three devices; the descriptions ranged from 15 to 90 seconds in length.

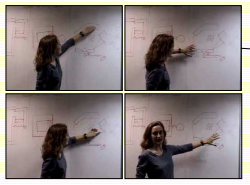


Their speech and gestures were transcribed and annotated. A total of 574 gesture phrases were recorded.









Gestures Refer to the Diagram

96% of gestures made reference to the diagram. This suggests that it is acceptable to require gestures to be performed on the display surface, as in touch-screens or pen-based user interfaces.

Many Speakers Use Both Hands

Of the nine speakers, six used both hands while gesturing. Two-handed gestures comprised 14% of the total, and speakers also alternated between hands while making single-handed gestures. Input modalities for gestural user interfaces should support two-handed gestures.

Deixis is More Frequent

The presence of a display significantly increases the proportion of deictic gestures. Studies without displays found ten times as many iconic as deictic gestures. In this study, we found 222 deictics and 338 iconics. The display enabled speakers to substitute deictic references for iconics, which are more difficult to recognize and interpret.

Gesture Units are Longer

Previous research has viewed gestures as isolated atoms, rather than as pieces of a larger structure. But the presence of a display significantly increases the length of gesture units. These gesture units are reliable predictors of discourse segment boundaries.