

**Figure 1:** Clips from the same video.  
**[Top]** Only a single frame is needed to describe this clip.  
**[Bottom]** Many frames are needed to describe the expressions in this clip.

