

are being recorded (for ethical issues that need to be addressed before recording speakers, see the chapters by Rice and Newman, Chapters 18 and 19 below).

6.5 Basic Annotation of Gesture

As discussed in §6.2, researchers studying gesture investigate the coordination of speech and gesture to answer questions concerning the forms and functions of gestural deployment, to determine the role of gesture in communication, and to explore what gesture reveals about cognition. A basic and minimal annotation of gesture consists of marking the periods of time in which the hands are gesturing. This type of basic annotation enables the analyst to locate the stretches in the video which are relevant for gesture analysis, and can be used to assess the time speakers spend gesturing. To identify gesture in the movement stream, the basic structure of gesture has to be taken into account.

p. 163 Gestural movement can be segmented into four distinct phases (Kita, van Gijn, and van der Hulst 1998; McNeill 1992; Kendon 2004a): preparation, stroke, hold, and retraction. Consider a speaker describing how a ball rolled down a hill by making a spiral movement downwards with the extended index finger. First the hand is lifted from the lap to chest height. This phase is the preparation: the hand is brought into the starting position for the stroke. The stroke then consists of the hand moving downward in a spiralling movement. The stroke is the only obligatory phase of a ↵ gesture, and it can be characterized as the most articulated phase during which the most force is exerted. A stroke can be preceded or followed by a hold (pre-stroke or post-stroke), a static phase in which the hand is held still in a position. The phase during which the hand is moved back into rest position is the retraction. A stroke can be static or dynamic, with meaning being expressed through a static configuration (e.g. when size or extension is indicated) or through motion (rotating movement downward). Note that this established segmentation scheme mixes formal (hold) and functional categories (preparation, stroke, retraction). To circumvent this problem one, can distinguish in the annotation between static and dynamic strokes.

The phases can be distinguished based on changes in the direction of movement and changes in the speed of motion (acceleration or deceleration) and changes in the articulation of the motion (e.g. tension of the hand, directedness of motion). The segmentation of the movement into discrete phases can be accomplished by looking at the video frame by frame to see when the points of change occur. When the starting and the end points are determined, the stretch can be tagged and categorized (for a discussion of how to determining moments of transition systematically, see Seyfeddinipur 2006).

The time speakers spend gesturing is the basic gesture unit (Kendon 2004a). This is when the speaker moves his hands away from a rest position (e.g. lap) or an activity (e.g. cutting, smoking) and begins to gesticulate until the hands return to a different activity again.

For multimedia annotation, software like ELAN² or ANVIL³ should be used. These annotation tools allow setting up multiple tiers for different levels of annotation as in a musical score. In each tier a different type of annotation can be provided which is time aligned with the video (see Fig. 6.3). Relevant phases in which the speaker gestures can be marked based on a frame-by-frame inspection of the video. Once a period is identified, it can be tagged and replayed.

For each hand, one tier should be set up, and the times during which gesturing occurs should be marked for each hand. During such a gesture unit, the speaker may use multiple gestures, and the next level of annotation for a more detailed analysis entails a breakdown of the stream of gestural movements into gesture strokes, the semantic nucleus of gestures. This level of annotation can be pursued when the researcher has a clearer idea of which gestures are of particular interest.