

entering, exiting). Because of the relatively static nature of picture books, retelling may or may not be successful in eliciting gesture.

More focused gesture elicitation tools have been developed for motion events (e.g. the Tomato man movie: Özyürek et al. 2008) and for spatial reference (see Majid, Chapter 2 above). However, as mentioned above, the applicability of the stimuli depends on the familiarity of a given community with such materials and tasks, and so some stimuli and tasks may not be appropriate for the cultural context, or for use with certain members of the community.

For those who are more interested in undertaking gesture studies, some examples in the literature of gesture research by field linguists are Enfield (2009), Green (2009), Haviland (1993), Kita et al. (2001), Kita and Essegbey (2001), Le Guen (2011), Sherzer (1991), Wilkins (1999). See also Goodwin (1986) and Kendon (2004a).

6.4.2 How to record

Gesture, just like speech, is a fleeting phenomenon. Because of its ephemeral nature, certain aspects are impossible to catch by casual observation alone. The temporal unfolding over time of the gesture form, its trajectory of motion, and its synchrony with specific elements in speech can only be captured systematically by video. Video recording, however, can be problematic for ethical reasons (Rice, Chapter 18 below), is technically challenging (Margetts and Margetts, Chapter 1 above), and speakers may alter their behaviour in the presence of a recording device.

Video need not only be shot by the fieldworker, but may in fact be better collected by trained community members. Having community members video record interactions can also provide insights into what they consider important in the interactions (for uses of video in documentation, see Ashmore 2008).

In any case, there are some guidelines for recording which best facilitate later analyses of gesture speech interaction. An ideal recording frames the conversational situation with the speakers' upper bodies, with at least an arm-length space around their head. If possible, in framing the span of the interlocutors' space, it is best to overcompensate so that the maximum breadth of possible movement (assuming fixed position) is taken into account. The camera should be positioned at 9 o'clock or frontal, and angled slightly down from eye level of the speakers. With such a framing of the speakers, we can observe the hand shapes and movements along with facial expressions, gaze, and eyebrow and head movements. This is important because these all tend to form elements in an orchestrated whole.

An obvious problem for this setup is that the way speakers are oriented towards each other depends on cultural as well as situational restrictions. In many places it is not very natural to have two or more people side by side facing the camera rather than each other, and any artificial composition placing the speakers in culturally unnatural positions should be avoided. So we always have to find a compromise, such as framing only the main speaker in this way while the other speakers' faces may not be fully visible.

Ideally, the social space created by the interlocutors (including all interlocutors and not just a single speaker) should be recorded. This is because the verbal and gestural behaviour of speakers is contingent on the behaviour of their interlocutors, where they are located, whether they look or not, whether they backchannel their understanding or not, etc. (Schegloff 1984; Streeck 1994; Goodwin 1986). Speakers in some cultures are known to orient their gestures and the direction of movement depending on the location of the interlocutors (Özyürek 2002). In addition, for Canadian English-speaking college students, it was found that speakers' gestures are more pronounced and better articulated when the speaker thinks the recipient does not have prior knowledge of what the speaker is talking about: when the information to be