The NHI and the emergence of video-based multimodal studies of social interaction. Response to Wendy Leeds-Hurwitz and Adam Kendon

Lorenza Mondada

5.6 Introduction

Initiated in 1955–1956 by an interdisciplinary group interested in the analysis of sound films, with the aim of developing a cybernetic vision of communication and pathology in psychiatry and a new model of social interaction more generally, the Natural History of an Interview (NHI) project constitutes a groundbreaking event in many respects. In particular, it offers an exemplary case for reflecting on the beginnings of video-based multimodal studies of social interaction and, more generally, on the nexus between technological innovations and analytical advances. While the NHI project has been clearly outlined and discussed by Leeds-Hurwitz and Kendon in this volume (see also Leeds-Hurwitz 1987), this response to their chapter aims at reflecting on what constitutes its novelty, by relating it historically and conceptually to contemporary endeavors in video analysis, with a special focus on the early work of Charles and Marjorie Goodwin, who are central contemporary references for multimodal approaches to video-based analyses of social interaction.

Exploring the new possibilities offered by sound film for studying human interaction, the NHI was the first project that made substantial use of film technologies for capturing the details of talk and embodied movements in their fine-grained temporality and precise coordination. This text reflects on the articulation between recording technologies, transcription practices and new models of social interaction (Kendon 1979; Erickson 2004; 2011), with a particular focus

on the use of video in conversation analysis (Goodwin 1993; 2018; Heath et al. 2010; Knoblauch et al. 2006; Mondada 2006; 2012; 2021). This articulation characterized the initial NHI discussions in 1955–1956 (§2) and was further developed as the project continued until 1968, especially in the work of Birdwhistell, Scheflen and Condon (§3). These scholars were also in contact with Goffman and the group which became a reference for micro-sociology, ethnography of communication and conversation analysis, within which future leading figures of present-day multimodal analysis, Charles and Marjorie Goodwin, were trained (§4). This response shows the enduring influence of the NHI for establishing a web of questions that are still central today, and for creating a network of scholars who opened up possible future pathways, and were crucial for the training of contemporary scholars in video studies and multimodal analysis.

5.7 Some innovations of the Natural History of an Interview project

The NHI was a revolutionary project for the study of social interaction in many ways. I insist here on the implications of the historical fact that it is the first project to take full advantage of the analytical potential of film with synchronized sound. As described in detail by Leeds-Hurwitz (1987) and Leeds-Hurwitz & Kendon 2021 [this volume], the NHI started as an interdisciplinary project involving psychiatry (Fromm-Reichman, Brosin), linguistics (Hockett, McQuown), kinesics (Birdwhistell), and anthropology/cybernetics (Bateson). The NHI members collectively engaged in the study of film materials provided by Bateson, most prominently what has been called the "Doris film" (film GB-SU-005 by Bateson and Myers) – in which Doris talks with Bateson in her living room, in presence of her young son Billy playing around, while Myers, the cameraman, films them. The NHI's exclusive focus on the film materials, rather than more generally on group communication or family therapy, privileged the *filmed* interaction, treating the film as preserving the naturalistic setting so documented (vs. the filmed practices that documented it, see Engelke 2021 [this volume] and Watter 2017: 47). This enabled a full appreciation of the analytical potentials of film for analysis (§2.1), the development of a sophisticated way of transcribing the action in the film (§2.2) and the elaboration of analyses sensitive to the details of social interaction (§2.3).

5.7.1 A naturalistic approach of filmed materials

Since the development of "chronophotography" and motion picture technology at the end of the 19th century (Tosi 2005), with Marey (1896) and Muybridge (1887), moving images have been used continuously in the scientific study of human behavior in the form of body movements (Mondada 2021), especially in anthropology (Hockings 1975; Ruby 2000). But for a long time, the film technologies that were available for social scientists in practice were limited to silent film. Even if commercial cinema introduced sound films in 1927, it was only after World War II that sound film became financially, materially, and practically accessible to the social sciences.

Bateson had experimented with film recordings when working with Mead in Bali already in the 1930s. Film, together with photography and fieldnotes, offered them a vivid way to document ordinary life. However, despite their richness, these materials were only loosely interconnected. Even if Mead's and Bateson's collaborative fieldwork was based on a division of labor in which she took notes about what people were saying while he was shooting silent films and photographs of what they were doing (Bateson & Mead 1942: 49–59, Jacknis 1988), this was not enough to provide for a fine-grained, precise analysis of the coordination of language and embodied action.

Later on, developing his interest in mental health with psychiatrists, Bateson became involved in the production of a series of 16mm films of therapy interviews, in a collaboration with Ruesch (a psychiatrist) and Kees (an experimental filmmaker) in Berkeley (Ruesch & Bateson 1951; Ruesch & Kees 1956) and later with Weakland and Haley at the Veterans Administration Hospital in Palo Alto (Bateson et al. 1956; Haley 1962). Kees played a crucial role in developing filming techniques for realistically capturing the ordered but fleeting patterns of communication (in "films that permit us to look at human beings as they actually are"; Ruesch & Kees 1956: 12). Film enabled them to make visible otherwise imperceptible details, small but crucial for revealing the mechanisms of mental life. This filmic approach was motivated – and at the same time made possible – by a conception of mental illness as generated by specific ways of communicating in groups like the family, as opposed to being located in individual pathologies. Mental illness was defined in terms of small recurrent behavioral patterns that could be captured precisely on film (Geoghegan 2017: 72-73). At the same time films were heavily edited for the purposes of communicating findings and to illustrate theories (Engelke 2021 [this volume]), rather than treated as raw materials serving as a basis for detailed analysis, as would be the case in the NHI.

The NHI films provided the basis for elaborating a *naturalistic* perspective on communication: "We start from a particular interview on a particular day between two identified persons in the presence of a child, a camera and a cameraman. Our primary data are the multitudinous details of vocal and bodily action recorded on this film. We call our treatment of such data a 'natural history' because a minimum of theory guided the collection of the data" (Bateson 1971: 6). Bateson's aim was the study of the natural history of human communication (Leeds-Hurwitz 2005). "Natural history" was a method, an approach, a perspective (McQuown 1971b; Scheflen 1971), which participants in the project opposed to the "experimental" approach. In the "natural history" approach, "the organism or group under study is maintained as far as possible under the customary conditions of living", while under the "experimental" method "the subjects are usually approached in the more-or-less specialized artificial conditions" (Brosin 1971: chap. 4, p. 50). Interestingly, a natural approach was also advocated, in the same period, in Sacks' lectures ("sociology can be a natural observational science"; 1989 [1966]: 211, see Lynch & Bogen 1994), and became a key concept in conversation analysis and ethnomethodology (Lynch 2002). The first video-based study by Charles Goodwin (1979) refers to "natural conversation" and this might not only be a reference to Sacks, but also to the NHI.

5.7.2 Transcribing sound film materials

The mere transcription of embodied behavior in the NHI project was not a novelty per se. There is a long history of attempts to annotate body movement. For example, there have long been various notation systems for dance and choreography, and these served as an inspiration to the participants in the NHI. Although dance notation and movement notation have different objectives, they share an interest in analytical precision and timing. Dance notation is intended for instructing future dancers, and aims to reproduce deliberate body forms set out by a choreographer, whereas movement notation captures fleeting movements as they happen/happened, including their uniqueness and contingencies. But both address the order of the moving assemblage, its gestaltic character and its unfolding in time—which film would make it possible to capture.

Influential dance notations emerged as early as the 16th century, such as Arbeau's *Orchésographie* (1588) and Feuillet's *Choréographie* (1700). Feuillet proposed a sophisticated system in which orientations of legs, arms, torso, head etc., with a specific focus on path of steps on the floor, are precisely described by combining elementary graphic symbols. Further systems of notation emerged later on, either based on the trajectories of dance steps on the floor, like Feuillet's, or

rather on time, in linear representations of several lines synchronized with the musical score. This is the case of Saint-Léon's *Sténochoréographie* (1852), in which each movement of each dancer is written above the corresponding note along a continuous temporal line. Saint-Léon is closer to the issues the NHI scholars faced when trying to associate the movements of the body and the progressivity of talk on a timeline. Moreover, Saint-Léon inspired late 19th century notations, such as Stepanov (1892) and Zorn (1987), who were searching for a systematic combination of signs to represent movements as body sentences. These notations might have inspired Birdwhistell too, even if he refers to other predecessors (Craighead 1942, a master thesis in dance, and Pollenz 1949, an anthropology study of ethnic dances), and dismisses the most famous notation at his time, Laban's *Schrifttanz* (also called *kinetography*; 1928, Hutchinson Guest 1970), as too complicated and as mainly used in industry rather than research (1970: 181). He developed his own notation: a temporally organized plurilinear score in which visual signs referring to a systematic inventory of positions and movements of body parts are used. Birdwhistell's notation is inspired by the combinatorial vision of units in structural linguistics and is perfectly fitted with Hockett's and McQuown's linguistic and paralinguistic notations of speech, which complete the NHI transcripts.

Transcripts constitute half of the NHI final report, which runs to 982 pages (McQuown 1971c). Its originality concerns several levels. First, it integrates "two systems of transcription, graphic symbolic, and alphabetic" (1971c: Foreword, p. 7), referring to phono-recording and kine-recording (1971c: Chapter 6, p. 2). More generally, it includes in a multi-layered system of paralinguistic (a term referring to voice qualities and non-linguistic vocalizations; Trager 1958; 1971), prosodic, phonetic and alphabetic notations for linguistics (developed by Hockett and McQuown), and multiple kinesic annotations for body movements (developed by Birdwhistell). These notations further distinguish various levels of granularity, metaphorically referred to by Bateson as the "ladder of Gestalten from the most microscopic particles of vocalization towards the most macroscopic units of speech" (1971: 24). Second, these layers of annotation are spatially arranged as multiple lines constituting a musical score, referring to a timeline that is not expressed in portions of seconds but in frame numbers. This time-based score enables the integration of very different annotations from various disciplines in a unique and cumulative multilayered object, the transcript. Third, these annotations were based on different recorded sources (chapter 6): the interview had been continuously recorded on audio tape, which was listened to "short section by short section" for detailed linguistic and sound annotations, whereas the sound film – which was interrupted every 3 ½ minutes because the 100ft

magazine had to be replaced, occasioning long gaps in the recording (Bateson 1971: 13ff.) – was watched again and again in a silent mode for the annotation of body behavior. Thus, different perceptive practices lay at the foundation of the transcripts, within different modes of listening and viewing the records (under the process that the participants in the NHI called "soaking", McQuown 1971a: 5, Leeds-Hurwitz & Kendon 2021 [this volume]) for more micro or macro details, supported by technologies such as the audio player, the moviola, and the slow-motion projector.

5.7.3 Filming and transcribing for the analysis of social interaction

The transcript is inspired by and in turn makes possible a rigorous conception of social interaction. In this respect, Bateson presents the NHI project as wanting "to see every detail of word, vocalization, and bodily movement as playing its part in determining the ongoing stream of words and bodily movements which is the interchange between the persons" (1971: 9). Likewise, Birdwhistell quotes Bateson's claim that "everything which occurs in a social interaction is meaningful in the sense of being part of the interchange as well as non-accidental" (1971: 1) – a position that is echoed by Sacks saying that "there is order at all points" (1992: 484, lecture 33, Spring 1966), which is considered a fundamental assumption of conversation analysis.

The availability of recording and transcription technologies enabled the NHI to develop a detailed, analytic and holistic view of social interaction as a complex web of relationships, organized in patterns in which all details might matter and which have a predictable order. Their order is conceived in the framework of structural linguistics, in terms of identification of minimal units combined at multiple levels of complexity. Although contemporary multimodal analyses rely on different theoretical principles, some of these issues are still central, such as the definition of the relevant units of social interaction, which continue to be crucially embedded in transcription choices (Ochs 1979).

5.8 Continuing the NHI: the work of Birdwhistell, Scheflen, and Condon

Initiated in 1955–1956, the NHI project continued until 1968, ending with an unfinished manuscript deposited at the library of the University of Chicago (McQuown 1971c; see Leeds-Hurwitz & Kendon 2021 [this volume] for a reconstruction of the (dis)continuities of these meetings). During these years, the participants in the project published their results individually – with all making use of

the film – and transcription-based approach developed in the project (Birdwhistell 1970; Brosin 1964; 1966; McQuown 1957) – and continued to work on ideas the NHI had made possible. Among these developments, I briefly consider further analytical usages of film and video technologies, and their consequences for future research on the multimodality of social interaction. These crystalized around the figures of Birdwhistell and Scheflen – who worked together at the Eastern Pennsylvania Psychiatric Institute (EPPI), the former from 1959 until he joined the University of Pennsylvania at the Annenberg School of Communication in 1969, and the latter from 1956 to 1967, when he moved to the Bronx State Hospital of New York. Condon collaborated with Scheflen and both are mentioned in later NHI chapters. Kendon – who later became a leading figure of gesture studies (1990; 2004) – came from the UK in the mid-1960s to work with Scheflen, and also collaborated with Condon. Scheflen is a key figure, considered as having made Birdwhistell's approach more explicit and systematic, and as having developed it for an overall comprehension of the structure of the encounter rather than only for a microscopic focus on single scenes (Kendon & Sigman 1996).

5.8.1 Developing film and video technologies

Sound film was crucial to making the NHI project possible, and the use of film and video continued to be refined during the 1960s. At the Eastern Pennsylvania Psychiatric Institute, Birdwhistell organized a film and video lab, with slow motion projectors for film and editing tables for video (Watter 2017: 52). In 1960 he hired a filmmaker, Van Vlack, who not only provided for professionally shot films, but also for their theoretical discussion (1965; 1966a; 1966b). He defined the scientific film as a "new type of motion picture […] produced under controlled, explicit conditions to produce a permanent record which may be repeatedly searched for the re-observation and re-analysis of the original ephemeral event" (1966a: 15–16). He also reflected on how technologies enable an analytical vision considering that "events have a regularity that can be isolated and described, so that they constitute predictable sub-patterns within the larger context" (1966b: 3). The identification of these patterns and searching for them depend on the use of slow-motion and stop-frame projectors (1966b: 5), which enables specific practices of perceiving these images, contrasting with practices of watching traditional movies: "observation of the film by trained observers. With repeated observation, scientists experienced in this technique can locate and abstract the patterns from their context. This is quite a different approach than our usual culture-bound passive acceptance of film" (1966b: 5). Practices of manipulating-for-looking also materialized in another device, the "B-roll frame number count

for every frame" (Van Vlack 1966b: 5). This technique involved duplicating the film to create a working copy (the B-roll), on which the references to frame numbers were printed in the upper central margin of the picture. This enabled the analyst to locate a scene within the film at a precise point (see Kendon 1979: 77). More generally, van Vlack shows that the use of moving images is not merely a matter of "watching" but more crucially a matter of equipping the audio-visual analytic perception of the observer.

At the Bronx State Hospital, Scheflen hired Schaeffer (who would write a dissertation on videotape techniques in anthropology, 1970, and who was also a musician) as field director of his Project *Human Communication*. The project included the continuous recordings of family life in six households over several weeks, with portable video and 16mm film cameras, which were remotely controlled. The data constitute a first impressive example of the use of several cameras for longitudinal filming sessions continuously made over several days, preceded by a careful ethnographic approach to securing acceptance of the filming on the part of the participants (Schaeffer 1975). The development of film as well as the emergence of video technologies made possible a further elaboration of the "naturalistic" way of filming, strongly articulated with fieldwork. The increasing portability of the camera equipment enabled the documentation of a diversity of social contexts.

However, even if video was becoming available, film was not superseded: both continued to be used together (Erickson 2011: 181, Schaeffer 1975; Scheflen et al. 1970). More flexible and easier to manipulate, video was used for initial filming sessions, and for rapid feedback to informants; however, film was of much better quality for a more detailed, high-resolution frame-by-frame analysis. Film projectors were of better quality than video projectors. Moreover, film could be projected in slow-motion; for example, with hand crank projectors, advanced by turning a handle, allowing frame-by-frame motion. Editing tables for video produced small and unfocused images, although they caused less wear and tear on the recordings (Kendon 1979: 77). For these reasons, both technologies were used at the same time – along with separate audio tape recording. Video progressively enabled researchers to make their own recordings, whereas the early uses of film relied on professional camera operators, who were also very often artists – like Kees with Bateson, Van Vlack with Birdwhistell and Schaeffer with Scheflen.

5.8.2 More transcriptions

Technologically supported manipulations of film/video produced an array of artifacts, diagrams, motion flow charts, scrolls containing transcripts organized

along horizontal timelines, like musical scores, and other forms of annotations on paper. Recorded by cameras and inspected with controllable projectors, events could be scrutinized for their real-time flow of details, units, boundaries of units, and recursive structures.

Consequently, a number of coding, transcribing, annotating, and charting conventions were proposed during this period. Linguistic, phonetic and paralinguistic transcription was elaborated by Hockett and others in a highly original book, experimenting with the material composition of the pages, integrating transcripts as a book within the book that could be leafed through independently of their analysis (Pittenger et al. 1960). Birdwhistell continued to work on his kinesic notations (1952; 1970). Condon proposed some careful notation of the temporality of embodied details (1970, see below). An alternative system was proposed by Hall (1963), founder of "proxemics" (1959; 1966), the study of spatial distribution of people in social activities, a factor that varies between ethnic groups. Hall's notation system includes not only distance, orientation and body arrangements of the participants, but also the wider sensorium (olfaction, touch) – a dimension that would be considered again much later on (Mondada 2019). Previously, Hall had written, with Trager, a handbook combining proxemic and paralinguistic analyses (1953). Hall, Trager, and Birdwhistell knew each other, after meeting at the Foreign Service Institute, a training facility for diplomats, where they were all teaching (Leeds-Hurwitz 1990; Kendon & Sigman 1996; Watter 2017: 38ff.).

5.8.3 The analysis of spatial and temporal details

Providing the necessary technological support for the collection of data, as well as for the inspection and exploitation of the collected data in transcripts, created the conditions for new forms of analysis that were able to document in detail phenomena that were previously only loosely described in ethnographic accounts or that even escaped notice. Birdwhistell repeatedly highlighted how many details can be seen in a body movement, how film/video enables the segmentation of movement into a multiplicity of units, and how their combination reveals ordered patterns (1970). Whereas Birdwhistell's aim of finding the "grammar" of the communicative system, built, by analogy with structural linguistics, on a hierarchy of units (kines, kinemorphs, kinesic sentences) seems to have failed (Birdwhistell 1970: 197–198), further work by Scheflen, Condon and Kendon explored these patterns, revealing more local but also more systematic orders. I briefly comment on two fundamental features made available by film/video: spatiality and temporality.

Space is the focus of Scheflen's work. Searching for the common denominator between various ways of interacting in therapy consultations, Scheflen proposed a hierarchy of units describing different levels of organization: the point, the position and the presentation (1964) – similar to linguistic units identified in structural linguistics (see Scheflen & Scheflen 1972: 46-47, fn.). The original contribution of Scheflen is to point out that these units, at all levels, refer to the territoriality of the body – and are visible, measurable, photographable and analyzable (Scheflen 1971). On a macro level, echoing Goffman's (1963) interest in forms of co-presence within space, and in distinct body assemblages for focused vs. unfocused interactions (see Scheflen & Scheflen 1972: 35-36 exemplifying these distinctions), Scheflen considered that the way bodies occupy space provides for a specific characterization of the event. On more micro levels, space intervenes in the orientation of fine embodied details, indexed by head movements, gestures, or even moves of the eyelid, working as "markers", such as indicating the end of a sentence (Scheflen 1964: 321, fig. 1; Scheflen & Scheflen 1972: 48ff.).

This primacy of spatiality enabled a view of human interactions focused on the entire body as it is mobilized – and constrained – in its ecology (from the architecture of the crowded Bronx households to the minimal space required for a body to gesticulate). This produced further research highlighting the "proxemic shifts" in an interaction (Erickson 1975), who obtained his PhD under Hall's supervision), the spatial dimension of "distant" vs "close" greetings (Kendon & Ferber 1973), and more generally the spatial "formations" (Kendon 1977) characterizing the relative positions of the participants engaged in an activity. Several decades later, these discussions would be influential for the study of interacting bodies considered in their entirety (Goodwin 2000; Mondada 2016; 2018), and in their ecology, enabling the treatment of the materiality surrounding them and used by them (artifacts, tools, objects, Heath 2012; Heath & Luff 2000; Goodwin & Gwyn 2003; Mondada 2019). They would also be influential for further work on "interactional space" (Mondada 2009) and mobility in interaction (Haddington et al. 2013).

Film also enabled the capturing of the dynamic temporality of these movements. Condon, who collaborated with Scheflen and participated in the late phase of the NHI, was particularly interested in the timing of speech and movements. He was famous for using a modified Bell and Howell time-motion analyzer to facilitate a manual scanning of the film frame-by-frame (1970). This made a detailed segmentation of speech and motion possible, by manually contrasting one segment with the previous one in order to detect isomorphisms and changes. Condon used speed cameras able to shoot 48 frames per second (vs. 24fs for

usual 16mm films), enabling him to double the precision of the sound/film analysis (see Condon 1976: 292, fig. 2 for an example of multilayered transcription using a 48fs-based segmentation). This provided a foundation for his description of synchrony: since different body parts can move independently, it is significant when they move together. By contrast, in pathological subjects, these relations are disarticulated (Condon & Ogston 1966). On this basis, Condon identified two forms of co-ordination: within the individual ("self-synchrony") and between different participants ("interactional synchrony", Condon 1976: 306, fig. 6), in which the participants adopt the same temporal structure for their embodied conduct.

This attention to the detailed timing of speech and body co-ordination anticipates other tools – such as computer-supported aligning software for transcription (e.g. ELAN) – which facilitate the creation of several lines of annotations and measures. They also anticipate transcription systems that are based on multiple superpositions of lines of annotations including the most diverse aspects of talk and the body (Mondada 2018). In addition, the focus on temporal details and their location within the emergent and dynamic flow of actions prefigures the interest not only in co-speech temporal regularities (Kendon 2004; Schegloff 1984) but also in more complex "multimodal Gestalts" (Mondada 2014), which consider holistically several types of resources, each of them characterized by a specific temporality, not mechanically synchronic but relevantly and reflexively adjusted one to another (Goodwin 2000).

5.9 The beginnings of multimodal conversation analysis: Charles and Marjorie Goodwin

At the end of the 1960s, Charles (Chuck) and Marjorie (Candy) Goodwin – whose work today serves as a model in video studies, multimodal analysis and conversation analysis, interconnected with linguistics, anthropology and communication – were beginning their PhDs at the University of Pennsylvania. Their emerging trajectory intersects with the legacy of the NHI, the work of Birdwhistell and colleagues, as well as the academic group around Goffman and Labov.

Chuck and Candy Goodwin repeatedly pointed to the importance of Goffman's and Labov's seminars at the Center for Urban Ethnography (newly founded in 1969, and which also funded Candy Goodwin's dissertation). Goffman arrived at the University of Pennsylvania in 1968, and Hymes was instrumental in bringing him there; Labov joined them from Columbia in 1971. Hymes and Goffman had both previously worked in Berkeley, where Goffman had collaborated with Garfinkel, Sacks and Schegloff (Schegloff 1992). Goffman had various connections

to Birdwhistell, who joined the University of Pennsylvania at the Annenberg School of Communication in 1969: he had been Birdwhistell's undergraduate student in Toronto (Leeds-Hurwitz & Sigman 2010), and both had participated in the Macy conferences on group processes, with Bateson and Mead from 1955 to 1958 (Winkin 1984). Although Goffman and Birdwhistell had different views on social interaction, they had an appreciation for each other's work. Goffman published Birdwhistell (1970) in his series at the University of Pennsylvania Press; Birdwhistell's famous analysis of the "cigarette scene" (1970: 227-240) was re-published in *Directions in Sociolinguistics*, a collective book edited by Gumperz and Hymes in 1972, in the same section as Garfinkel, Sacks, and Schegloff.

The fact that Goffman's group was caught up in the wake of the NHI can also be traced back to a letter Goffman sent to Hymes on October 26, 1967, proposing the organization of a workshop with McQuown, Bateson, Hall, Sommer, Ekman and "people from the Garfinkel school", mentioning Schegloff and Sudnow, as well as Garfinkel himself (correspondence Goffman/Hymes, Goffman archive). This project shows the converging interest on social interaction of scholars coming from the NHI project, Goffman's group, and Californian conversation analysts and ethnomethodologists.

It is in this context that Chuck and Candy Goodwin began their PhDs in 1969. Candy Goodwin was doing fieldwork with Afro-American children in working-class neighborhoods, under the guidance of Goffman, whose approach to social interaction inspired the project, and Labov, who was researching the grammar of Black English Vernacular. Chuck Goodwin wanted to work with scholars who had collaborated with Bateson: in 1969, he became the research assistant of Krippendorff, who was interested in cybernetics and soon became his PhD supervisor at the Annenberg School of Communication. He was also hired as a filmmaker at the Philadelphia Child Guidance Clinic (1971–1974) under the supervision of Haley, who had been Bateson's collaborator. In a way, Goodwin was joining the long line of filmmakers associated with protagonists of the NHI project. Furthermore, Chuck and Candy Goodwin attended Birdwhistell's classes in 1969, where they learned to use his kinesic annotations and analyses, and also saw numerous films shot by Bateson's, Birdwhistell's, Condon's and Scheflen's teams (such as the cigarette scene, Schaeffer's and Scheflen's films on the crowded households in the Bronx, and Bateson's and Kees' as well as Birdwhistell's and van Vlack's films shot in zoos). They were also in contact with Kendon, with whom they intensively examined film data (for example on greetings, Kendon & Ferber 1973). These connections show that even if the NHI did not directly inspire the work of Chuck and Candy Goodwin, they were still exposed to its intellectual spirit and representatives.

The work of Chuck and Candy Goodwin was at the crossroads of several emergent scientific programs. Goffman was influential for his approach of the interactional order, his conception not only of focused interaction, but also of the co-presence of bodies of not-yet-co-participants in space, intelligible thanks to their body-glosses (Goffman 1963; 1971). Although Goffman was critical of recordings and preferred participant observation, his vision of participation as an embodied phenomenon in its material and spatial ecology resonated with Scheflen's work on territoriality as well as Birdwhistell's interests in bodies in motion. Birdwhistell represented another way to approach bodies in interaction, not only through film but importantly through his sophisticated system of annotations and coding. For Chuck and Candy Goodwin, however, his emphasis on the exhaustive annotation of all possible details, within a notational system that integrated large *a priori* lists of pre-defined body-parts, strongly influenced by structural linguistics, contrasted with the approach to transcription they were learning from Gail Jefferson. Jefferson, who had just finished her PhD with Sacks at UCLA, had joined Labov at Penn in 1972, working for him in transcribing data and co-teaching with him a seminar in conversation analysis. She proposed a practice of transcription that was necessarily selective and related to the relevance of details along the emergent sequential organization of social interaction.

In this context, Chuck and Candy Goodwin – who had met in 1969 – developed a way of doing fieldwork collaboratively that was enhanced by the technologies at hand. Candy was strongly engaged in fieldwork and Chuck was developing professional film skills at the Child Guidance Clinic. Together they wanted to effect a shift from the therapeutic room-cum-TV studio toward ordinary social settings, indoors and outdoors – i.e. from arrangements of the participants *for* the camera view to an adjustment *of* the camera work to the filmed events. Confronted with the challenges of naturalistic settings they realized all the advantages represented by the mobile Portapak video camera newly produced by Sony (AV-3400/AVC-3400). Although very heavy and still relatively bulky (it actually required a wagon to be moved, and was almost impossible to carry over the shoulder), the Portapak enabled them to engage in the recording of everyday life, among family, friends, and neighbors, as well as encounters at the meat market. The Portapak made it possible to engage in new forms of naturalistic fieldwork-with-a-camera.

Presented in Goffman's and Labov's seminars, the analysis of these data turned out to be influenced by Jefferson, who guided Chuck's and Candy's Goodwin first steps into conversation analysis (they had already accessed Sacks' lectures thanks to Labov, but met Sacks and Schegloff only later, at the Summer Institute of Linguistics of 1973). In weekly meetings at their house with Jefferson and

Malcah Yaeger-Dror, a phonetician collaborating with Labov, they began to develop a conversation analytic approach to the videos they had collected.

Jefferson's contribution to the history of conversation analysis has been recognized in relation to her effective way of transcribing talk, her system of conventions still used in conversation analysis today (Jefferson 1983; 1985; 2004), and importantly, her analytical attention to the fine-grained temporality and sequentiality of talk – studied in relation to turn-taking (Sacks et al. 1974), overlapping talk as revealing the timing precision of turns-at-talk (Jefferson 1973; 1984; 1988), and the systematic order of repair (Jefferson 1972; 1974; 2017) among other topics. Jefferson's role mentoring Chuck and Candy's Goodwin early career reveals an underestimated aspect of her analytical eye: she embraced the study of video materials very early on, supporting Chuck's dissertation on gaze and embodiment in conversation (1981 [1977]). The intellectual role she played in relation to video has not been documented until now, and casts some new light on the importance of her contribution to the history of conversation analysis.

Jefferson transcribed the verbal aspects of some of Chuck Goodwin's data—for instance, the "cigarette utterance" fragment (so called perhaps in a kind of *clin d'oeil* to Birdwhistell's "cigarette scene"), which would later be published as Goodwin's first paper (1979). This intensive transcription work gave Jefferson a deep understanding of the data, which were discussed in their weekly sessions. More fundamentally, she was drawing their attention to details of gaze and body postures, and under her guidance they progressively discovered and learned to see these details. She also imagined practical ways of preparing the materials for analysis, such as making the suggestion of connecting the Portapak to a TV-monitor. This would enable them not only to watch the activities studied repeatedly (the Portapak camera had playback and slow-motion functions) but, more crucially, to stop the image in order to scrutinize its details.

On this basis, Jefferson developed a technique of tracing, consisting in drawing the silhouettes of the participants by putting a transparent plastic paper ("saran wrap") on the monitor. These drawings are very simple: they focus on the co-participants' body postures, and enable the tracking of their mutual involvement, gaze and manipulations of objects (for instance, cigarettes). Tracings are produced for an entire fragment of talk, tracking the moment-by-moment unfolding of postures: the verbal transcript and pauses are reproduced below each image, and their segmentation adjusts to the relevant shifts in postures and gaze (see Goodwin 1981 [1977]: 144–147 for an example). In the working transcripts, fragments of a few seconds are rendered by dozens of drawings. These tracings are in striking contrast to Birdwhistell's kinesic annotations: whereas the latter

are based on a textual representation, articulated in words and specifically pre-defined semiotic codes corresponding to body movements, Jefferson's notation privileges the holistic representation of the image drawn on the basis of the video footage, enabling one to catch at a glance the participants' reciprocal adjustments. In the NHI, the absence of images is noticeable. In Jefferson's tracings, the image is predominant – this connects her rather to the tradition of studies that made extensive use of photographic stills (Mead & Byers 1968; Ruesch & Kees 1956; Scheflen & Scheflen 1972). However, Jefferson's tracings are distinct in the sense that they enable understanding of the sequential organization of embodied talk, rather than just illustrating specific types of postures and typologies of spatial bodies arrangements.

This contrast between ways of re-presenting action refers to contrasting analytic approaches. Birdwhistell's coding approach was based on hierarchies and combinations of units inspired by structural linguistics, whereas Jefferson's approach was based on the local relevance of units shaped by the participants in the course of their situated emergence, assembled and adjusted by them in response to their recipient's actions. Likewise, in his dissertation, Chuck Goodwin (1981 [1977]: 25-27) reflects explicitly on the conception of units inspired by the structural vs. the sequential view: the former, associated with Duncan (1972) – also relying on Trager & Smith (1957 [1951]), for the suprasegmental phonemic transcription – is a "signal" model of turn-taking, in which some markers indicate that the speaker has come to the end of their utterance; the latter is an "anticipatory" model in which the speaker projects and the recipients anticipate what is to come, and mutually adjust to it, shaping the dynamic emergence of the utterance in an interactive way. Although recognizing Birdwhistell's and Scheflen's work (1981 [1977]: 27), and referring to Condon's (1981 [1977]: 28) notion of synchrony, Goodwin also shows that within a sequential perspective, the emergence of units, especially at transitions between one unit and the next, are made intelligible and anticipable thanks to differentiated projections achieved by distinct multimodal resources (1981 [1977]: 29).

Jefferson's mentoring of Chuck and Candy Goodwin, contains *in nuce* a multimodal sequential analysis of social interaction, which combines the "analytic mentality" (Schenkein 1978) generated by findings concerning the organization of turns-at-talk in interaction (Sacks et al. 1974), together with the analysis of video materials and embodied postures.

This intense analytic engagement with video data is not restricted to Jefferson's collaboration with Chuck and Candy Goodwin; Christian Heath reports very similar experiences: he also reports on the capacity of Jefferson to see things in the video data that nobody else had noticed, her holistic vision of sequential

organization including body postures, her skilled use of tracings on a diversity of data. Moreover, both Goodwin and Heath talk about Jefferson's generosity, not only offering ideas during data sessions but also providing extensive comments and annotations on their manuscripts (this resonates with the experience Goffman himself had with his manuscript on "response cries" 1978, reviewed by Jefferson in a detailed, eleven-page critical analysis, Bergmann & Drew 2015: 10-11). This contrasts with the fact that today Jefferson is above all recognized for her analyses of talk and for the excellence of her verbal transcripts, while her video analytic expertise is never acknowledged in the literature. Jefferson herself never published on video analysis. But her mentoring had a tremendous impact on the work of Chuck and Candy Goodwin as well as of Christian Heath, pioneers in the further development of multimodal analysis within ethnomethodology and conversation analysis.

In sum, at that time, Charles and Candy Goodwin were at the crossroads of various intellectual traditions. They were exposed to a variety of options – fieldwork à la Goffman, attention to speech patterns à la Labov, systematic coding of embodied communication à la Birdwhistell – which made Gail Jefferson's inspiration so valuable and innovative.

5.10 From NHI to contemporary multimodal analysis

This response has shown how the use of technologies for recording, listening/watching and transcribing sound films (and then videos) has shaped notation systems and analytical possibilities. It has shown the permanence of questions about how to adequately record human activities, how to transcribe/annotate/code a multiplicity of resources articulated in time, and how to analyze these resources in order to enrich our understanding of social interaction. It has demonstrated how it is possible to reconstruct a lineage from the NHI to the expanding work of Birdwhistell and his collaborators, and to the training of Goodwin, leading to contemporary multimodal analysis, by tracking the changing articulations between technologies and analytical issues.

Technologies evolve at a fast pace, but the conceptual constraints on adequate recordings for social interaction are relatively stable (cf. Kendon 1979; Heath et al. 2010; Mondada 2012). Despite other theoretical frameworks having replaced structural linguistics and cybernetics, the NHI project remains remarkably current in its holistic-and-analytical vision of social interaction and of the assembling of bodies in their communicative ecology. After a number of decades in which the transcription/annotation/coding of a smaller number of dimensions

was privileged, which enabled systematic analyses, there is again a scientific interest in the ordered complexity of the action of entire bodies in interaction. This is demonstrated by contemporary debates about social interaction and mobility, material and spatial environments, sensoriality, multiactivity, and coordination in silent activities. The study of these phenomena relies crucially on the reflections initiated by the NHI project on recording and transcribing for analysis.

Acknowledgments

References

Arbeau, Thoinot. 1588. *Orchésographie*. Paris.

Bateson, Gregory. 1971. Chapter 1: Communication. In Norman A. McQuown (ed.), *The natural history of an interview*. Chicago: Microfilm Collection of Manuscripts on Cultural Anthropology, University of Chicago, Joseph Regenstein Library, Department of Photoduplication.

Bateson, Gregory, Don D. Jackson, Jay Haley & John Weakland. 1956. Toward a theory of schizophrenia. *Behavioral Science* 1. 251–264.

Bateson, Gregory & Margaret Mead. 1942. *Balinese character: A photographic analysis*. New York: New York Academy of Sciences.

Bergmann, Jörg & Paul Drew. 2015. Introduction. In *Talking about troubles in conversation*. New York: Oxford University Press.

Birdwhistell, Ray L. 1952. *Introduction to kinesics: An annotation system for analysis of body motion and gesture*. Washington, D.C.: Departement of State, Foreign Service Institute.

Birdwhistell, Ray L. 1970. *Kinesics and context: Essays on body motion communication*. Philadelphia: University of Pennsylvania Press.

Birdwhistell, Ray L. 1971. Chapter 3: Body motion. In Norman A. McQuown (ed.), *The natural history of an interview*. Chicago: Microfilm Collection of Manuscripts on Cultural Anthropology, University of Chicago, Joseph Regenstein Library, Department of Photoduplication.

Brosin, Henry W. 1964. Studies in human communication in clinical settings using sound film and tape. *Wisconsin Medical Journal* 63. 503–506.

Brosin, Henry W. 1966. Linguistic-kinesic analysis using film and tape in a clinical setting. *American Journal of Psychiatry* 122(12S). 33–37.

Brosin, Henry W. 1971. Chapter 4: Implications for psychiatry. In Norman A. McQuown (ed.), *The natural history of an interview*. Chicago: Microfilm Collection of Manuscripts on Cultural Anthropology, University of Chicago, Joseph Regenstein Library, Department of Photoduplication.

Condon, William S. 1970. Method of micro-analysis of sound films of behavior. *Behavior Research Methods and Instrumentation* 2(2). 51–54.

Condon, William S. 1976. An analysis of behavioral organization. *Sign Language Studies* 13. 285–318.

Condon, William S. & William D. Ogston. 1966. Sound film analysis of normal and pathological behavior patterns. *Journal of Nervous and Mental Disease* 143(4). 338–347.

Craighead, Jean. 1942. *A system of notation for the modern dance*. Baton Rouge: University of Louisiana. (MA thesis).

Duncan, Starkey D., Jr. 1972. Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology* 23(2). 283–292.

Engelke, Henning. 2021. Perception, awareness, and film practice: A natural history of the "Doris Film". In James McElvenny & Andrea Ploder (eds.), *Holisms of communication: The early history of audio-visual sequence analysis*, 105–138. Berlin: Language Science Press. DOI: 10.5281/zenodo.5142286.

Erickson, Frederick. 1975. One function of proxemic shifts in face-to-face interaction. In Adam Kendon, Richard M. Harris & Mary R. Key (eds.), *The organization of behavior in face-to-face interaction*, 175–188. Berlin/New York: De Gruyter Mouton.

Erickson, Frederick. 2004. Demystifying data construction and analysis. *Anthropology & Education Quarterly* 35. 486–493.

Erickson, Frederick. 2011. Uses of video in social research: A brief history. *International Journal of Social Research Methodology* 14(3). 179–189.

Feuillet, Raoul-Auger. 1700. *Chorégraphie, ou l'art de décrire la dance par caractères, figures et signes démonstratifs*. Paris.

Geoghegan, Bernhard D. 2017. The family as machine: Film, infrastructure, and cybernetic kinship in suburban America. *Grey Room* 66. 70–101.

Goffman, Erving. 1963. *Behavior in public places. Notes on the social organization of gatherings*. New York: Glencoe.

Goffman, Erving. 1971. *Relations in public: Microstudies of the public order*. New York: Basic Books.

Goffman, Erwin. 1978. Response cries. *Language* 54(4). 787–815.

Goodwin, Charles. 1979. The interactive construction of a sentence in natural conversation. In George Psathas (ed.), *Everyday language: Studies in ethnomethodology*, 97–121. New York: Irvington.

Goodwin, Charles. 1993. Recording interaction in natural settings. *Pragmatics* 3(2). 181–209.

Goodwin, Charles. 2000. Action and embodiment within situated human interaction. *Journal of Pragmatics* 32. 1489–1522.

Goodwin, Charles. 2018. Why multimodality? Why co-operative action? *Social Interaction: Video-Based Studies of Human Sociality* 1(2).

Goodwin, Charles. 1981 [1977]. *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.

Goodwin, Charles & Richard Gwyn. 2003. The semiotic body in its environment. In Justine Coupland (ed.), *Discourses of the body*, 19–42. New York: Palgrave/-Macmillan.

Gumperz, John J. & Dell H. Hymes. 1972. *Directions in sociolinguistics: The ethnography of communication*. New York: Holt McDougal.

Haddington, Pentti, Lorenza Mondada & Maurice Nevile (eds.). 2013. *Interaction and mobility*. Berlin: De Gruyter.

Haley, Jay. 1962. Family experiments: a new type of experimentation. *Family Process* 1(2). 265–293.

Hall, Edward T. 1959. *The silent language*. Garden City: Doubleday.

Hall, Edward T. 1963. A system for the notation of proxemic behavior. *American Anthroplogist* 65(5). 1003–1026.

Hall, Edward T. 1966. *The hidden dimension*. Garden City: Doubleday.

Hall, Edward T. & G. L. Trager. 1953. *The analysis of culture*. Washington: American Council of Learned Societies.

Heath, Christian. 2012. *The dynamics of auction: Social interaction and the sale of fine art and antiques*. Cambridge: Cambridge University Press.

Heath, Christian, Jonathan Hindmarsh & Paul Luff. 2010. *Video in qualitative research*. London: Sage.

Heath, Christian & Paul Luff. 2000. *Technology in action*. Cambridge: Cambridge University Press.

Hockings, Paul (ed.). 1975. *Principles of visual anthropology*. The Hague/Chicago: Mouton/Aldine.

Hutchinson Guest, Ann. 1970. *Labanotation or kinetography Laban: The system of analyzing and recording movement*. New York: Theatre Arts Books.

Jacknis, Ira. 1988. Margaret Mead and Gregory Bateson in Bali: Their use of photography and film. *Cultural Anthropology* 3(2). 160–177.

Jefferson, Gail. 1972. Side sequences. In David Sudnow (ed.), *Studies in social interaction*, 294–338. New York: Free Press.

Jefferson, Gail. 1973. A case of precision timing in ordinary conversation: Overlapped tag-positioned address terms in closing sequences. *Semiotica* 9(1). 47–96.

Jefferson, Gail. 1974. Error correction as an interactional resource. *Language in Society* 3(2). 181–199.

Jefferson, Gail. 1983. On exposed and embedded correction in conversation. *Studium Linguistik* 14. 58–68.

Jefferson, Gail. 1984. Notes on some orderlinesses of overlap onset. In V. D'Urso (ed.), *Discourse analysis and natural rhetoric*, 11–38.

Jefferson, Gail. 1985. An exercise in the transcription and analysis of laughter. In Teun A. van Dijk (ed.), *Handbook of discourse analysis*, vol. 3, 25–34. London: Academic Press.

Jefferson, Gail. 1988. On the sequential organization of troubles talk in ordinary conversation. *Social Problems* 35(4). 418–441.

Jefferson, Gail. 2004. Glossary of transcript symbols with an introduction. In Gene H. Lerner (ed.), *Conversation analysis: Studies from the first generation*, 13–31. Amsterdam/Philadelphia: John Benjamins.

Jefferson, Gail. 2017. *Repairing the broken surface of talk: Managing problems in speaking, hearing, and understanding in conversation*. Oxford: Oxford University Press.

Kendon, Adam. 1977. The F-formation system: spatial orientational relations in face-to-face interaction. In Adam Kendon (ed.), *Studies in the behavior of social interaction*. Bloomington, IN: Indiana University Press.

Kendon, Adam. 1979. Some theoretical and methodological aspects of the use of film in the study of social interaction. In Gerald P. Ginsburg (ed.), *Emerging strategies in social psychological research*, 67–91. London: John Wiley.

Kendon, Adam. 1990. *Conducting interaction: Patterns of behavior in focused encounters*. Cambridge: Cambridge University Press.

Kendon, Adam. 2004. *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.

Kendon, Adam & Andrew Ferber. 1973. A description of some human greetings. In Richard M. Michael & John H. Crook (eds.), *Comparative ecology and behavior of primates*, 591–668. London: Academic Press.

Kendon, Adam & Stuart Sigman. 1996. Commemorative essay. Ray L. Birdwhistell (1918–1994). *Semiotica* 112(3–4). 231–262.

Knoblauch, Hubert, Bernt Schnettler, Jürgen Raab & Hans-Georg Soeffner (eds.). 2006. *Video analysis: Methodology and methods: Qualitative audiovisual data analysis in sociology*. Frankfurt: Lang.

Laban, Rudolf. 1928. *Schrifttanz*. Universal-Edition. Wien/Leipzig.

Leeds-Hurwitz, Wendy. 1987. The social history of *The Natural History of an Interview*: A multidisciplinary investigation of social communication. *Research on Language and Social Interaction* 20(1–4). 1–51.

Leeds-Hurwitz, Wendy. 1990. Notes in the history of intercultural communication: The Foreign Service Institute and the mandate for intercultural training. *Quarterly Journal of Speech* 76. 262–281.

Leeds-Hurwitz, Wendy. 2005. The natural history approach: Bateson's legacy. *Cybernetics and Human Knowing* 12(1–2). 137–146.

Leeds-Hurwitz, Wendy & Adam Kendon. 2021. The Natural History of an Interview and the microanalysis of behavior in social interaction: A critical moment in research practice. In James McElvenny & Andrea Ploder (eds.), *Holisms of communication: The early history of audio-visual sequence analysis*, 145–200. Berlin: Language Science Press. DOI: 10.5281/zenodo.5142288.

Leeds-Hurwitz, Wendy & Stuart Sigman. 2010. The Penn tradition. In Wendy Leeds-Hurwitz (ed.), *The social history of language and social interaction research: People, places, ideas*, 235–270. Cresskill: Hampton Press.

Lynch, Michael. 2002. From naturally occurring data to naturally organized ordinary activities: Comment on Speer. *Discourse Studies* 4(4). 531–537.

Lynch, Michael & David Bogen. 1994. Harvey Sacks's primitive natural science. *Theory, Culture & Society* 11(4). 65–104.

Marey, Étienne J. 1896. La chronophotographie sur plaque fixe en France et à l'Etranger. *Revue générale internationale, scientifique, littéraire et artistique* 2.

McQuown, Norman A. 1957. Linguistic transcription and specification of psychiatric interview materials. *Psychiatry* 20(1). 79–86.

McQuown, Norman A. 1971a. Chapter 9: Collation. In Norman A. McQuown (ed.), *The natural history of an interview*. Chicago: Microfilm Collection of Manuscripts on Cultural Anthropology, University of Chicago, Joseph Regenstein Library, Department of Photoduplication.

McQuown, Norman A. 1971b. Natural history method: A frontier method. In Alvin R. Mahrer & Leonard Pearson (eds.), *Creative developments in psychotherapy*, vol. 1, 431–438. Cleveland: The Press of Case Western Reserve University.

McQuown, Norman A. (ed.). 1971c. *The natural history of an interview*. Chicago: Microfilm Collection of Manuscripts on Cultural Anthropology, University of Chicago, Joseph Regenstein Library, Department of Photoduplication.

Mead, Margaret & Paul Byers. 1968. *The small conference: An innovation in communication*. Paris/Den Haag: Mouton.

Mondada, Lorenza. 2006. Video recording as the reflexive preservation-configuration of phenomenal features for analysis. In Hubert Knoblauch, Bernt Schnettler, Jürgen Raab & Hans-Georg Soeffner (eds.), *Video analysis*, 51–68. Bern: Lang.

Mondada, Lorenza. 2009. Emergent focused interactions in public places: A systematic analysis of the multimodal achievement of a common interactional space. *Journal of Pragmatics* 41(10). 1977–1997.

Mondada, Lorenza. 2012. The conversation analytic approach to data collection. In Jack Sidnell & Tanya Stivers (eds.), *The handbook of conversation analysis*, 52–56. London: Wiley-Blackwell.

Mondada, Lorenza. 2014. The local constitution of multimodal resources for social interaction. *Journal of Pragmatics* 65. 137–156.

Mondada, Lorenza. 2016. Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics* 20(3). 336–366.

Mondada, Lorenza. 2018. Multiple temporalities of language and body in interaction. Challenges for transcribing multimodality. *Research on Language and Social Interaction* 51(1). 85–106.

Mondada, Lorenza. 2019. Contemporary issues in conversation analysis: Embodiment and materiality, multimodality and multisensoriality in social interaction. *Journal of Pragmatics* 145. 47–62.

Mondada, Lorenza. 2021. *Sensing in interaction*. Cambridge: Cambridge University Press.

Muybridge, Eadweard. 1887. *Animated locomotion*. Philadelphia: University of Pennsylvania.

Ochs, Elinor. 1979. Transcription as theory. In Elinor Ochs & Bambi B. Schieffelin (eds.), *Developmental pragmatics*, 43–72. New York: Academic Press.

Pittenger, Robert E., Charles F. Hockett & John J. Danehy. 1960. *The first five minutes: A sample of microscopic interview analysis*. Ithaca: Paul Martineau.

Pollenz, Philippa. 1949. Methods for the comparative study of the dance. *American Anthropologist* 51. 428–435.

Ruby, Jay. 2000. *Picturing culture: Explorations on film & anthropology*. Chicago: University of Chicago Press.

Ruesch, Jurgen & Gregory Bateson. 1951. *Communication: The social matrix of psychiatry*. New York: Norton.

Ruesch, Jürgen & Weldon Kees. 1956. *Nonverbal communication. Notes on the visual perception of human relations.* Berkeley: University of California Press.

Sacks, Harvey. 1992. *Lectures on conversation [1964–72].* Oxford: Basil Blackwell.

Sacks, Harvey. 1989 [1966]. Introduction. *Human Studies* 12(3–4). 211–215.

Sacks, Harvey, Emanuel A. Schegloff & Gail Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50. 696–735.

Saint-Léon, Arthur. 1852. *La sténochorégraphie, ou L'art d'écrire promptement la danse.* Paris.

Schaeffer, Joseph H. 1970. *Videotape techniques in anthropology: The collection and analysis of data.* New York: Columbia University. (PhD dissertation).

Schaeffer, Joseph H. 1975. Videotape: New techniques of observation and analysis in anthropology. In Paul Hockings (ed.), *Principles of visual anthropology*, 255–284. The Hague/Chicago: Mouton/Aldine.

Scheflen, Albert E. 1964. The significance of posture in communication systems. *Psychiatry* 27(4). 316–331.

Scheflen, Albert E. 1971. Natural history method in psychotherapy: Communicational research. In Alvin R. Mahrer & Leonard Pearson (eds.), *Creative developments in psychotherapy*, 393–422. original-date: 1966. Cleveland, OH: Case Western Reserve University Press.

Scheflen, Albert E., Adam Kendon & Joseph A. Schaeffer. 1970. A comparison of videotape and moving picture film in research in human communication. In Milton M. Berger (ed.), *Videotape techniques in psychiatric training and treatment.* New York: R. Brunner.

Scheflen, Albert E. & Alice Scheflen. 1972. *Body language and the social order: Communication as behavioral control.* Englewood Cliffs: Prentice Hall.

Schegloff, Emanuel A. 1984. On some gestures' relation to talk. In John M. Atkinson & John Heritage (eds.), *Structures of social action: Studies in conversation analysis*, 266–295. Cambridge: Cambridge University Press.

Schegloff, Emanuel A. 1992. Introduction. In Harvey Sacks (ed.), *Lectures on conversation*, ix–lii. Oxford: Blackwell.

Schenkein, Jim. 1978. Sketch of an analytic mentality for the study of conversational interaction. In Jim Schenkein (ed.), *Studies in the organization of conversational interaction*, 1–6. New York: Academic Press.

Stepanov, Vladimir I. 1892. *L'alphabet des mouvements du corps humain.* Paris.

Tosi, Virgilio. 2005. *Cinema before cinema: "The origins of scientific cinematography".* London: British Universities Film & Video Council.

Trager, George L. 1958. Paralanguage: A first approximation. *Studies in Linguistics* 13(1–2). 1–12.

Trager, George L. 1971. Appendix 1: Paralanguage: A first approximation. In Norman A. McQuown (ed.), *The natural history of an interview*. Reprint of Trager (1958). Chicago: Microfilm Collection of Manuscripts on Cultural Anthropology, University of Chicago, Joseph Regenstein Library, Department of Photoduplication.

Trager, George L. & Henry L. Smith. 1957 [1951]. *An outline of English structure.* original-date: 1951. Washington: American Council of Learned Societies.

Van Vlack, Jacques. 1965. The research document film. *American Science Film Association Notes* April.

Van Vlack, Jacques D. 1966a. Filming psychotherapy from the viewpoint of a research cinematographer. In Louis A. Gottschalk & Arthur H. Auerbach (eds.), *Methods of research in psychotherapy*, 15–24. New York: Appleton-Century-Crofts.

Van Vlack, Jacques D. 1966b. Research filming for studies in human communication. *Journal of the University Film Producers Association* 18(2). 3–6.

Watter, Seth B. 2017. Scrutinizing: Film and the microanalysis of behavior. *Grey Room* 66. 32–69.

Winkin, Yves. 1984. *La nouvelle communication.* Paris: Seuil.

Zorn, Friedrich A. 1987. *Grammatik der Tanzkunst.* Leipzig: J. J. Weber.