
Vor, hinter, für und mit der Kamera: Viergliedriger Video-Analyserahmen in der Qualitativen Sozialforschung

Christine Moritz

Zusammenfassung Der vorliegende Beitrag stellt den Versuch dar, auf der Basis der aktuellen Bestandsaufnahme gegenwärtiger Literatur zur Thematik der Video- und Filmtranskription (zu denen die Beiträge der AutorInnen des vorliegenden Sammelbandes, versehen mit dem Kürzel *idB* für „in diesem Band“, hinzuzuzählen sind) unterschiedliche Weisen der Transkription im Spiegel eines „viergliedrigen Video-Analyserahmen“ zu reflektieren. Dieser Rahmen besteht aus den Kategorien „vor“, „hinter“, „für“ und „mit“ der Kamera im Video-/Filmmaterial, und dient – durchaus wie eine „Checkliste“ – dazu, methodenneutral und disziplinunabhängig das eigene Forschungsmaterial hinsichtlich dieser vier Aspekte in Augenschein zu nehmen.

Im Anschluss an einen *Streifzug durch die aktuelle Forschungsliteratur* (Textabschnitt 1) beschäftige ich mich mit den Fragen zur *Materialität des Videos* (Textabschnitt 3), mit den *Gründen für und gegen eine Videotranskription*, mit ihren möglichen *Funktionen* und auch mit ihren gegenwärtigen *Problemfeldern* (Textabschnitt 4). Ab Textabschnitt 5 werden forschungspraktische Fragen erläutert: Die methodenneutrale *Konzeption des Multikodaltranskripts* erlaubt es den Forschenden, sowohl repräsentative (Einzelbilder aus dem Video, Skizzen etc.), deskriptive (Verbaltranskriptionen, Notationen) als auch interpretative und/oder hermeneutische (Memos, Interpretationssitzungen etc.) Verfahren in einem *Feldpartitur-Design (FP-D)* zu bündeln, und auf diese Weise die qualitative Arbeit mit einem Video forschungspraktisch zu bewältigen. Der Beitrag endet in der Zusammenfassung der bisherigen Erläuterungen in der Form des „viergliedrigen Video-/Film-Analyserahmen“ (Textabschnitt 6), um abschließend die genannten Aspekte für die gegenwärtige Forschungspraxis in einem *Fazit* (Textabschnitt 7) zusammen zu fassen.

1 Einführung

„Vor, hinter, für und mit der Kamera?“ Die Kamera steht im Mittelpunkt – und mit der Kamera das Produkt, welches die Kamera für den oder die Qualitative/n Sozialforscher/in herstellt oder hergestellt hat: das aufgezeichnete Video- oder Filmmaterial¹ als Datensatz im Rahmen einer Forschungsanlage². Die Bezeichnungen „vor der Kamera“ und „hinter der Kamera“ wurden bereits mehrfach zur Bündelung eher komplexer Inhalte des Datenmaterials in der Forschungsmethodenliteratur verwendet (siehe Schwerpunktausgabe FQS Knoblauch et al 2008; Jones et al 2008 sowie Textabschnitt 2), und werden daher in diesem Beitrag unter Bezugnahme auf die gegenwärtige Forschungsliteratur ergänzt um zwei weitere Kategorien, „für die Kamera“ und „mit der Kamera“. Auch diese beiden weiteren Kategorien sind wie die ersten beiden an sich nicht neu; neu ist die Bündelung der vier Betrachtungsweisen „vor-hinter-für-mit“ als einem leicht handhabbaren „viergliedrigen Video-/Film-Analyserahmen“. Dieser entstand im Rahmen von Forschungswerkstätten mit Videomaterial der Forschenden, die ich im Rahmen unterschiedlicher Lehrveranstaltungen³ und Fachdisziplinen durchführte und auf diese Weise erprobte. Der viergliedrige Video-/Film-Analyserahmen kann für andere Forschende und v. a. auch für Lernende von Nutzen sein, weshalb er einen Beitrag zur Methodenlehre darstellt.

-
- 1 Im vorliegenden Beitrag verwende ich die beiden Bezeichnungen *Film* und *Video* nach Möglichkeit gemeinsam, teilweise jedoch auch allein, wenn der sprachliche Ausdruck eines Satzes dies erfordert. Die Begriffe, auch die Bezeichnungen *audiovisuelle Aufzeichnung* oder *Bewegtbild* wurden, wie auch weitere Begriffe des vorliegenden Beitrages, im Buch zur Feldpartitur (Moritz 2011) definiert. Übereinstimmungen zwischen der genannten Publikation und der hier vorliegenden sind in den Textabschnitten 5 und 6 gegeben, wobei eine Anpassung, Aktualisierung und Erweiterung der entsprechenden Textpassagen vorgenommen wurde
 - 2 Nicht im Vordergrund stehen demnach Forschungsweisen, welche das Datenmaterial Video/Fotografie zwar in – z. B. partizipativer Forschung – nutzen und einsetzen, jedoch nicht oder kaum in ihre Forschungsdokumentation einbeziehen, wie es etwa bei Emmison, Smith 2000, Holliday 2004 der Fall ist.
 - 3 Videoanalyse-Lehrkonzeption und Videotranskriptionen – und -analyseworkshops siehe in deutscher www.feldpartitur.de/schulung oder in englischer Sprache unter www.feldpartitur.de/en/training

2 Ein kurzer Streifzug: Videoanalyse in der Qualitativen Sozialforschung

Ein Florieren der Beschäftigung mit Formen audiovisueller Medialität und neuen Kommunikationspraktiken innerhalb unterschiedlicher Semiosen unserer Gesellschaft hat nach etlichen „turns“⁴ gerade in den letzten drei bis fünf Jahren⁵ zu einer Vielzahl an Auseinandersetzungen zu interpretativen Verfahren der Videoanalyse geführt, die nach einer Phase der Exploration und einer daran anschließenden Phase der Stagnation die beginnende forschungsmethodische und -methodologische Auseinandersetzung mit dem Medium anzeigt. Diese Auseinandersetzung spannt sich weitgehend entlang der zwei Pole „vor der Kamera“ und „hinter der Kamera“ auf. Während in Soziologie (Banks 2001, Van Leeuwen, Jewitt 2001, Mohn 2002, Bohnsack 2009, Corsten et al 2010, Reichertz 2014 idB, Englert 2014 idB) und Bildungs- und Erziehungswissenschaften (Ehrenspeck et al. 2003, Niesyto 2003, Frieberthäuser et al. 2007, Dinkelaker 2009, Schluß, May 2013, Kamper 2014 idB, Köhler 2014 idB, Tucek et al 2014 idB, Schuler 2014 idB) Anwendungsweisen der Videoanalyse rund um Verhaltensbeobachtung im sog. natürlichen Feld vor der Kamera (etwa im Bereich der Qualitätssicherung oder der erziehungswissenschaftlichen Professionsdebatten) bereits eine Anzahl grundlegender Schriften aufweist, sind eine hohe Vielzahl interessanter Perspektiven durch Ansätze zur Beobachtung nichtsprachlicher Kommunikation, wie im weitesten Sinne der Körperhaltung, der Körpersprache, der Gestik oder der Mimik vorfindlich (⁶Abfalter 2014 idB, Arellano 2014 idB, Hauke, Moritz 2014 idB., Helzle 2014 idB, Welling 2014 idB) oder auch der virtuellen Kommunikation unter dem Stichwort online-Diskurse oder mediale Kommunikation (Bliesener 2014 idB, Schüler, Stertkamp 2014 idB) zu sichten. Nicht unerwähnt bleiben sollen aufgrund der starken Verbreitung vernakuläre (also historisch gewachsene, nicht von WissenschaftlerInnen entwickelte) Ansätze der Videoanalyse im Bereich alltagsweltlicher Praktiken (wie etwa videobasiertes Führungskräfte-training, Bewegungsoptimierung im Hochleistungsbereiche wie Musik, Sport oder der

4 In chronologischer Folge „iconic turn“ oder „pictorial turn“ (Boehm 1996), „Video Revolution“ (o. A.), „pictorial turn“ (Mitchell 1997, Mitchell, Frank 2008), Postmodern visual turn (Moritz 2011, S. 13 f), „performative turn“ (Fischer-Lichte 2004, S. 25–27). Gründe für die – auch unter ähnlich plakativen Titeln wie bekanntgewordene Bewegung werden von Schnettler und Raab 2008 insbesondere in drei Faktoren gesehen: „... surely three factors have had a decisive influence: (a) the end of the logocentric paradigm, (b) the massive dissemination of visual media, and (c) the proliferation and easy access to visual devices in scientific research practice during the last decade.“ (Schnettler, Raab 2008, S. [25])

5 Siehe bis zum Jahr 2011 die Literaturrecherche Moritz 2011, S. 13–33. Zur Entstehungsgeschichte visueller Forschungsmethoden vergleiche Reichert 2007, Heath, Hindsmith, Luff 2010 (vorw. Ethnomethodology und Konversationsanalyse), Schnettler, Raab 2008, zur Video-Analyse in der Qualitativen Sozialforschung Tuma et al 2013, S. 19 f und die beiden bereits erwähnten Schwerpunktausgaben FQS Jones et al 2008, Knoblauch et al 2008

6 Aus Gründen des Umfangs seien hier nur die im vorliegenden Sammelband enthaltenen Beiträge dieser Gruppe genannt

immer differenzierter sich ausbildenden Videosurveillance). Die beiden weiteren Kategorien „für die Kamera“ und „mit der Kamera“ werden in dieser Ausdrucksweise in der Methodendiskussion meines Wissens bislang nicht genutzt. Es wird unter diesen Kategorien auf die verhältnismäßig wenig zitierten Film- und Designwissenschaften (siehe überblickend Hilt 2010, S. 212–217 und Joost 2008, S. 9–50, Hilt 2014 idB) die Filmsemiotik (Metz 2000, Metz, Blüher 2000, Barthes 1990, Eco 1977), Mediensemiotik (Hess-Lüttich et al 1990) sowie die Einflüsse zur Deutung filmischer Elemente mit Methoden der Psychoanalyse (Lacan 1973) hingewiesen. Annäherungen an das Material unter ästhetischen Aspekten (Reuter 2012, Fischer-Lichte 2004, Reichertz/Englert 2010), auch unter dem Begriff des Performativen⁷ (Mersch 2002, Krämer 2004, Liamputtong 2010, Hauke, Moritz 2014 idB, Hilt 2014 idB), der Intermedialität (Paech, Schröter 2007, Fraas 2008, Buchner 2010, Bliesener 2014 idB, Stertkamp, Schüler 2014 idB), beleuchten zunehmend – und keineswegs undiskutiert im Rahmen erkenntnistheoretischer Grundlagentebatten – die Schnittstellen zwischen Kunst und Wissenschaft (Kurt 2010, Troendle 2011). Als derzeit populärste Richtung der klassischen *Filmanalyse* gelten die Cultural Studies. Diese nehmen Filme – ebenso wie Texte oder andere Alltagsdokumente – im Sinne einer „radikal-kontextualistischen“ Vorgehensweise (Winter 2010, Bromley, Ang 1999) vor allem semiotisch in den Blick (Winter 1991, 2010). Das Problem der Polysemie (Barthes 1990), also der Vieldeutigkeit der Zeichen innerhalb eines Videos (Hess-Lüttich 1990, 2003), wird forschungsmethodisch allein durch die Kontextualisierung (interaktiver Kontext, Kontexte im Sinne der sozialen Lagerung) aufgelöst (siehe Willis 1981, Hepp 2004). Nach Bohnsack (2009, S. 120–121) wenden VertreterInnen der Cultural Studies insbesondere zwei unterschiedliche Verfahrensweisen an: entweder sie verfolgen eine Fokussierung der Rezeptionsweisen oder aber eine theoretische resp. historische Produktanalyse. Mikos (2003 S. 135 f) etwa fokussiert in zweitgenannter Gruppe mit der struktur-funktionalen Analyse die ästhetischen Strukturen des Films, wodurch die Beschränkung auf die Strukturkomponenten eines Videos und der Konventionen, die sie repräsentieren, vorgenommen wird.

Die fachspezifische Auseinandersetzung mit dem Material Video und Film in der Qualitativen Sozialforschung⁸ zog historisch betrachtet eine entsprechende *Methodenentwicklung* nach sich, die sich *multidisziplinär* auf die einzelnen Fächer auswirkte.

Der älteste Ansatz wird in der *Video-Ethnografie* im Rahmen der Ethnomethodologie auf das Ende des 19. Jh. datiert (Regnault 1885, 1886), die Gründungsschriften zu foto- und videoanalytischen Vorgehensweisen (Garfinkel, Bateson, Mead) werden auch heute noch zitiert (siehe Hohenberger 1988, Pink 2007, zur Weiterentwicklung der aus der *Konversationsanalyse* Sacks et al 1974, Selting et al 2009, etwa im Rahmen

7 Überblickend siehe bitte die Sonderausgabe im Forum Qualitative Sozialforschung Jones et al 2008

8 Quantifizierende Verfahren der Videoanalyse, wie sie in der Psychologie, den Bildungs- und Erziehungswissenschaften, der Markt- und Verhaltensforschung angewendet wird, werden in diesem Beitrag außen vor gelassen

der daraus hervorgehenden ethnomethodologisch orientierten Videoanalyse Heath et al 2010, vom Lehn 2014 idB). Neben video-ethnografischen Ansätzen sind vor allem die Vielfalt der *hermeneutischen Ansätze* (wie etwa objektiv-hermeneutische, Loer 2010, hermeneutisch wissenssoziologische Witte, Rosenthal 2007, Raab 2008, Kissmann 2009, Reichertz/Englert 2010) von Bedeutung. *Wirkungsorientierte Ansätze* lassen sich in der Filmtheorie finden, welche den Film aus der Perspektive der Rezeption fokussieren. Insb. der neoformalistische Ansatz von Bordwell, Thompson 2010 ist hier von Bedeutung, weiterhin für die Qualitative Sozialforschung jedoch auch Kaemmerling 1991, Grodal 1997, Emmison 2000, Lindsay 2000, Doelker 2001, Korte 2010. *Bildtheorien* (Panofsky 1987, Boehm 1996, 2006a, 2006b, Imdahl 1996, Mitchell 2008, 2012, Sach-Hombach 1999, 2009, Scholz 2009, Lucht et al 2013) werden in den videobasierten Forschungsmethoden dabei auffallend häufig, *Filmtheorien* (siehe überblickend Elsaesser, Hagener 2007) oder *Montagetheorien* (einführend Beller 1999) auffallend selten rezipiert. Während Panofsky (1987) und Balasz (2001) betonen, der Film als ein „moving picture“ bezöge seine Kraft in der Hauptsache eben aus dem Bild (zit. n. Bohnsack 2009, S. 140), wenden sich andere Klassiker als eigentlichem filmischen Ausdruckselement eher der Bewegung im Film (Deleuze 1985; Kracauer 1993), der ästhetischen Struktur eines Films (Metz 1972, S. 27; Souriau 1951) sowie dem Narrativen im Film (Metz 1972, S. 42; Bazin 1975; Thompson 1999; Eisenstein 1975) zu. *Bildfokussierende Methoden* thematisieren entsprechend in unterschiedlicher Weise das stehende Einzelbild aus dem Video als zentralen Informationsträger. Die Reduktion des Filmischen auf das stehende Bild erfolgt dabei zu einem Teil aus forschungsmethodologischen Gründen (siehe hierzu etwa Bohnsack 2009, Loer 2010), zu einem anderen Teil ist diese Vorgehensweise das Resultat auch heute noch in einigen Fällen mangelnde Kenntnis über technische Möglichkeiten der sozialwissenschaftlichen Handhabung eines Videos. Das *stehende Einzelbild* aus dem Video wird in vielen dieser genannten Fälle ergänzend neben den Verbaltranskripten eines Videos (die nur die gesprochene Sprache menschlicher Akteure berücksichtigen) interpretiert. Zu bildfokussierenden Verfahren zählt etwa (zumindest bis 2012) die Video-Interaktions-Analyse und die Videografie⁹ (Knoblauch 2004, Knoblauch et al 2006, Knoblauch et al 2008 Tuma et al 2013), sie beschäftigt sich, wie der Name andeutet, mit Videos als Beobachtungsmittel von Interaktion, welche zum Zweck der Forschung produziert wurden, nicht jedoch mit filmkulturellen Produkten etwa des kommerziellen Films oder der medialen Eigenproduktion. Die auf der *struktural-hermeneutischen Bildhermeneutik* (Oevermann 1979) und der *wissenssoziologischen Bildhermeneutik* (Soeffner 2004; Raab 2008, Kap. 5) beruhenden Verfahren beschäftigen sich inhaltlich, in Abgrenzung zur Kunstwissenschaft, mit den profanen Produkten (Willis 1981) visueller Alltags- oder Subkulturen mit den Methoden

9 Videografie wird bei Tuma et al 2013 nicht im Sinne einer „Videoaufzeichnung“ verstanden, sondern als Zusammensetzung der Worte Videoanalyse + Ethnografie = Videografie

der Wissenssoziologie (Schütz, Berger/Luckmann, Soeffner), so etwa in Keppler (2006), Raab (2008, 2012), Reichertz (2000) oder filmsoziologisch Winter (1991). Es besteht die Auffassung, dass Alltagskulturen Repräsentationen des impliziten Handlungswissens einer Gesellschaft darstellen, die entsprechend in den Gerinnungen eines Standbildes ihren Niederschlag finden (Bohnsack 2009). Die *Dokumentarische Forschungsmethode* als ein dezidiert bildfokussierender Ansatz innerhalb der Videoanalyselandschaft zollt in gleicher Weise „Respekt vor der profanen Alltagskultur“ (ebd S. 118). Hier wird das Bild – sowohl aus der Verhaltensbeobachtung wie auch aus Produkten des kommerziellen Films – jedoch nicht als lediglich ergänzende Information aufgefasst, sondern mithilfe der theoretisch auf der Kunst- und Bildwissenschaft (Imdahl 1996; Ikonologie Panofsky 2006) sowie der Semiotik (Eco 1977; Barthes 1979; Sachs-Hombach, Rehkämper 1999) fundierten Methode werden bildvermittelte Informationen systematisch und nachvollziehbar erschlossen. Das Bild soll eine gleichberechtigte, eine „emanzipierte“ Stellung im Forschungsmaterial gegenüber den Textdaten erhalten. Forschungspraktisch wird dabei weniger die Bedeutung des Bildes im situativen Gesamtzusammenhang, sondern im Gegenteil in seiner sog. „Eigenlogik“ in den Vordergrund gestellt. Forschungsmethodisch wird eine „paradigmatische Trennung“ (Burkard 2007, S 63) zwischen einer (vor-)ikonografischen, auf deskriptive Vorgehensweisen zurückführbare Ebene auf der einen Seite, einer ikonologischen, sinnkonstituierenden Ebene auf der anderen Seite vorgenommen. Erst die zweite Ebene konstituiert demnach das Bild in der/dem Betrachtenden durch die Zuschreibung einer Bedeutung. Die Rekonstruktion von Bedeutung erfordert bei diesem Prozess eine „Begabung zur synthetischen Intuition“ (Panofsky 1987 b, S. 221; siehe auch Sachs-Hombach, Rehkämper 1999) von den Forschenden.

Bei diesem stark verdichteten Streifzug durch unterschiedliche Ansätze wird deutlich, dass die Möglichkeiten einer Aufbereitung eines solch informationsreichen Datentypus wie dem Video oder dem Film im Sinne einer Videotranskription sich nicht einfach gestalteten, und eine grundlegende Auseinandersetzung erfordern. Dieser Gedanke wird fortgeführt in den folgenden Textabschnitten des vorliegenden Beitrages.

3 Was ist eigentlich ein Video?

Ein so komplexer Datentyp wie ein Film oder ein Video hat nicht nur heterogene Erscheinungsweisen (Goodman 2007, Seel 2009) im Sinne einer Materialität (Wagner 2001, Holl 2010), sondern auch heterogene Verwendungsweisen im Rahmen unterschiedlicher Forschungsansätze und Forschungsphasen. Ein Video oder ein Film kann in der mediatisierten Gesellschaft einseitig-technologisch hinsichtlich seiner formalstrukturellen Erscheinung für Forschende bedeutend sein. Oder es kann, als gedachter Gegenpol, im konstruktivistischen Sinn eher als ein Wahrnehmungs- oder Rezeptions-

phänomen betrachtet werden¹⁰. Soziokulturell ist es möglich, das Video bezogen auf sein implizites kulturelles Wissen, systemisch oder auch strukturell-semiotisch im Rahmen weiter oder enger gefasster Semiosen innerhalb einer vorfindlichen Gesellschaft aufzufassen (Posner 2003, Hess-Lüttich 2003).

Das für qualitativ Forschende heterogene Video- und/oder Filmmaterial wird in der forschungsmethodischen Fachliteratur entsprechend unterschiedlich kategorisiert. Knoblauch erwähnt im Jahr 2004 (S. 128 ff) neben den beiden oben bereits genannten Ebenen der Bildanalysen (in denen Videos mit einer Hermeneutik untersucht werden, wie sie auch für stehende Bilder gilt) und der Übertragung in sprachliche Bildbeschreibungen (um anschließend diese – also textuelle – Daten zu interpretieren) noch folgende drei Richtungen: Erstens die standardisierten codierten Analysen wie etwa experimentelle, gestellte oder nachgespielte Situationen, zweitens die Erstellung von Code-Kategorien für Verhalten und konversationsanalytisch orientierte Analyse audiovisueller Aufzeichnungen natürlicher Situationen sowie drittens die hermeneutisch geprägte Sequenzanalyse. Raab (2002) und Reichertz, Englert (2010) unterscheiden den Datentypus Video dahingegen innerhalb hermeneutisch-wissenssoziologischer Fallstudien in Datenkomponenten, welche sich als „vor der Kamera“, und Datenkomponenten, welche sich als „hinter der Kamera“ verorten lassen. Weitere Klassifikationen des Forschungsmaterials werden bei Corsten 2010, S. 8 ff, Hilt 2010, Reichertz, Englert 2010; Bohnsack 2009; Friebertshäuser et al. 2007; Marotzki et al. 2006; Ehrenspeck et al. 2003 diskutiert.

Im hier präferierten Ansatz werden, wie bereits erwähnt, zusätzlich zu den beiden Rahmenkategorien „vor“ und „hinter“ der Kamera zwei weitere hinzugefügt: „für“ die Kamera – dies ist die wechselseitige Re-Aktivität¹¹ zwischen dem Dargestellten und dem Darstellenden – und „mit“ der Kamera – darunter werden Aspekte zur medialen Zeigehandlung des Videos im Rahmen der jeweiligen Forschungsarbeit umfasst – so dass sich der Video-/Film-Analyserahmen als ein eben viergliedriger aufspannen lässt¹².

Der Video-Analyserahmen erlaubt es, ein vorliegendes Video oder einen vorliegenden Film nicht in der Manier einer die Forschenden schnell einengenden Werks-, Material- oder Situationsanalyse zu behandeln, sondern kulturelle, situative, institutionelle

10 Siehe zu dieser Dichotomie die Literatur der Filmtheorie, insb. die vielzitierte Gegenüberstellung der Rahmen- und Fenstermetapher bei Elsaesser, Hagener 2007, S. 23 ff.

11 Der Begriff der Aktivität und der Re-Aktivität wurde in einer vorangehenden Studie anhand eines Modells (Dialogischer Kubus) in Abgrenzung zu den Begriffen der „Interaktion“, „Reaktion“, „Reaktanz“ und auch der „Tätigkeit“ in vier verschiedenen Aktivitäts-Kategorien eines menschlichen Akteurs definiert: Moritz 2010 S. 116–127

12 Anschlüsse sind zu finden in den Grundlagentexten der Konversationsanalyse und der Wissenssoziologie, welche die Daten der Forschung als soziale Produkte interpretiert, wodurch ihre Genese nicht ohne Bezugnahme zu den Kontexten der Herstellung verstanden werden können. Siehe hierzu umfassend Keller 2012

oder auch mediale (Verweisungs-)Zusammenhänge zu integrieren, und auf diese Weise die Grenzen eines Genres, eines Verwendungszweckes des Films/Videos oder des Produktionskontextes verstehend zu überschreiten. *Forschungspraktisch* dient nicht eine vorfindliche Werkskategorie, sondern die materialen Niederschläge des Videos/Films (seine verräterischen „Spuren“ darin) das *empirische Zentrum*, um welches die Forschungsarbeit kreist und kreisen kann, und um welche entsprechend weitere Informationen (Kontext- oder Interpretationsmaterial etc.) gruppiert werden.

4 Warum Video-/Filmtranskription in der Qualitativen Sozialforschung?

4.1 Gründe für und gegen Video- und Filmtranskription

Die Frage, in welcher Weise und in welchem Umfang auditives und visuelles Daten-Rohmaterial zum Zweck der Forschung technisch aufbereitet (also konvertiert, komprimiert) und inhaltlich aufbereitet (also transkribiert, transformiert, verschlagwortet, annotiert, kodiert, archiviert u. v. a.) werden soll/muss/kann, bewegt Forschende seit dem Aufkommen visueller Datenmaterialien in der Qualitativen Sozialforschung (z. B. Bergmann 1984, Soeffner 2004, Garz 1994).

Anders als in textbasierten Forschungsarbeiten (z. B. in der Arbeit mit auditiven Aufzeichnungen von Interviews, überblickend Dittmar 2012), wo es seit den 80er Jahren zum (heute meist wenig in Frage gestellten) Usus gehört, das originäre auditive Material durch eine (meist die para- und nonverbalen Elemente eliminierenden, siehe hierzu Textabschnitt 5.1.) Sprechtranskription zu ersetzen, befindet sich dieser Forschungsbereich der Transkription im Bereich videobasierter Arbeiten derzeit in einem Prozess der Reflexion. Während bei der Audiotranskription das Transkript in den meistverwendeten Praktiken der Qualitativen Sozialforschung¹³ reduziert wird auf den deskriptiven Wortgehalt, und nichtverbale Inhalte wie etwa die Tonhöhenveränderung, der Ausdruck der Stimme oder „Subtexte“ des wortlos Mit-Gemeinten weitestgehend ohne Reflexion außen vor verbleiben (kritisch hierzu etwa Reichertz 2013d), scheint im Bereich der Videotranskription ein gerade umgekehrter Fall vorzuliegen. Mancherorts wird die Existenz (etwa bei Loer 2010) einer deskriptiven oder strukturellen Ebene im Video – jedoch mit Ausnahme der gesprochenen Anteile im Video – verneint mit dem Verweis auf den ausschließlich durch interpretative Verfahren zugänglichen Sinn

13 Siehe überblickend zu konversations- und gesprächsanalytischen Studien Konventionen, überblickend auf der Internetseite des ids Schuette 2010, siehe auch Deppermann 1999, Selting et al 2009, Bakhtin 2010. Hier sind grundlegende Schriften und Literaturhinweise zu den Aspekten der Transkription von nonverbaler Kommunikation oder körperlich-räumlicher Interaktion, wie etwa bei Schmitt 2013 zu finden. Kritisch zur gegenwärtigen Transkriptionspraxis in der Qualitativen Sozialforschung siehe hierzu z. B. Reichertz 2013d

der Bedeutungsgehalte eines Films oder eines Videos und zu Gunsten einer bleibend paradigmatischen Trennung zwischen Text und Bild auch für das Material eines Videos. Während also, wie die zitierten Quellen zeigen, im Bereich der *Audioaufzeichnung* eine Transformation des flüchtigen auditiven Materials zum Zweck der Fixation und auch der Validierung einer Arbeit selbst von hartnäckigsten KritikerInnen der Videotranskription stets gefordert und in Form der Verwendung eines ausgewählten Zeichensystems (hier: die Alphabetschrift) praktiziert wird, wird derselbe Vorgang bei der *Videoaufzeichnung* in manchen Fällen unter Bezugnahme auf Visualität und somit Nicht-Diskursivität pauschal abgelehnt. Beide Datensorten enthalten jedoch neben der gesprochenen Sprache weitere Informationen, weil im Video (Montage, Filmstruktur, Musik, Mimik, Gestik, Geräusch, Raumkonstellationen, Artefakte ...) wie auch in der Audio-Aufzeichnung (Stimmfarbe, Stimmverzögerungen, Geschwindigkeitsänderungen, Sprecherüberlappungen¹⁴ etc.) weitere Kommunikationssysteme (Codesysteme) Verwendung finden, die für die Verstehens-Arbeit der Qualitativen Sozialforscher/innen von Bedeutung sind. Aus meiner Sicht ist daher nicht nur eine Neudiskussion zu den Praktiken der Videotranskription erforderlich, sondern aus dieser Debatte heraus leiten sich alte Fragen in neuer Weise auch für die vielerorts unreflektiert praktizierte Audiotranskription ab.

Die Videotranskription mit der Feldpartitur entwickelt im Rahmen der Erfassung des Videos/Films eine differenzierende Position, indem sie sich unter dem *Konzept der Multikodalität* (siehe Textabschnitt 5.2.) auf die Transformationsfähigkeit der Konstituenten eines Videos/Films und die Möglichkeiten der Verschriftung und/oder Visualisierung (zurück)besinnt. Dies können, müssen jedoch keineswegs die sprachlich-diskursiven Bausteine eines Videos/eines Films sein. Sondern unter Bezugnahme auf die jeweils vorliegende Forschungsarbeit („Kriterium der Gegenstands-Angemessenheit“ Steinke 2000) werden in der Videotranskription mit der Feldpartitur a) relevante Konstituenten zunächst einmal *identifiziert*, und b) Transkriptions- sowie Interpretationsweisen einer Videopartitur in ihrer Unterschiedlichkeit zum Zweck des Nachvollzugs der jeweiligen Transkriptionsweise (sog. „Editier-Modi“) *gekennzeichnet* (siehe Textabschnitt 5).

Mit anderen Worten: Bereits dieser Prozess der Identifikation der für die eigene Forschungsarbeit zweckdienlichen Konstituenten (benannt soeben als Punkt a) und seine Entscheidung für eine der Formen der Verschriftung ist ein *qualitativ-heuristischer*, das heißt: Er wird nicht (immer) nach einem vorgegebenen (etwa standardisierten) Verfahren gesetzt, sondern (in vielen Fällen) aus dem Material selbst, also „grounded“, entwickelt.

Auch *mit* einer Multikodaltranskription bleibt bei der Interpretation eines Videos und der Erforschung des jeweiligen „Sinn“ im Rahmen seiner gesellschaftlichen und

14 Für die QSF bislang aufgrund der Komplexität der Schriftsysteme schwer zugänglichen Systeme siehe Dittmar 2010, überblickend die Internetseite des ids Mannheim Schütte 2010

forschungspraktischen Verortung ein *Rest*, den es nicht festzuhalten, sondern dem es sich aneignend resp. vermittelnd anzunähern gilt (s. z. B. Dewey n. Strübing, Schnettler 2004, S. 233–244).

Dass demnach eine Feldpartitur bei diesem Transformationsprozess nicht gleichzusetzen ist mit dem Video selbst (auch dies vermutlich ein Fehlschluss, der aus der Audiotranskriptionspraxis direkt übertragen wurde), sondern als ein Instrument der Forschungspraxis aus vielerlei Gründen (siehe in diesem Textabschnitt unten) dienlich ist, sollte bereits an dieser Stelle, spätestens jedoch am Ende des vorliegenden Beitrages klar geworden sein.

Die Frage nach der Integration auch visueller Kommunikate in fixierende Formate wird, so steht zu fürchten, nicht mehr lange die Vormachtstellung im forschungsmethodologischen Diskurs innehaben angesichts der rasanten technischen Entwicklungen innerhalb des letzten Jahrzehnts und des damit einhergehenden gesellschaftlichen Wandels, der nicht nur aus zusätzlichen visuellen Formen der Kommunikation neben den bewährten schriftsprachlichen besteht. Es wird die Forderung nach geeigneten Forschungsinstrumenten laut, die neben visuellen auch multimediale, elektronische, virtuelle und technische Kommunikationsformen der Gesellschaft berücksichtigt (siehe im vorliegenden Band die sich an diese Thematik herantastenden Beiträge von Stertkamp/Schüler 2014 idB oder Bliesener 2014 idB). Denn die vielzitierte Nichtdiskursivität des Materials kleidet sich zwar augenscheinlich als unbezwingbares Argument, dies scheint aber nahezu ohne Effekt angesichts der praktizierten Forschungsarbeiten in den einzelnen Disziplinen. Nicht nur die Film- und Kunstwissenschaften, die Techniksoziologie oder die Computer Sciences (Arellano 2014 idB, Helzle 2014 idB), auch weit verzweigte Fächer rund um die Gesundheitswissenschaften beschäftigen sich seit vielen Jahren mit multikodalen und -medialen Daten, die entsprechende Erfassungsmethoden erfordern. Schritt zu halten bei der Entwicklung entsprechender Hilfsmittel ist eine stets bleibende Herausforderung an die Technik und naturgemäß hinkt diese (hier: die technische Verwirklichung der Anforderungen an eine multikodale Videotranskription) den Anforderungen der wirklichen (Forschungs-)Welt hinterher, – was durch eine enge Verbindung mit der praktizierenden Forschungscommunity weitestgehend reduziert werden soll und auch kann.

Doch nach diesen Erläuterungen möchte ich die Ausgangsfrage in der Überschrift dieses Textabschnittes beantworten. Wann besteht eine Anforderung nach einer multikodalen, diagrammatischen Annäherung an ein Video-/Filmmaterial? Eine multikodale Videotranskription ist erforderlich in Forschungsanlagen,

- in denen die Herstellung eines fixierenden Dokuments aufgrund der Flüchtigkeit des Mediums erforderlich ist (Soeffner 2004, Reichertz 2013b, S. 45 ff, Loer 2010, Raab 2008) (Aspekt der **Fixation**),
- in denen forschungsethische Aspekte die Ablösung des Datenmaterials von personengebundenem (Bild)Material zum Zweck der Anonymisierung erfordern, oder,

Transkription von Video- und Filmdaten in der
Qualitativen Sozialforschung
Multidisziplinäre Annäherungen an einen komplexen
Datentypus
Moritz, C. (Hrsg.)
2014, XII, 536 S. 86 Abb., 13 Abb. in Farbe., Softcover
ISBN: 978-3-658-00878-9