

REINHOLD SCHMITT

DER „FRAME-COMIC“ ALS DOKUMENT MULTIMODALER INTERAKTIONSANALYSEN¹

1. Vorbemerkung

Die folgenden Ausführungen zum Frame-Comic sind in die Struktur des Sammelbandes integriert. Auf der einen Seite führt der Beitrag punktuell die Überlegungen zur Relevanz von Standbildern für die multimodale Interaktionsanalyse weiter, wie sie in dem Beitrag zur Standbildanalyse von Hausendorf/Schmitt (i.d.Bd.) vorgestellt werden. Zum anderen steht der Beitrag in einem engen Verhältnis zur Studie „Freiraum schaffen im schulischen Funktionsraum“ (Schmitt/Dausendschön-Gay 2015). Dieses Verhältnis wird neben der grundsätzlichen bidirektionalen Verweisstruktur beider Beiträge, die sich hinsichtlich ihrer methodologischen Reflexion wechselseitig stützen und ergänzen, vor allem durch den Aspekt der Arbeitsteilung charakterisiert. Die Studie zum Freiraum stellt im Rahmen dieses arbeitsteiligen Verhältnisses eine fallspezifische Exemplifizierung des methodischen Vorgehens bei der Frame-Comic-Analyse dar. Hier wird fallbezogen demonstriert, wie die Frame-Comic-Analyse funktioniert und welches Erkenntnispotenzial damit verbunden ist. Aus diesem Grund habe ich im vorliegenden Text auf eine exemplarische Analyse verzichtet.

2. Frame-Comic? Eine erste Annäherung

Die videobasierte Entdeckung der Relevanz des Visuellen für die Interaktionskonstitution führt im Kontext der multimodalen Interaktionsanalyse zur Entwicklung von speziell auf die visuell wahrnehmbaren Aspekte des interaktiven Verhaltens bezogenen Analyseverfahren und Verfahren der Datenkonstitution. Ich will nachfolgend unter dem Begriff „Frame-Comic“ einen bestimmten Typus visuell konstituierter (Sekundär-)Daten präsentieren, der im Zusammenhang mit der Entwicklung der „visuellen Erstanalyse“ (Kap. 4) als eigenständiges Analyseverfahren der multimodalen Interaktionsanalyse entstanden ist. Zum besseren Verständnis will ich an dieser Stelle vorgehend eine kurze Definition dessen geben, was ich unter einem Frame-Comic verstehe. Im Laufe der nachfolgenden Ausführungen, bei der die Genese des Frame-

¹ Dieser Beitrag ist eine stark komprimierte, speziell auf den Sammelband zugeschnittene Version eines detaillierter ausgearbeiteten Manuskripts mit dem Titel *Methodische Verfahren der multimodalen Interaktionsanalyse: „Visuelle Erstanalyse“ und „Frame-Comic“* (Schmitt i.Vorb.).

Comics dargestellt wird, werden alle relevanten Aspekte dieser Definition wieder aufgegriffen und detailliert diskutiert.

Der Frame-Comic ist – wie ein Transkript – ein Sekundärdokument, das für die Analyse bestimmter Fragestellungen im Rahmen von Interaktionsanalysen und raumanalytischen Erkenntnisinteressen als empirische Grundlage entwickelt wurde.

Der Frame-Comic besteht aus einer Abfolge von Standbildern, die in motivierter Weise als Frames aus einer Videoaufnahme extrahiert werden. Diese Standbildfolge bildet – wie ein gezeichneter Comic – wichtige Stadien und Veränderungen eines kohärenten und konsistenten Handlungszusammenhangs ab (beispielsweise eine zeitlich begrenzte, situative Raumnutzung oder eine raumgebundene und raumbezogene soziale Praktik wie „Warten“ (Dausendschön-Gay/Schmitt i.d.Bd.)). Der Frame-Comic stellt eine handlungskonsistente Standbildfolge dar, wodurch der abgebildete Zusammenhang, obwohl zwischen den einzelnen Frames vieles ausgelassen wird, in seiner grundlegenden Struktur versteh- und analysierbar wird. Dies liegt primär an der Wirksamkeit zweier Prinzipien: dem Kontinuitätsprinzip und dem Kontiguitätsprinzip. Das Kontinuitätsprinzip garantiert, dass die ausgewählten Bilder nicht aus unterschiedlichen und inkonsistenten Handlungszusammenhängen stammen. Das Kontiguitätsprinzip gewährleistet, dass die einzelnen Frames in ihrer Sequenzialität tatsächlich benachbart sind, wodurch keine Sprünge entstehen und die Bilder auch nicht in ihrer Platzierung getauscht werden können.

Aus analytischer Perspektive ist der Frame-Comic gerade wegen der motivierten Auslassungen von besonderem Interesse, weil die dadurch bewusst herbeigeführte „Informationsverknappung“ und „Ent-Dynamisierung des Vollzugs“ spezifische Rekonstruktionsverfahren verlangt.

Für den Frame-Comic sind zwei Aspekte konstitutiv: zum einen die methodisch motivierte Auflösung der zeitlichen Kontinuität des Videos durch die Selektion einzelner Frames und die damit geschaffene Statik sequenziell geordneter Einzelbilder; zum anderen die ebenfalls methodisch induzierte Abwesenheit von Verbalität. Der Frame-Comic ist also ein ausschließlich visuell wahrnehmbares Dokument ohne Anbindung an ein verbales Transkript: Aufgrund seines Status als autonomes Sekundärdokument fehlt dem Frame-Comic nicht nur kein Transkript, sondern er will auch keins haben!

Aufgrund dieser Eigenschaften ist der Frame-Comic beispielsweise eine adäquate und notwendige empirische Grundlage für die Analyse situativer Raumnutzung. Für die sozialtopografische Analyse sind die einzelnen Entwicklungsstadien des Frame-Comics wesentlich besser geeignet als der voll-

ständig dokumentierte Vollzug dieser Raumnutzung (Einleitung von Hausendorf/Schmitt i.d.Bd.; Schmitt/Dausendschön-Gay i.d.Bd.). Und auch im Kontext visueller Erstanalysen ist der Frame-Comic ein adäquates empirisches Dokument.

Und noch ein klärender Hinweis zum Abschluss dieser Vorinformation: Die Konstitutionslogik des Frame-Comics ist gänzlich unterschieden von aktuellen Versuchen (beispielweise Laurier 2014a, b), zur Gestaltung multimodaler Transkripte tatsächlich Anleihen beim Comic zu machen. Laurier (2014a) lässt beispielsweise die gemeinsam im Auto Sitzenden in gezeichneten Sprechblasen sprechen, so wie wir sie aus Comic-Zeichnungen kennen. Er geht noch einen Schritt weiter und repräsentiert mittels Sprechblasen im Bild verbale Aktivitäten von Beteiligten, die selbst nicht zu sehen sind (Laurier 2014b). Aus multimodaler Sicht ist das nicht unproblematisch, da dadurch die grundsätzliche Egalität auditiv und visuell wahrnehmbarer Informationen zu Lasten des Visuellen „überschrieben“ wird. Das ist im engen Wortsinn zu verstehen, denn die Sprechblasen überdecken einen nicht gerade geringen Teil des im Bild Sichtbaren und heben dadurch das Verbale deutlich hervor.

Im Zentrum des Beitrags steht die Beschäftigung mit folgenden Fragen:

- Was ist der methodische Eigenwert des Frame-Comics als empirische Grundlage für multimodale Interaktionsanalysen?
- Wie sieht der Prozess des visuellen Segmentierens als zentrale Prozedur der Konstitution eines Frame-Comics aus?
- Wie gestaltet sich das Verhältnis des Frame-Comics zum Verbaltranskript?
- Was sind die Potenziale und Grenzen des Frame-Comics?

Bevor ich mich der Beantwortung dieser Fragen zuwende, will ich zum besseren Verständnis kurz den übergeordneten Rahmen skizzieren, in dem der Frame-Comic steht.

Der Frame-Comic steht im Zusammenhang mit Überlegungen zur visuellen Erstanalyse (detailliert hierzu Schmitt i.Vorb.: Kap. 6.1) als einem systematischen Versuch der gegenstandsadäquaten Entwicklung eigenständiger multimodaler Analyseverfahren. Bei der visuellen Erstanalyse handelt es sich um ein methodisches Verfahren, das zu Beginn der Konstitutionsanalyse audiovisueller Interaktionsdokumente bewusst das gesamte akustisch wahrnehmbare Interaktionsgeschehen ausklammert. Die visuelle Erstanalyse² stellt je-

² Die Bezeichnungen ‘visuelle Erstanalyse’ und die in Kapitel 7 verwendete Bezeichnung ‘visuelles Segmentieren’ sind im Sinne der unökonomischen Langform zu lesen als: ‘Analyse visuell wahrnehmbaren Verhaltens ohne Bezugnahme auf verbale Aktivitäten’ und ‘Segmentierung visuell wahrnehmbaren Verhaltens ohne Bezugnahme auf verbale Aktivitäten’.

doch kein Gegenverfahren oder substitutiven Zugang zur Verbalanalyse dar. Es gelangt vielmehr entweder bei spezifischen Fragestellungen oder bei der bewussten Entwicklung methodischer Verfahren der multimodalen Interaktionsanalyse zum Einsatz.

Der methodisch motivierte Ausschluss auditiver Erscheinungsformen (insbesondere solcher sprachlicher Natur) produziert an Stelle von Verbaltranskripten Standbilder und Standbildfolgen als empirische Basis der Analyse. Diese haben den gleichen Stellenwert wie Verbaltranskripte und werden wie diese sequenzanalytisch bearbeitet (siehe dazu den Beitrag zur Standbildanalyse von Hausendorf/Schmitt i.d.Bd.). Visuelle Erstanalysen stehen am Anfang eines mehrschrittigen methodischen Verfahrens, bei dem mit jeweils spezifischer Fokussierung auf bestimmte Aspekte (Ausdrucksressourcen oder Aspekte der Interaktionskonstitution) wiederholt durch denselben Dokumentausschnitt „gegangen“ wird. Ich bezeichne diese generelle Struktur der konstitutionsanalytischen Rekonstruktion als „systematische analytische Rekurrenz“ (Schmitt i.Vorb.). Der Frame-Comic ist ein spezifisches Ergebnis dieser visuell wahrnehmbaren und konstituierten Sekundärdaten in Form von Standbildern.

3. Wie analysiert man eine Gesprächspause?

Die methodischen Überlegungen zum Frame-Comic gehen auf die Rekonstruktion der interaktiven Struktur einer längeren Gesprächspause³ zurück (Schmitt 2004). Bei der Sichtung so genannter Pitching-Situationen im Kontext der Ausbildung von Filmstudenten⁴ (Heidtmann 2009) fiel eine etwa 14-sekündige Gesprächspause auf und rückte in den analytischen Fokus. In dieser Phase passierte – ohne dass gesprochen wurde – ausgesprochen viel, und es war möglich, das interaktive Geschehen und den Gang der Interaktionskonstitution gänzlich ohne den vorherigen und nachfolgenden Kontext, in dem gesprochen wird, allein aufgrund des visuell wahrnehmbaren Verhaltens der Beteiligten detailliert zu rekonstruieren. Möglich war dabei nicht nur die Rekonstruktion einzelner Entwicklungsetappen der Interaktion in der Gesprächspause. Es konnten darüber hinaus auch unterschiedliche Beteiligungsweisen der Studierenden identifiziert, der pragmatische Gehalt ihrer interaktiven Präsenz formuliert und eine klare Hypothese (eine Seh-Art) entworfen werden, die erklären konnte, warum es gerade ein spezifisches Mitglied der studentischen Gruppe war, das diese Gesprächspause beendet.

³ Der Begriff ‘Gesprächspause’ ist aus multimodal-interaktionistischer Perspektive problematisch, da das Gespräch, d.h. die verbalen Aktivitäten, immer noch den normativen Bezugspunkt darstellen. Die korrekte, wenn auch etwas „eckige“ Formulierung wäre demnach ‘Interaktionsphase ohne verbale Aktivitäten’.

⁴ Pitchings sind Arbeitssitzungen, in denen Filmstudenten gemeinsam mit Dozenten Ideen für Filme entwickeln.

Da es in der Gesprächspause keine verbalen Aktivitäten gibt, konnte bei der Analyse ihrer interaktiven Struktur kein klassisches Transkript als empirische Grundlage eingesetzt werden. Gleichwohl war es aus Gründen der Intersubjektivierung der empirischen Grundlage und der Analyseergebnisse nötig, eine für das spezifische Erkenntnisinteresse adäquate empirische Basis zu erstellen. Es war klar, dass sie in der Dokumentation ausschließlich visuell wahrnehmbaren Verhaltens bestehen musste. Vergleichbar dem Verbaltranskript sollte es sich um ein Sekundärdokument handeln, das die Dynamik der Videoaufzeichnung in systematischer Weise in die Statik einzelner, sequenziell geordneter und motiviert ausgewählter Standbilder überführt.

Für die Logik der Standbildextraktion war die Orientierung bestimmend, primär den Wechsel von statischen Phasen und Momenten im Auge zu behalten, in denen sich die Beteiligten in signifikanter Weise bewegten, ihre Kopfhaltung und Blickrichtung veränderten, ihre Sitzposition modifizierten etc. Als Ergebnis dieses Transformationsprozesses entstand eine Abfolge von Standbildern, die nun in vergleichbarer Weise wie ein Transkript als Bezugspunkt der konstitutionsanalytischen Rekonstruktion dienen konnte. Sie sah wie folgt aus:



Beginn der Gesprächspause



Blickveränderung Özkan



Blick Anja zu Michael



Anja Hand am Kinn



Michael Hand am Kinn



Özkan Hand am Kinn

Da der fallspezifische Charakter im Vordergrund meines Erkenntnisinteresses stand, habe ich mir damals über die Tragweite der methodischen Implikationen dieses Vorgehens der durch die Relevanz der Daten selbst konstituierten Standbildfolge keine weiterführenden Gedanken gemacht. Ich wollte (vor allem in Bezug auf die systematische Untersuchung von Bergmann 1982) wissen, was man mehr und anderes als die klassische Konversationsanalyse über Gesprächspausen erfahren kann, wenn man sie zum Gegenstand der multimodalen Interaktionsanalyse macht.

Ich verfügte damals noch nicht über das Konzept 'Frame-Comic' und die dafür nötige methodologische Sensibilität, um die konstituierte Standbildfolge im expliziten Vergleich und Kontrast mit Verbaltranskripten in ihrer spezifischen, interaktionsdokumentarischen Eigenwertigkeit zu reflektieren. Dieses methodologische Bewusstsein entstand schrittweise, primär motiviert durch die Zusammenarbeit mit Heiko Hausendorf und der sich dabei entwickelnden Frage nach der erkenntnismäßigen Eigenständigkeit und methodischen Salienz von Standbildern für interaktionistische Erkenntnisinteressen. Gleichwohl war die damalige analytische Beschäftigung mit der interaktiven Struktur einer längeren Interaktionsphase, in der nicht gesprochen wird, der erste Schritt in Richtung Frame-Comic. Es war die erste Konstitution einer ausschließlich aus visuellem Verhalten bestehenden Folge motivierter Standbilder als eigenständigem Dokument multimodaler Interaktionsanalyse.⁵

Die Analyse der Gesprächspause ist ein *clear case* im Kontext der Notwendigkeit, visuell wahrnehmbare Interaktion aufgrund ihrer eigenen Vollzugssystematik zu rekonstruieren. Turn-Design, Wortwahl, Formulierungsdynamik, Prosodie etc. fallen als verbalitätsspezifische, akustisch wahrnehmbare Hilfstruppen in diesem Falle ja aus. Hat man sich dazu entschieden, die interaktive Struktur und den Konstitutionsgang der Gesprächspause zum Gegenstand zu machen, ist man alternativlos auf eine Abfolge von systematisch extrahierten Standbildern als empirische Grundlage angewiesen. Solche Analysen sind kein akademischer Selbstzweck. Es sind primär zwei Gründe, die ihre Relevanz ausmachen.

Zum einen sind Analysen visuell wahrnehmbaren Verhaltens Lösungen für Probleme, die bei der Verfolgung fallspezifischer Fragestellungen als methodisches Erkenntnisinstrumentarium entstehen. In dieser Hinsicht können – und wollen – visuelle Analysen die Rekonstruktion verbaler Strukturen als wesentliche Bestandteile der Interaktionskonstitution natürlich nicht ersetzen. Die eigenständige Analyse visuell wahrnehmbaren Verhaltens ist vielmehr integraler Bestandteil der Analyse des Gesamtverhaltens der Interaktionsbeteiligten auf der Basis einer systematischen methodischen Rekurrenzstruktur. Zum anderen ist die Analyse visuell wahrnehmbaren Verhaltens das zentrale methodische Verfahren im Kontext spezifischer Fragestellungen, die sich auf die Rekonstruktion situativer Ordnungsstrukturen beziehen, welche die Beteiligten unter motivierter Abstinenz von Verbalität produzieren.⁶

⁵ Der Frame-Comic kam (ohne methodologische Reflexion) bereits bei der Analyse von „Gehen als situierte Praktik“ (Schmitt 2012a) und der Rekonstruktion der „Störung und Reparatur eines religiösen Ritus“ (Schmitt 2012c) zum Einsatz.

⁶ Zu weiteren Untersuchungen, die ausschließlich visuell wahrnehmbares Verhalten zum Gegenstand haben, siehe beispielsweise „Gehen als situierte Praktik“ (Schmitt 2012a, 2012c) sowie „Be-Greifen“ relevanter Aspekte der unmittelbaren interaktionsarchitektonischen Umgebung (Schmitt 2013a).

4. Visuelle Erstanalyse

Es gibt im Rahmen multimodaler Interaktionsanalysen die Möglichkeit, mit der Rekonstruktion der visuell wahrnehmbaren Interaktionsphänomene auch in solchen Fällen zu beginnen, in denen auditiv wahrnehmbares Verhalten für die Interaktionskonstitution zentral ist. Eine solche Entscheidung führt dann zur visuellen Erstanalyse, mit der die multimodale Komplexität des Interaktionsdokuments motiviert reduziert werden kann. Dieser Analysegang kann auf der Grundlage von Videosegmenten, Standbildfolgen und in bestimmten Fällen auf der Grundlage von Frame-Comics durchgeführt werden.

Der Begriff 'visuelle Erstanalyse'⁷ verweist auf die Position und Funktionalität dieses Verfahrens: Nur wenn die Analyse des visuellen Interaktionsverhaltens noch nicht durch das konstitutionsanalytische Wissen um das verbale Geschehen beeinflusst ist, ist das Verfahren hinsichtlich der Identifikation neuer, für die Interaktionskonstitution relevanter Aspekte und in deskriptiv-kategorialer Hinsicht produktiv. Die visuelle Erstanalyse ist in dieser Hinsicht ausgesprochen implikationsreich und erkenntnisgenerierend. Bei der detaillierten Deskription visueller Verhaltensaspekte bildet sich gerade durch das Fehlen eines auf Verbalität bezogenen kategorialen Vorverständnisses eine strukturreflexive, deskriptive Beschreibungssprache aus. Dieser deskriptive Beschreibungsmodus ist geeignet, visuell wahrnehmbare Aspekte der Interaktionskonstitution auch in ihrer „versteckten“ pragmatischen Relevanz in den Blick zu nehmen.

Bei der visuellen Erstanalyse eines Videoausschnitts vom Filmset (Schmitt/Deppermann 2007) konnte man so z.B. sehen, wie die Regisseurin ihre Kamerafrau durch eine Umarmung und eine gleichzeitig realisierte Bewegung um die eigene Achse „dreht“. Nachdem anhand der sprachlichen Erscheinungsformen der Szene nachfolgend eine Auseinandersetzung zwischen den beiden Protagonistinnen rekonstruiert werden konnte, wurde das Umdrehen im Sinne eines gesprächsrhetorischen Zuges sichtbar.⁸ Es hat in einem ganz unmittelbaren Sinne die pragmatische Implikation von „jemanden von seiner ursprünglichen Position abbringen und ihn in die eigene Richtung orientieren“. Dies ist ein instruktives Beispiel dafür, dass relevante pragmatische Orientierungen der Beteiligten nicht nur verbalisiert, sondern auch in einem unmittelbaren Sinn verkörpert werden.⁹

⁷ Erste Hinweise auf das Verfahren „visuelle Erstanalyse“ und seine Bedeutung im Kontext der multimodalen Interaktionsanalyse finden sich in Schmitt (2007a, S. 407ff.) sowie in Schmitt (2007b, S. 40-43), eine prototypische Umsetzung des Verfahrens stellt Schmitt/Knöbl (2013) dar.

⁸ Zu den theoretischen Implikationen der Gesprächsrhetorik siehe Kallmeyer (1996) sowie als Exemplifikation des Ansatzes Kallmeyer/Schmitt (1996).

⁹ Weitere Beispiele hierfür sind „jemanden umdrehen“ (Schmitt/Deppermann 2007) oder „jemanden auf den Boden zurückholen“ (Pustički/Schmitt 2015).

Es gibt unterschiedliche Argumente für die Produktivität und Notwendigkeit, das Verfahren der visuellen Erstanalyse einzusetzen:

- Das Verfahren ist die methodologisch motivierte, analysefaktische Umsetzung des aus der *context analysis*¹⁰ stammenden theoretischen Postulats der Egalität aller Ausdrucksressourcen,
- es sensibilisiert für autonome Konstitutionsleistungen des visuell wahrnehmbaren Interaktionsverhaltens,
- es erweitert unser Wissen über kulturelle Praktiken jenseits von Verbalität (beispielsweise „Gehen als situierte Praktik“ (Schmitt 2012a)),
- es führt zur multimodalen Substantiierung bislang mono-modaler Konzepte wie „recipient design“ (Schmitt/Knöbl 2013, 2014) und
- es erweitert den Bereich verstehensdokumentarischer Aktivitäten (Deppermann 2008, Deppermann/Schmitt 2009, Schmitt 2010) und macht Intersubjektivitätskonstitution in ihrer faktischen empirischen Vielfalt und Komplexität deutlich.

Visuelle Erstanalysen (als Bestandteil einer multimodalen Interaktionsanalyse) kann man grundsätzlich auf der Basis von Primär- oder Sekundärdokumenten durchführen. Entscheidet man sich dazu, die Dynamik der Videoaufnahme an relevanten Stellen einzufrieren, um sich detailliert dem eingefrorenen Visuellen zuzuwenden, konstituiert man Standbildfolgen, die unter bestimmten Bedingungen den Status eines Frame-Comics haben.

5. Zur Begriffsbildung

Der Begriff ‘Frame-Comic’ ist ein Kompositum, das aus den beiden Teilen ‘Frame’ und ‘Comic’ besteht. Beide Begriffsteile haben eine jeweils eigene Bedeutung, die teils wissenschaftlicher, technischer und alltagsweltlicher Natur ist.

Der Frame: ‘Frame’ verweist auf den videotechnischen Zusammenhang und bedeutet zunächst nichts anderes als „Einzelbild“. Der zeitliche Abstand zwischen zwei Frames ist im PAL-System, das mit insgesamt 25 Bildern pro Sekunde arbeitet, die kleinste definierte Einheit des Videos pro Sekunde. Der Zeitraum zwischen zwei Frames hat im PAL-System die Dauer von 0,04 Sekunden. Der Frame selbst hat keine messbare zeitliche Dimension, zumindest keine, die für die Analyse relevant wäre.

Er ist also eine Einheit der videotechnischen Verarbeitung und wird aus der dynamischen Videoaufnahme extrahiert. Man bewegt sich durch die Video-

¹⁰ Zur *context analysis* siehe unter anderem Birdwhistell (1970), Heilman (1979), Kendon (1990) und Schefflen (1964, 1972).

aufzeichnung und trifft dabei Extraktionsentscheidungen auf der Grundlage spezifischer Erkenntnisinteressen (siehe visuelles Segmentieren: Kap. 7). Man wählt dann beispielsweise Frame 8, 17 oder 23 als für die weitere Interaktionsanalyse relevant aus (oder Frame 5, 125 und 437). Im ersten Fall unterscheidet sich Frame 8 von Frame 17 dadurch, dass 9 weitere Frames zwischen ihnen liegen. Frame 8 zeigt also ein anderes Bild als Frame 17, was jedoch nur dann ins Auge fällt, wenn sich zwischen den beiden Frames etwas verändert hat. Ist dies nicht der Fall, zeigen Frame 8 und 17 das gleiche Bild, und man kann im Zweifelsfall nicht erkennen, dass es sich um zwei unterschiedliche Frames handelt, die 0,36 Sekunden auseinanderliegen.

Unter einer videotechnischen Perspektive ist ein Frame also das Ergebnis einer durch die Bildtechnik eines bestimmten Videosystems produzierten systematischen Segmentierung, bei der die Dynamik des Videos auf die Statik einzelner Bilder reduziert wird, auf die man alle 0,04 Sekunden Zugriff hat.

Unter einer *interaktionsanalytischen* Perspektive ist ein Frame der Grundbaustein für die Konstitution eines visuell basierten Sekundärdokumentes und damit die empirische Grundlage konstitutionsanalytischer Investigation. Dieser konstitutionsanalytische Prozess nutzt die videotechnischen Möglichkeiten, um das zu untersuchende Interaktionsgeschehen in einer spezifischen Weise festzuhalten und in genau 0,04 Sekunden auseinanderliegenden Schnitten einzufrieren. Die Selektion der Frames 8, 17 und 23 geschieht nun nicht zufällig, sondern folgt einer eng an der Handlungs- und Aktivitätsstruktur orientierten Weise. Es ist genau diese Systematik, die nun mit dem Comic den zweiten Begriffsteil in den Fokus rückt.

Der Comic: Anders als beim 'Frame' handelt es sich beim 'Comic' um einen Begriff aus der Alltagskultur, der uns – auch wenn wir keine aktiven Leser bzw. Betrachter dieser erzählerischen Klein- und Kunstform sind – geläufig ist. Nach den gängigen Vorstellungen handelt es sich bei einem Comic um eine gezeichnete Darstellung eines Geschehens oder einer Geschichte in Form von Einzelbildern, die zu einer sequenziellen Bilderfolge zusammengestellt werden.

Ein Comic ist eine gewissermaßen auf das Minimum reduzierte, aber Verstehen sichernde Darstellung einer Handlungssequenz. Bei dieser reichen die gezeichneten Bilder aus, um der Geschichte, die erzählt werden soll, in ihrer Entwicklungsdynamik und Handlungslogik folgen zu können. Obwohl die Bilder in einem offensichtlichen zeitlichen Verhältnis zueinander stehen, ist klar, dass viel Geschehen dazwischen ausgelassen wurde, weil dessen Darstellung für das Verstehen der Geschichte nicht unbedingt wichtig ist. Die gezeichneten Bilder tragen also als Eckpunkte die Geschichte und enthalten – aus Sicht des Comic-Zeichners und Autors – alle für das Verständnis relevanten Informationen.

Der Frame-Comic: Der 'Frame-Comic' ist die Kombination beider Begriffsbestandteile, wobei 'Comic' mit seinem Prinzip der Ausschnitthaftigkeit und der systematischen Auslassung auf die Spezifik verweist, in der das dokumentierte Geschehen für die Analyse zur Verfügung gestellt wird.¹¹ Die 'Frames' sorgen hingegen für die empirische Grundlage der durch die Videotechnik vorgegebenen, selektiven Repräsentation des Dokuments. Die Kombination beider Begriffe entzieht das Konzept dem alltagsweltlichen Zusammenhang kulturspezifischer Kleinformen bildlichen Erzählens und verortet es im Kontext der multimodalen Interaktionsanalyse.

Der Frame-Comic ist im Unterschied zum alltagsweltlichen Comic kein Primärdokument, anhand dessen das Erzählen und Bebildern einer Geschichte analysiert werden kann. Vielmehr verweist er als Sekundärdokument auf die Videoaufnahme, aus der er extrahiert wurde. Er ist auch nicht das Ergebnis der Phantasie oder des autonomen Gestaltungswillens eines Autors, und er erzählt auch keine (von den im Primärdokument festgehaltenen Ereignissen) unabhängige, von einem Autor (oder wie beim Comic: Zeichner) ausgedachte Geschichte. Der Frame-Comic gibt vielmehr als Sekundärdokument ausschnitthaft und in Form von Einzelbildern das im Video dokumentierte, unabhängig vom Gestaltungswillen des Analytikers bereits existierende Interaktionsgeschehen wieder. Der wesentliche Unterschied zwischen dem Frame-Comic und dem klassischen Comic ist also der Aspekt der Abbildung eines dokumentierten Interaktionsgeschehens gegenüber der narrativen Gestaltung eines Ereignisses. Der Zeichner entwirft seine Bilder entlang der Dramaturgie, die er (s)einer Geschichte geben will; der Wissenschaftler hingegen extrahiert aus einem bereits vorliegenden Interaktionsdokument die von ihm als relevant erachteten Frames, um für die eigene erkenntnisgesteuerte Analyse eine angemessene empirische Basis zu gewinnen. Und natürlich beinhaltet ein Frame-Comic aus methodischen Gründen auch keine Sprechblasen oder sonstigen Hinweise auf das verbale Interaktionsgeschehen, die das abgebildete Geschehen verstehbar machen sollen.

Jenseits dieser strukturellen Unterschiede gibt es aber eine Gemeinsamkeit, die letztlich für die Begriffsbildung verantwortlich ist. Im Frame-Comic sollen – wie beim herkömmlichen, gezeichneten Comic – alle vom Analytiker als wesentlich erachteten Veränderungen des im Video vollständig dokumentierten Interaktionsgeschehens in einzelnen, sequenziell geordneten Standbildern präsentiert sein. Der Frame-Comic gibt in diesem Sinne die Eckdaten einer Handlungslogik oder eines Aktivitätsstrangs – wie beim Comic – auf der Grundlage ausgewählter Frames/Bilder wieder.

¹¹ Beim gezeichneten Comic spielt zudem das gesprochene Wort, das oft genug in einer Sprechblase Platz finden muss, häufig nicht die zentrale Rolle.

Beim Comic rekonstruiert der Leser bzw. Betrachter aufgrund der gezeichneten Einzelbilder und der darin festgehaltenen Veränderungen kontinuierlich diejenige(n) Aktion(en), die notwendigerweise realisiert worden sein muss/müssen, um die festgehaltenen Veränderungen zu bewirken. Er erschließt so das Gesamtgeschehen aus den Einzelbildern. Das Gleiche gilt auch für den Frame-Comic. Auf dessen empirischer Grundlage ist es dem Analytiker ebenfalls möglich, das zwischen den ausgelassenen Frames dokumentierte Geschehen auf Grundlage der ausgewählten Standbilder zu rekonstruieren.

Während der Comicleser die beschriebenen Prozesse beim Lesen jedoch automatisch und habituell realisiert, handelt es sich bei der Analyse eines Frame-Comics um ein methodisches Vorgehen, das bestimmten Prinzipien der Bearbeitung folgt (Schmitt i.Vorb., Kap. 7.6). Wie bei jedem anderen methodischen Vorgehen im Rahmen interaktionsanalytischer Anstrengungen spielt die Rekonstruktion der De-facto-Methodologie (Schmitt 2001) für die Kontrolle und Absicherung des methodischen Vorgehens bei der Konstitution eines Frame-Comics und bei seiner Analyse eine wesentliche Rolle. Nur wenn man sich damit beschäftigt, was man *de facto* tut, hat man die Möglichkeit, die einzelnen Bestandteile und die Angemessenheit des methodischen Vorgehens zu erkennen und einzuschätzen (Kap. 7.2).

6. Frame-Comic als Interaktionsdokument

Der Frame-Comic hat wie das Verbaltranskript (siehe Kap. 8) seinen Platz im Rahmen interaktionsanalytischer Erkenntnisinteressen. Die einzelnen Frames, die den Comic konstituieren, sind einzig und allein durch die Struktur und die Dynamik der im Video dokumentierten Interaktion motiviert und bleiben immer an die Interaktion gebunden, die in ihnen ausschnitthaft eingefroren ist. Sie besitzen keinen davon unabhängigen Zweck, mit dem man sich etwa unter bildästhetischen Gesichtspunkten beschäftigen kann (Hausendorf/Schmitt i.d.Bd.). Es ist diese interaktions-dokumentarische Qualität der Frames, die das Sekundärdokument konstituiert und strukturiert. Und es ist das jeweilige interaktionistische Erkenntnisinteresse, das darüber entscheidet, wie viele Frames ein solcher Comic hat und wie klein- bzw. großschrittig das Interaktionsgeschehen im Frame-Comic abgebildet wird.

Ungeachtet der durch die spezifische Erkenntnislage motivierten Schnittfolge und dem zwischen den einzelnen Frames bestehenden zeitlichen Verhältnis (zur Bedeutung der Zeitlichkeit siehe Kap. 8.3) wird der Frame-Comic entlang der dokumentierten Interaktionsentwicklung konstituiert. Der Bezug des Frame-Comics zur Interaktionsstruktur ist jedoch nicht nur ein formalstruktureller, für den Sequenzialität die zentrale Konstitutionsgrundlage darstellt. Der Frame-Comic besitzt vielmehr einen direkten Bezug zur Hand-

lungsstruktur. Es sind konsistente und gestalthafte, thematisch-pragmatische Ausschnitte aus einem audiovisuellen Interaktionsdokument, die es überhaupt sinnvoll machen, einen Frame-Comic zu konstituieren. Hier gibt es eine Parallele zur gezeichneten Bildergeschichte: Ein komplexes – und theoretisch unendlich verfeinerbares – Interaktionsgeschehen wird durch die Auswahl weniger Standbilder auf die basalen Konstituenten seiner Handlungsstruktur „eingekocht“.¹²

Auch in anderen Zusammenhängen und mit anderen Erkenntnisinteressen werden Standbilder aus Videoaufzeichnungen analysiert. Diese müssen dabei jedoch nicht automatisch den Status konsistenter Interaktionsdokumente haben, obwohl deren Analyse ebenfalls aus einer interaktionistischen Perspektive erfolgt. Aufgrund der fehlenden Aktivitätsbindung verliert in solchen Fällen die Sequenzialität des Interaktionsdokuments, die für den Frame-Comic bestimmend ist, mehr und mehr ihre Bedeutung für die Analyse.

Analysiert man beispielsweise den Altarbereich einer Kirche im Rahmen interaktionsarchitektonischer Erkenntnisinteressen (wie Hausendorf/Schmitt i.d.Bd.), ist die Sequenzialität des Interaktionsdokuments, über den der Altarraum empirisch zugänglich wird, weitgehend ausgeblendet. Sequenzen und Bilder können in einem solchen Kontext zur Verdeutlichung bestimmter Aspekte relativ frei von hinten nach vorne gezogen werden und umgekehrt. Während hier also eher das Kriterium der Vollständigkeit und der Sichtbarmachung interaktionsarchitektonischer Relevanzen für die Ordnung der Standbilder ausschlaggebend ist, spielen bei der Frame-Comic-Konstitution und -Analyse die Zeitlichkeit der Interaktionsentwicklung und die zwischen einzelnen Frames bestehenden zeitlichen Beziehungen eine zentrale Rolle.

Die Bezeichnung 'Frame-Comic' ist daher ausschließlich für sequenziell konstituierte Sekundärdokumente aus Standbildern eines audiovisuellen Interaktionsdokuments reserviert, die eine konsistente Handlungsstruktur aufweisen.

7. Visuelles Segmentieren

Im Folgenden will ich zum einen die interaktionstheoretischen Grundlagen visueller Segmentierung klären; zum anderen will ich die einzelnen Arbeitsschritte skizzieren, die dem visuellen Segmentieren als methodischem Verfahren eigen sind.

¹² Der gezeichnete Comic ist letztlich nichts anderes als eine motivierte Standbildfolge, die eigentlich ein Video voraussetzt, aus dem der Zeichner seine Bilder auswählt. Ob Comic-Zeichner tatsächlich ein Video „im Kopf“ haben, weiß ich natürlich nicht.

7.1 Interaktionstheoretische Voraussetzungen

Segmentierung ist ein analysepraktisches Verfahren, bei dem ein größerer Interaktionsausschnitt zum Zwecke seiner besseren Analysierbarkeit in kleinere Teile gegliedert wird. Die visuelle Segmentierung orientiert sich an Relevanzen des Videoausschnitts und geht auf der Grundlage des Postulats der theoretischen Egalität aller Ausdrucksressourcen davon aus, dass das visuell wahrnehmbare Verhalten der Beteiligten – genau wie das akustisch wahrnehmbare – bereits segmentindikative Qualität besitzt. Die Interaktionsbeteiligten produzieren in ihrem visuell wahrnehmbaren Verhalten Hinweise darauf, in welchen sinnvollen Einheiten, Mikrozusammenhängen und Aktivitätseinbindungen sie sich aktuell verhalten. Diese Hinweise werden im Rahmen der allgemeinen Intersubjektivitätskonstitution und damit zur Absicherung der Verstehbarkeit individueller Verhaltensweisen in den Relevanzen der praktischen Zwecke der Interaktion realisiert. Zu solchen Hinweisen gehören beispielsweise

- der markante Wechsel der Körperpositur,
- die Veränderung der räumlichen Position,
- der Wechsel der Präsenzform,
- das Einfrieren von Arm- und Handhaltungen im Kontext gestikulativer Aktivitäten oder
- die blicklichen Fixierungen von Interaktionsbeteiligten oder Objekten.

Bei der Betrachtung eines Videoausschnitts können auch Analytiker/innen diese Hinweise nicht nur identifizieren. Es muss sogar möglich sein, die visuelle Segmentierung – genau wie die Beteiligten selbst – ohne größere analytische Anstrengung vorzunehmen. Es kommt bei der visuellen Segmentierung nicht darauf an, in einem analytisch gehaltvollen und umfassenden Sinne zu verstehen, worum es den Beteiligten in der zu analysierenden Situation geht. Es geht vielmehr darum, aus der „oberflächennahen Wahrnehmung“¹³ ihrer Verhaltensweisen eine Intuition für die von ihnen produzierten Segmentierungshinweise zu entwickeln. Dabei muss man sich fragen, mit welchen Ausdrucksressourcen und Verfahren sie sich – unter welchen situativen Bedingungen – wechselseitig relevante Segmentierungshinweise geben. Was einzelne Ausdrucksweisen konkret bedeuten, ist dabei nicht von primärer Wichtigkeit. Es geht bei der visuellen Segmentierung vielmehr darum, sich für allgemeine Segmentierungsprinzipien – wie beispielsweise die Produktion und Funktionalität von Kontrastivität – zu sensibilisieren.

¹³ Ich spreche hier bewusst von Wahrnehmung und nicht etwa von Beobachtung, um damit die „Oberflächennähe“ der von den Beteiligten produzierten Verhaltensaspekte zu fokussieren, denen man aus analytischer Perspektive die Qualität von „Segmentierungshinweisen“ zuschreiben kann.

Die grundlegende Prämisse einer solchen Sensibilisierung lässt sich wie folgt formulieren: Sollen diese Segmentierungshinweise im interaktiven Vollzug unter online-analytischen Bedingungen und unter Orientierung auf praktische, situative Handlungsziele funktionieren, müssen sie ohne größere analytische Anstrengung sichtbar und verstehbar sein. Dies gilt auch für die extrakommunikative Perspektive der Analytiker. Es kommt in diesem Zusammenhang beispielsweise darauf an, Stellen als Segmentierungshinweise zu befragen,

- an denen sich das Verhalten einzelner Beteiligter sichtbar ändert,
- an denen ein Wechsel von Dynamik zu Statik oder umgekehrt stattfindet und
- an denen sich Verdichtungen bzw. die Synchronisierung von Ausdrucksressourcen finden lassen („Modalitätssynchronisierung“ im Sinne von Putzier (2011) bzw. „multimodal density“ (Norris 2004)).

Das Ergebnis der Segmentierung eröffnet dann die Möglichkeit, alle Frames, die an den Schnittstellen von zwei Videosegmenten liegen, als sequenzierte Standbildreihe zusammzusetzen, um die wesentlichen visuell wahrnehmbaren Ausdrucksveränderungen zu identifizieren. Dies ist ein praktikabler Versuch, die zunächst intuitiv vorgenommene Segmentierung durch die Identifikation segmentierungsrelevanter Verhaltensaspekte für zukünftige Segmentierungen des visuell wahrnehmbaren Verhaltens zu methodisieren. Man darf sich den Versuch der Methodisierung allerdings nicht so vorstellen, dass man irgendwann zu einer Liste kommt, die man – ungeachtet der Spezifik des zu analysierenden Interaktionsdokuments – in einer bestimmten Reihenfolge abarbeiten kann. Gleichwohl werden wiederkehrende Beobachtungen gemacht, die beispielsweise in Abhängigkeit von der Präsenzform, der Beteiligtenstruktur, der interaktionsräumlichen Konstitutionsleistungen etc. zu einer wichtigen Wissensgrundlage sedimentieren. Diese ist nicht formalistisch als Relevanzstruktur der Bearbeitung einzelner Aspekte organisiert, sondern reflektiert gewissermaßen auf der Grundlage einer „motivierten Heuristik“ generelle Aspekte der jeweiligen Interaktionsdokumente.

7.2 Methodisches Vorgehen

Wie genau funktioniert nun die visuelle Segmentierung eines Videoausschnitts? Geht man von dem Fall aus, dass das audiovisuelle Interaktionsdokument bereits vorliegt und schließt damit alle vorgängigen Schritte aus, die z.B. für die Felderschließung und die Datendokumentation faktisch realisiert sein müssen, ereignet sich die visuelle Segmentierung zu einem Zeitpunkt, zu dem bereits die Selektion und Definition eines für die Analyse relevanten Videoausschnitts und die Entscheidung für das Verfahren der visuellen

Erstanalyse erfolgt sind. Die visuelle Segmentierung gliedert sich dann in unterschiedliche Arbeitsschritte. Ich selbst gehe folgendermaßen vor:

- 1) Zunächst schaue ich mir den definierten Videoausschnitt ohne spezifische Fokussierung an. Dabei ist eher eine auf die Gestalthaftigkeit und Segmenthaftigkeit des Geschehens ausgerichtete Wahrnehmung und nicht die Fokussierung einzelner Verhaltensaspekte der Beteiligten relevant.
- 2) Dann stoppe ich – bewusst voranalytisch – das Video immer dann, wenn ich den Eindruck habe, dass sich etwas für meine spezifischen Erkenntnisinteressen Relevantes ereignet bzw. verändert. Den bei diesem intuitiven Stoppen produzierten Frame speichere ich dann als Standbild ab. Wichtig ist dabei, dass man in der anfänglichen voranalytischen Segmentierungsorientierung bleibt und nicht etwa beginnt, nach dem ersten Schnitt darüber nachzudenken, warum man gerade an dieser Stelle das Video gestoppt hat. Hat man sich auf diese Weise durch den ausgewählten Ausschnitt bewegt, liegen damit alle intuitiv ausgewählten Kandidaten für den Frame-Comic vor.
- 3) Als nächster Schritt schließen sich – genau wie beim Transkribieren – Korrekturgänge an. Der erste Segmentierungsgang wird mehrfach wiederholt, wobei sich das Procedere des ersten Gangs wiederholt: Wieder stoppe ich das Video an relevant erscheinenden Stellen intuitiv und extrahiere an den jeweiligen Haltestellen einen Frame.
- 4) Danach vergleiche ich die Schnittstellen und die Anzahl der bei den einzelnen Gängen produzierten Frames. Hat man es dabei, was durchaus vorkommen kann, mit einer gewissen Streuung zu tun, muss man entscheiden, welche der extrahierten Frames Eingang in den Comic finden sollen.
- 5) Da die Grundlage der bisherigen Schritte in der Person des Segmentierers liegen und damit allesamt zwar methodisch motiviert, jedoch von ihrem Status her subjektiv sind, ist es nötig, den Frame-Comic zu intersubjektivieren. Auch diesen Arbeitsschritt kennen wir von der Transkripterstellung, bei denen unterschiedliche Korrekturgänge von unterschiedlichen Transkribenten durchgeführt werden. Ich selbst lasse die von mir konstituierten Frame-Comics im Hinblick auf ihre Konsistenz im Kolleg/innen-Kreis diskutieren und evaluieren.
- 6) Bei der kollektiven Absicherung der Frames wurde deutlich, dass wir in einer Gruppe von sechs Personen sehr schnell ein hohes Maß intersubjektiver Sicherheit hinsichtlich der Relevanz der Frames erreicht hatten. Diese bezog sich auch auf den einzigen Frame, der in einem ersten Durchgang als „nicht passend“, „irgendwie anders“ oder als „nicht so wie die anderen“ charakterisiert wurde. Dass es diesen einen „falschen Frame“ gab,

war ein ausgesprochener Glücksfall. Er verdeutlichte nämlich in seiner kontrastiven Qualität nicht nur seine dokumentarische Andersartigkeit. Vielmehr stellte er die Frage nach den Kriterien für die Homogenität der anderen Frames.

- 7) Der wichtigste Erkenntnisgewinn und damit auch ein erster Ansatzpunkt zur Methodisierung der visuellen Segmentierung als Verfahren ist jedoch die Reflexion der De-facto-Methodologie, das heißt das Bewusstmachen der Kriterien, die bei der intuitiven Entscheidung in systematischer Weise relevant sind. Dieser Punkt öffnet die Tür zur Reflexion der visuell-segmentalen Kompetenz des segmentierenden Analytikers.
- 8) Es ist nicht überraschend, dass es bei den einzelnen Frames, die bei den unterschiedlichen Segmentierungsgängen extrahiert wurden, nicht immer den exakt gleichen Treffer gab. Die Streuung lag jedoch in einem Bereich, der bei allen Frames die wesentliche Veränderung, die im Video zu sehen war, fokussierte. Es ist also keine Frage der Phänomenqualität, die zu der Streuung führte. Diese resultierte vielmehr aus der jeweils unterschiedlichen Schnelligkeit, mit der bei den einzelnen Segmentierungsgängen per Mausclick jeweils auf dasselbe Phänomen reagiert wird.
- 9) Der nächste Schritt besteht dann darin, die gestreuten Frames der jeweiligen Schnitte miteinander zu vergleichen. Dieser Frame-Vergleich ist bezogen auf die Qualität und das Erkenntnisinteresse konstitutionsanalytischer Investigation voranalytisch und auf die Abbildqualität des wahrgenommenen, für die Selektion relevanten Phänomens bezogen. In der Regel ist es unproblematisch, sich für den Frame in einem Streuungsbereich zu entscheiden, der den besten dokumentarischen Wert besitzt.

Man kann bei den zurückliegenden Ausführungen verschiedentlich Gemeinsamkeiten bei der Erstellung von Frame-Comics und Verbaltranskripten erkennen. Für seriös arbeitende Konstitutionsanalytiker ist es fraglos, aufgrund von Unvollständigkeiten und Ungenauigkeiten bei der Wiedergabe des Gehörten, von Falschidentifizierungen von Sprechern, „falschen Freunden“ aufgrund zu vorverständigen Hörens etc. nicht mit Erstversionen von Transkripten zu arbeiten. In der Regel durchläuft ein Transkript, ehe es empirische Grundlage konstitutionsanalytischer Rekonstruktion wird, verschiedene Korrekturgänge. Diese mehrfache Korrektur wird fraglos verstanden als eine notwendige Investition zur Qualitätssicherung des Transkripts. Niemand würde auf die Idee kommen, aus der Korrekturnotwendigkeit ein Argument gegen den Status und die Bedeutung von Transkripten als methodisch konstituierte empirische Grundlagen für die Untersuchung monomodal-verbaler Fragestellungen zu entwickeln. Das Gleiche gilt auch für den Frame-Comic. Natürlich spielten bei den notwendigen Korrekturen auf dem Weg des Frame-

Comics zu seinem Status als methodisch konstituierte empirische Grundlage für die Untersuchung multimodaler Fragestellungen andere Formen der Korrektur eine Rolle.

Diese beziehen sich beispielsweise auf folgende Fragen:

- Ist die Anzahl der Frames ausreichend, um die visuelle Repräsentation der Homogenität des Aktivitätszusammenhangs zu gewährleisten?
- Sind die Frames tatsächlich nach dem Kontinuitäts- und Kontiguitätsprinzip ausgewählt worden?
- Ist bei den einzelnen Segmentierungsentscheidungen jeweils der dokumentarisch „wertvollste“ Frame ausgewählt worden?

Wie bei der Erstellung eines Transkripts, so werden die Korrekturaktivitäten im Zusammenhang mit der Erstellung eines Frame-Comics im Modus einer voranalytischen Haltung durchgeführt. Sie sind selbst nicht Teil der konstitutionsanalytischen Rekonstruktion, sondern notwendige Vorarbeiten.

8. Das Verhältnis von Transkript und Frame-Comic

Die systematische Reflexion des Verhältnisses von verbalem und multimedialem Transkript und dem Frame-Comic ist eine wichtige Aufgabe der weiteren Entwicklung der multimodalen Interaktionsanalyse und eine eigenständige, systematische Fokussierung wert. Diese kann im hiesigen Zusammenhang nicht geleistet werden. Gleichwohl will ich kurz und auf der Grundlage eines nur schwer aufzuholenden Vorsprungs in der wissenschaftlichen Praxis des Transkribierens verbaler Aktivitäten, der Systematisierung der über die Jahrzehnte hinweg produzierten Ergebnisse und der Reflexion der dem Transkribieren inhärenten Implikationen einen für meine Darstellungszwecke kurzen, jedoch ausreichenden Vergleich angehen.

Das Verhältnis des Frame-Comics zu klassischen Verbaltranskripten und den unterschiedlichsten Varianten multimedialer Transkripte¹⁴ ist gleichermaßen einfach und komplex.

Es ist einfach, weil es sich bei dem Frame-Comic – wie beim klassischen monomodalen Verbaltranskript und beim Multimedial-Transkript – um ein zur empirischen Bearbeitung spezieller Fragestellungen angefertigtes Sekundärdokument handelt.

Es ist komplex, weil der Frame-Comic – aufgrund seiner Konstitutionslogik und seiner besonderen empirischen Funktionalität – kein Transkript ist. Man darf also nicht den Fehler begehen, den Frame-Comic als Endpunkt eines

¹⁴ Siehe Mondada (i.d.Bd.) zur Frage, wie Informationen zum visuell wahrnehmbaren Verhalten und Bilder in die Transkripte kamen.

Kontinuums zu verstehen, das im Rahmen konversationsanalytischer Erkenntnisinteressen in Form unterschiedlicher Typen von Transkripten entstanden ist. Ein solcher Fehlschluss würde den Frame-Comic – auf der Grundlage des dokumentkonstitutiven Fehlens von Verbalität – als „verbalitätsfreies“ Transkript einordnen.

Transkripte (egal, ob in ihnen nur auditiv oder auch visuell wahrnehmbares Verhalten notiert wird) und Frame-Comics sind diesbezüglich kategorial unterschiedlich. Sie gehen nicht etwa derart ineinander über, dass es immer weniger Verbalität gibt, bis sie beim Frame-Comic gänzlich verschwindet.

Beim Versuch, die beiden Sekundärdokumente zu vergleichen, stellt sich die Frage nach den relevanten Kriterien, auf deren Grundlage die beiden Sekundärdokumente zueinander in Beziehung gesetzt werden können. Für den hiesigen Zusammenhang habe ich mich für die drei folgenden Basiskriterien entschieden:

- Vollständigkeit der Abbildung (Kap. 8.1),
- Sequenzialität der Abbildung (Kap. 8.2),
- Repräsentation von Zeitlichkeit (Kap. 8.3).

Aufgrund der Tatsache, dass Verbaltranskripte einen deutlichen wissenschaftshistorischen Vorsprung gegenüber rein visuell konstituierten Sekundärdokumenten haben, fällt das Augenmerk zudem stärker auf die Reflexion, die dieser Vergleich für den Frame-Comic besitzt.

8.1 Vollständigkeit der Abbildung

Beim Vergleich der beiden Sekundärdokumente stellt sich sofort die Frage danach, was die Orientierung an der vollständigen Dokumentation des zu analysierenden Geschehens für das Verbaltranskript und den Frame-Comic jeweils bedeutet.

Bei der Erstellung eines Verbaltranskripts orientiert man sich daran, im Prinzip alles, was auditiv wahrnehmbar ist und für die spätere Analyse als relevant eingeschätzt wird, zu erfassen, also z.B. nicht nur zu transkribieren, *was* gesprochen wird, sondern auch und gerade *wie* gesprochen wird. In gewisser Hinsicht ist das Ideal eine Duplizierung des Gesprochenen in Gestalt eines schriftlich fixierten Textes. Der Testfall für das Gelingen dieses Versuchs ist die maximale (Wieder-)Annäherung an das Primärdokument bei der Re-Transformation (dem mündlichen Vortrag) des transkribierten Textes unter Berücksichtigung aller kodierten Informationen.

Der Versuch, die im Primärdokument festgehaltene Interaktion möglichst exakt, detailreich und vollständig im Prozess ihrer zeitlichen Entwicklung ab-

zubilden, erfolgt nach dem Prinzip „order at all points“ (Sacks 1984).¹⁵ Für die schriftliche Abbildung des Gesprochenen werden Symbole definiert und in Form von Transkriptionssystemen systematisch zusammengestellt. Im sprachwissenschaftlichen Kontext gehen alle aktuellen Transkriptionssysteme (beispielsweise Selting et al. 1998, 2009) mehr oder weniger direkt auf die „Urversion“ von Jefferson zurück, die erstmalig in Sacks et al. (1974) publiziert wurde (siehe auch Jefferson 2004).

Für die Erstellung eines Frame-Comics spielt Vollständigkeit als relevante Orientierung ebenfalls eine zentrale Rolle. Anstatt jedoch das Interaktionsgeschehen entlang der Zeitlichkeit seiner Entwicklung vollständig und detailliert zu dokumentieren, geht es beim Frame-Comic um das Erfassen relevanter Veränderungen der Interaktionsentwicklung. Will man also auch in diesem Zusammenhang von einer Vollständigkeitsorientierung sprechen, kann es beim Frame-Comic nur darum gehen, relevante Veränderungen auf der Grundlage einer aussagekräftigen Kontrastivität, die in einzelnen Frames dokumentiert wird, möglichst vollständig festzuhalten. Es handelt sich also um eine spezifische Variante von Vollständigkeit, für die das systematische Weglassen von nicht aussagekräftigen Frames konstitutiv ist. In diesem Sinne handelt es sich – im Gegensatz zur Transkription und im Gegensatz zur Audio- und Videodokumentation – nicht um einen „registrierenden“ Modus der Datenaufbereitung, sondern im Sinne von Bergmann (1985) um eine spezifische Form „methodischer Fixierung sozialer Wirklichkeit“.

Die Konstitutionslogik des Verbaltranskripts besteht in der weitgehend interpretationsfreien, vollständigen Überführung des Hörbaren ins Lesbare. Entsprechend fordern neuere Transkriptionsverfahren, die in der Frühphase der Gesprächsanalyse oft verwendeten interpretativen Kommentare zur Sprechweise (wie „ironisch“ oder „empört“) durch objektivierbare Parameter zu ersetzen (vgl. GAT 2.0). Dabei stellen die Gesamtheit des einzelsprachlichen Wissens und die korrekte Umsetzung der in den jeweiligen Transkriptionsvorschriften festgelegten Transformationsregeln die zentralen Voraussetzungen dar. Diese sind dafür verantwortlich, dass unter bestimmten akustischen Bedingungen und einer gewissen Konzentration der Transkribenten zumindest der Wortlaut und die basale segmentale Struktur des Gesagten als intersubjektiv gesichert betrachtet werden können.

Die Konstitutionslogik eines Frame-Comics verfügt (noch) nicht über eine vergleichbare Intersubjektivitätsgrundlage. Diese muss erst entwickelt werden (siehe aber oben). Im Vergleich zum Transkribierenden besteht die Kompetenzgrundlage des visuell Segmentierenden in der Wahrnehmung relevanter Verhaltensveränderungen von zentralen Beteiligten. Dies kann dazu

¹⁵ Zu den Prinzipien, denen bei Verbaltranskripten gefolgt wird, siehe beispielsweise Deppermann (1999, S. 46-48).

führen, dass die Segmentierungsentscheidung unterschiedlicher Segmentierender stärker als bei der Transkription des Wortlauts divergiert.

Zudem verändern sich die für die Segmentierung relevanten Aspekte und damit letztlich auch die Größe der Binnensegmente in Abhängigkeit vom jeweiligen Erkenntnisinteresse. Das hängt im Wesentlichen mit der Relevanz der Ausdrucksressource zusammen, auf deren Grundlage die Segmentierung primär durchgeführt und der Frame-Comic konstituiert wird.

8.1.1 Spezifik/Relevanz der Ausdrucksressource

Besteht die dominante Ressource in auditiv wahrnehmbarem Verhalten, wirken sich mögliche Implikationen des Erkenntnisinteresses in erster Linie auf die Granularität aus, in der die Transkription erfolgt. Dies stellt beispielsweise GAT (Selting et al. 2009) systematisch durch unterschiedliche Detaillierungsniveaus (Minimal-, Basis- und Feintranskript) in Rechnung, die in Bezug auf prototypische Erkenntnisinteressen ausgearbeitet worden sind.¹⁶ Die grundsätzliche Orientierung an der verbalen Vollständigkeit ist jedoch die gemeinsame Grundlage aller Detaillierungsniveaus.

Besteht die dominante Ressource in visuell wahrnehmbaren Verhaltensweisen, wirkt sich das spezifische Erkenntnisinteresse unmittelbar auf die Selektion unterschiedlich weit auseinanderliegender Frames aus. Ist man beispielsweise an der Positionierung der Beteiligten im Raum interessiert, wird man nur dann entsprechende Frames extrahieren, wenn es tatsächlich räumliche Dynamik und Neupositionierungen gibt. Konzentriert man sich hingegen auf Gestikulation, wird man beispielsweise bei der Segmentierung Wechsel von links- zu rechtshändiger, von ein- zu beidhändiger, von reduzierter zu expressiver Gestikulation berücksichtigen. Da Gestikulieren eine wesentlich dynamischere Ausdrucksressource darstellt als etwa Positionierung im Raum oder auch die Veränderung von Posituren, wird man als Ergebnis der gestikulationsbezogenen Segmentierung desselben Videoausschnitts gänzlich andere und vor allem wesentlich mehr Frames produzieren und dadurch einen gänzlich anderen Frame-Comic konstituieren.

Im Bereich der Verbaltranskripte wurde über die Jahrzehnte hinweg eine relativ hohe Standardisierung erreicht. Demgegenüber zeichnet sich die aktuelle Entwicklung multimedialer Transkripte durch eine hohe Varianzbreite mit teilweise idiosynkratischen Lösungen für bestimmte Abbildprobleme aus.¹⁷

¹⁶ Im Kontext der 'Interaktionalen Linguistik' wurde beispielsweise der Bereich der Abbildung prosodischer und rhythmischer Phänomene ausgebaut (beispielsweise Auer/Couper-Kuhlen/Müller 1999).

¹⁷ Siehe auch Stukenbrock (2009) und vom Lehn (2014) zu den Anforderungen an multimediale Transkripte.

Es gibt Transkripte, die dem Prinzip ‘visueller Additivität’ folgen (beispielsweise Hausendorf/Schmitt i.d.Bd.). Bei solchen Transkripten werden Standbilder, die aus den zugrundeliegenden Primärdokumenten extrahiert werden, zur Visualisierung relevanter körperlich-räumlicher Informationen eingesetzt. Auf die Verbalisierung dieser Informationen im Transkript wird verzichtet. Diese werden bei der analytischen Beschreibung nachgeliefert. Solchen Lösungen ist eine Tendenz der Diffundierung des Transkripts in den analytischen Beschreibungstext zu eigen. Dies wird – als „kleiner Tod“ – im Hinblick auf eine bessere Lesbarkeit in Kauf genommen.

Dieser Lösung stehen Transkripte gegenüber, die nach dem Prinzip der ‘Verbalisierung visuell wahrnehmbarer Aspekte’ konstituiert werden. Hier ist die klassische Untersuchung zur Blickorganisation von Goodwin (1981) zu nennen, der noch gänzlich ohne Bilder auskam. Mondada arbeitet bei ihren Untersuchungen mit sehr detaillierten Transkripten, in denen sie Standbilder aus den Primärdokumenten und Verbalisierung kombiniert. Sie notiert so alle für die nachfolgende Analyse relevanten audio-visuell wahrnehmbaren Informationen (beispielsweise Mondada 2015). Solchen Lösungen ist eine Tendenz zur informativen Überfrachtung zu eigen, die im Zweifelsfalle dazu führt, die detaillierte empirische Grundlage bei der Lektüre zu überspringen.

Transkripte unterscheiden sich zudem hinsichtlich der Frage, ob (und wenn ja, in welchem Ausmaß) sie Standbilder aus den Primärdokumenten durch verschiedene Zusatzinformationen aufbereiten: durch Pfeile und Sprechblasen (Norris 2004), eingefügten Text (Goodwin 1994), Hervorhebungen (Schmitt 2012b) oder auch Musiknoten (Erickson 2004). Mittels solcher Markierungen wird – durchaus gewollt – interpretative Qualität in die empirischen Grundlagen hineingeschrieben. Solchen Lösungen ist eine Tendenz zur interpretativen Aufladung der empirischen Grundlagen zu eigen, die eigentlich das „interpretationsfreie“ Dokument der Analyse darstellen sollen.¹⁸

8.1.2 Kontinuität und Kontiguität

Die Systematik und die Schnittweite der Extraktion von Frames aus dem Video basiert auf zwei unterschiedlichen Aspekten einer gemeinsamen Orientierung: Zum einen zeigen die Frames den durch die Fragestellung relevant gesetzten Aktivitätszusammenhang in der ausschnitthaften Kontinuität seiner Entstehung. Zum anderen schließen alle Frames, die aufeinander folgen, abbildlogisch auch tatsächlich aneinander an.

¹⁸ Weiterhin realisieren dieselben Autoren in Abhängigkeit von ihren jeweiligen Adressaten, bezogen auf die gleichen Daten, unterschiedliche Abbildverfahren.

Die Abfolge der Frames ist so organisiert, dass – beispielsweise bezogen auf räumliche Dynamik und damit verbundene Positionierungen – alle Veränderungen der Positionen von Beteiligten im Raum in den Frame-Comic integriert werden. Diese Kontinuitätsorientierung gewährleistet, dass man als Analytiker die Sicherheit hat, zwei benachbarte Frames als sequenziell zusammengehörig betrachten zu können. Dies gilt natürlich nur in der Logik des Frame-Comics und nicht in der des Transkripts. Man muss also nicht fragen, ob sich die abgebildeten Personen „zwischen den Frames“ etwa an einem andern Ort aufgehalten haben und von dort wieder zurückgekehrt sind.

Die Kontiguität zweier Frames kann zeitlich sehr unterschiedlich sein. Sie kann bei einer interaktiven Eröffnungssequenz im Bereich von Sekundenbruchteilen (Schmitt/Knöbl 2014) und bei der Analyse von Raumnutzungsformen im Bereich mehrerer Minuten liegen (Dausendschön-Gay/Schmitt i.d.Bd.). Kriterien für die Konstitution beider Fälle sind übereinstimmend die Orientierung an der Kontinuität der Interaktionsentwicklung und der damit verbundenen Vermeidung von „Sprüngen“.

8.1.3 Extraktionsweite und Seh-Arten-Potenzial

Beim Frame-Comic existiert ein direkter Zusammenhang zwischen der Extraktionsweite der Frames und der Anzahl und Varianz möglicher Seh-Arten hinsichtlich dessen, was zwischen zwei Frames an interaktivem Vollzug wahrscheinlich realisiert worden ist. Je weiter die Schnitte zeitlich auseinanderliegen, desto größer wird die Anzahl möglicher Seh-Arten für das nicht sichtbare zwischenzeitliche Verhalten. Dieser systematische Zusammenhang zwischen Extraktionsweite und „interpretativer Sicherheit“ hat Konsequenzen für Fragestellungen und Erkenntnisinteressen, die auf der Grundlage von Frame-Comics verfolgt werden können.

Mit einer großen Extraktionsweite lässt sich nur eingeschränkt und letztlich spekulierend der tatsächlich zwischen zwei Frames liegende interaktive Vollzug „rekonstruieren“. Maximale Extraktionsweite ist daher eher geeignet, nicht-dynamische und vom kontinuierlichen Vollzug unabhängige(re) Phänomene zu untersuchen. Hierzu zählen beispielsweise die Implikationen von Position und Positur von im Raum Anwesenden zu anderen oder die Nutzung des Raums durch Anwesende aufgrund ihrer Positionierung in dem durch die Architektur zur Verfügung gestellten Möglichkeitsrahmen. Solche Frames fordern und fördern die Rekonstruktion sozialtopografischer Implikationen, ohne durch Vollzugsdynamik abgelenkt zu werden. Frames, die hierfür extrahiert werden, repräsentieren eher statische Momente, Augenblicke der Ruhe, Transitionen am Umschlagplatz von Abgeschlossenem und noch nicht neu Begonnenem sowie Start- und/oder Endpunkte von Bewegungen.

Ein instruktives Beispiel hierfür sind die beiden folgenden Frames, die unter dem Aspekt der individuellen Raumnutzung in ihrer Statik und symbolischen Eigenständigkeit funktional sind:¹⁹



Die beiden Frames aus dem Alpha-Gottesdienst zeigen dieselbe Person in zwei unterschiedlichen Positionen und Positionierungen, für die eine maximal kontrastierende Raumnutzung und ein spezifisches Beteiligungsformat charakteristisch sind. Um sich dafür zu interessieren, ist es weder nötig zu wissen, wie die konkrete zeitliche Beziehung der beiden Frames aussieht, noch, wie die Person vom Sitzplatz zum Mikro gelangt ist. Interessiert man sich jedoch für die Rekonstruktion der Laufwege und die Art des Gehens, kommt man mit diesen beiden Frames nicht weit. Man muss vielmehr neue Frames extrahieren, die eine Beschäftigung mit dem Aspekt „Gehen als situierte Praktik“ ermöglichen. Das sieht dann im vorliegenden Fall wie folgt aus:



¹⁹ Diese beiden sowie die nachfolgenden Standbilder stammen aus der Fallanalyse von Hausendorf/Schmitt (i.d.Bd.).



8.1.4 Zunehmende Detaillierung

Ein weiterer Unterschied zwischen Frame-Comic und Verbaltranskript besteht also offensichtlich in den Auswirkungen, die mit der Möglichkeit verbunden sind, das Detaillierungsniveau der Sekundärdokumente zu verfeinern.

Wie man ein Basistranskript mit seiner Konzentration auf wenige äußerungs- und interaktionsstrukturelle Phänomene durch die Integration suprasegmentaler Informationen schrittweise anreichern kann, so kann man auch in einen bereits existierenden Frame-Comic neue Frames einfügen. Zieht man jedoch zwischen zwei Frames neue Frames ein, geht das in der Regel nicht nur mit einer visuellen Anreicherung einher, sondern zuweilen mit der Konstitution eines neuen Dokuments.

Der zentrale Unterschied ist also, dass sich bei einem Verbaltranskript die Detaillierung auf der Grundlage einer letztlich unveränderten Basisstruktur (in der Regel ist das der gesprochene Wortlaut) vollzieht. Folglich gibt es bei der Detaillierung keinen neuen Wortlaut zu hören, sondern das bereits gehörte und schriftlich Fixierte wird durch neue Informationen spezifiziert. Beim Frame-Comic hingegen entsteht ein neuer visueller Eindruck, der das vorher fixierte visuelle Interaktionsgeschehen in offensichtlicher und teilweise signifikanter Weise verändert.

Natürlich wählen wir auch bei der Arbeit mit Transkripten aus. Bestimmte Segmente, die für die Rekonstruktion der Handlungsstruktur weniger wichtig sind, landen in Beschreibungen, die zum nächsten Segment überleiten. Die analysierten Segmente sind oft dadurch charakterisiert, dass sich in ihnen markant etwas an der Handlungsstruktur ändert: Eine Erzählung beginnt, eine bestimmte Positionierung startet, ein Anwesender ändert erkennbar seine interaktive Beteiligung etc.²⁰

8.2 Sequenzialität (der Abbildung)

Ein weiterer Aspekt beim kontrastiven Vergleich der beiden Sekundärdokumente ist die Ermöglichung eines sequenziellen Vorgehens bei der Analyse. Zumindest im ethnomethodologischen Zusammenhang ist es unstrittig, ein Verbaltranskript entlang seiner zeitlichen Entwicklung möglichst kleinschrittig zu rekonstruieren und nicht wahllos von vorne nach hinten zu springen. Die gleiche Unstrittigkeit gilt auch für den Frame-Comic. Auch hier muss das Dokument ermöglichen, das relevant gesetzte Phänomen in der Zeitlichkeit seiner Entstehung sequenzanalytisch zu rekonstruieren, ohne dass die Entwicklung von Seh-Arten bei Frame 1 bereits auf der Grundlage des Wissens um nachfolgende Frames eingengt wird. Vielmehr wird durch die umfassende und detaillierte Rekonstruktion der zeitlich geordneten, akustisch und visuell wahrnehmbaren Phänomene ein Zugang zu den generativen Mechanismen der Interaktionskonstitution eröffnet. Dadurch wird die von den Beteiligten realisierte Methodologie ihres Alltagshandelns sichtbar.

Beim Verbaltranskript werden durch die Vollständigkeit aller hörbaren Phänomene im Transkript Zusammenhänge deutlich, die sich nur durch die im Sekundärdokument sichtbare Segmentalität und den Bezug einzelner Aspekte aufeinander zeigen. Bei der Analyse des Frame-Comics liegt die erkenntnisgenerierende Produktivität gerade darin, die methodisch gewollte Reduktion explizierend aufzufüllen. Zum einen soll dadurch im Rahmen der Fallspezifik möglichst präzise ausgelotet werden, was „zwingend“ zwischen zwei Frames geschehen sein muss und was motivierte Anschlussersparungen des aktuellen Frames sind. Zum anderen geht es um die Reflexion der Annahmen, Idealisierungen und Erwartungen, die für diese fallspezifische Rekonstruktionsarbeit bestimmend sind. Es ist vor allem dieser letzte Punkt, der uns zwingt, unser eigenes sozialtopografisches Wissen zu explizieren, um dieses dann in einem zweiten Schritt zu intersubjektivieren. Um dies zu gewährleis-

²⁰ Mit bestimmten Einschränkungen kann man also sagen: Auch im Transkript gibt es markante Stellen, die sich wie bestimmte Frames im Frame-Comic als konstitutive Segmente für die Analyse anbieten.

ten, ist jedoch zunächst die autonome Analyse der einzelnen Frames notwendig, der dann die Fokussierung von Veränderung von Frame zu Frame folgt (Schmitt i.Vorb., Kap. 7.6).

8.3 Repräsentation von Zeitlichkeit

Beim Verbaltranskript wird durch die vollständige Verschriftung des Interaktionsgeschehens eine implizite „Time-Line“ abgebildet, die es ermöglicht, einzelne Äußerungen zeitlich zueinander in Beziehung zu setzen. In der Regel weiß man bei der Analyse eines Verbaltranskripts nicht wirklich exakt, wie viel Zeit seit dem Ende von Äußerung 2 und dem Beginn von Äußerung 4 vergangen ist. Für viele Erkenntnisinteressen ist es ausreichend, sich die Zeitlichkeit über Äußerungslänge und Pausenstruktur erschließen zu können.

Bei einem Frame-Comic geben die einzelnen Frames keine Auskunft über das zeitliche Verhältnis, das zwischen ihnen besteht. Sind die extrahierten Frames sehr ähnlich, kann man ohne Time-Line nicht sicher sein, ob die Frames die Sequenzialität des Interaktionsdokuments widerspiegelt. Hinsichtlich der nachfolgenden drei Frames²¹ sind beispielsweise folgende Sequenzen möglich: 1-2-3, 1-3-2, 2-1-3, 2-3-1, 3-1-2 sowie 3-2-1.



Zeit: 06:57:20



Zeit: 07:05:17

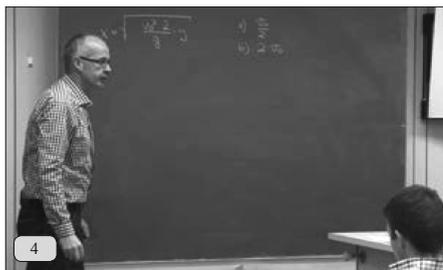
²¹ Die Abbildungen stammen aus Schmitt/Knöbl (2014).



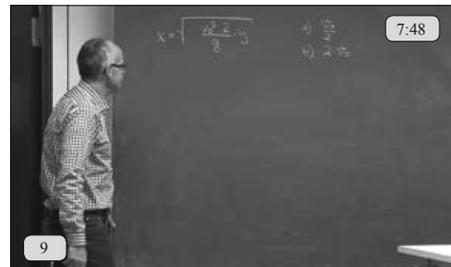
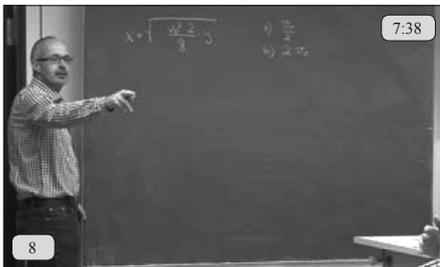
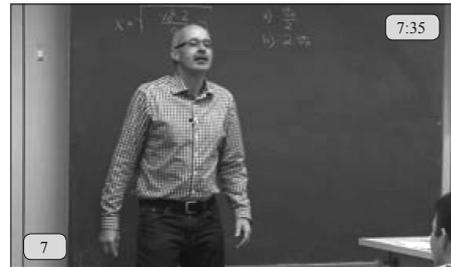
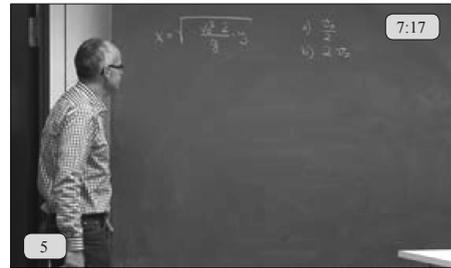
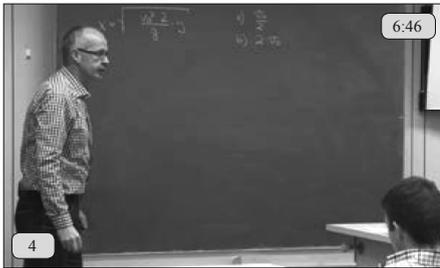
Zeit: 07:11:17

Zeitlichkeit kann nur durch eine Time-Line und die darin repräsentierte sequenzielle Position der Frames abgebildet werden. Das zeitliche Verhältnis der Frames zueinander ist für die Frame-Analyse nicht nur zur Absicherung der richtigen Frame-Abfolge wesentlich. Sie sind vielmehr in einem basalen Sinne Grundlage für angemessene Interpretationen des Frame-Comics.

Die zwei folgenden Standbilder sind sich so ähnlich, dass aufgrund von Kontinuitäts- und Kontiguitäts-Idealisierungen der Eindruck entsteht, sie folgen unmittelbar aufeinander. Eine naheliegende Interpretation könnte sein: Nachdem der Lehrer kurz in die Klasse geschaut hat, orientiert er sich auf die Tafel, indem er seinen Kopf nach links wendet und auch mit seiner linken Hand die Tafel berührt. Wenn man Auskunft über die dabei vergangene Zeit geben sollte, würde man sich wahrscheinlich im Bereich weniger Sekunden bewegen.



Faktisch liegt jedoch zwischen den beiden Frames eine Zeitspanne von über einer Minute! In der Zwischenzeit hat der Lehrer (siehe nachfolgend) sein Pult aufgesucht, hat von dort aus mit Schülern gesprochen und erst danach wieder seinen Platz an der Tafel eingenommen und sich dort stabil verankert.



Die Zeit, die zwischen einzelnen Frames vergangen ist, ist für die Frame-Comic-Analyse nicht nur zur Absicherung der richtigen Frame-Abfolge wesentlich. Sie ist auch für die inhaltliche Analyse der einzelnen Frames und den darin dokumentierten Veränderungen von zentraler Wichtigkeit. Denn die Zeitlichkeit ist entscheidend dafür, wie wir eine Veränderung zwischen zwei Frames interpretieren.

Zum Abschluss dieses Vergleiches der beiden Sekundärdokumente will ich den wesentlichen Unterschied zwischen Transkript und Frame-Comic noch einmal aus der Perspektive der wissenschaftlichen Praxis reformulieren. Transkribieren ist die prozesshaft organisierte wissenschaftliche Praxis (Bezemer/Mavers 2011), durch die auf der Grundlage der vollständigen Verschriftlichung des im Primärdokument Gehörten (und partiell auch des Gesehenen) Sekundärdokumente in Form schriftlicher Texte, den Transkripten, entstehen. Beim Frame-Comic entsprechen dem die wissenschaftliche Praxis des Segmentierens des audiovisuellen Primärdokuments auf der Grundlage visuell

wahrnehmbaren Verhaltens sowie das Auswählen und das sequenzielle Anordnen der beim Segmentieren extrahierten Frames in Form einer handlungskonsistenten Standbildfolge.

Der Frame-Comic ist also zum einen kein verbalitätsfreies Transkript, und er ist zum anderen die methodische Antwort auf Fragen, die sich außerhalb des konversationsanalytischen Erkenntniszusammenhangs stellen.

9. Potenzial und Grenzen des Frame-Comics

Die Funktionalität und Fruchtbarkeit des Einsatzes handlungskonsistenter Bildfolgen ist vor allem für sozialtopografische Analysen im Kontext konsistenter Handlungszusammenhänge – wie etwa bei Warten (Dausendschön-Gay/Schmitt i.d.Bd.) und Freiraum schaffen im Klassenraum (Schmitt/Dausendschön-Gay 2015) evident.

Bei der mikroanalytischen Rekonstruktion kleiner und dynamischer lokaler Handlungszusammenhänge, für die neben dem visuell wahrnehmbaren Geschehen auch das Hörbare konstitutiv ist, stellt sich das etwas anders dar. In solchen Zusammenhängen ist die Funktionalität des Frame-Comics von vornherein eine andere, da dem visuell-erstanalytischen Zugang notwendigerweise die Rekonstruktion des verbalen Geschehens folgen muss, um im Sinne von Goffman (1983) Zugriff auf die multimodal konstituierte „interaction order“ zu bekommen. Hier ist der Frame-Comic also integriert in die grundsätzlich multimodal orientierte Konstitutionsanalyse und aus spezifischen Erkenntnisgründen der Verbalanalyse vorgeordnet.

Ein weiterer Aspekt, der sich in grundsätzlicher Weise auf das Erkenntnispotenzial des Frame-Comics auswirkt, ist die Größe des zugrundegelegten Handlungszusammenhangs und das Erkenntnisinteresse, das darauf bezogen verfolgt wird. Je kleiner der festgelegte Ausschnitt und je statischer die visuelle Seite der Interaktion ist, desto geringer scheint das Erkenntnispotenzial des visuellen Sekundärdokuments zu sein. Vor allem die Intersubjektivität der visuellen Segmentierung und die Rekonstruktion der wahrscheinlichen Anschlussimplikationen nachfolgender Frames schrumpfen hier doch merklich.

Ob sich dieser momentane Eindruck verfestigen wird oder primär damit zusammenhängt, dass es im Moment noch keine dem Transkript vergleichbaren substanziellen Erfahrungen mit dem Frame-Comic gibt, muss zukünftigen empirischen Erfahrungen überlassen bleiben. Es wird vor allem interessant sein zu sehen, wie sich der Frame-Comic im Vergleich mit dem Verbaltranskript, dem Universalwerkzeug für die Analyse verbal konstituierter Ordnungsstrukturen, schlägt. Im Moment ist noch offen, ob die Einsatzmöglich-

keiten der handlungskonsistenten Standbildfolge auf spezifische Erkenntnisinteressen beschränkt bleiben oder ob die gerade erst einsetzende systematische Entwicklung der multimodalen Interaktionsanalyse auch zur allgemeinen „methodischen Aufwertung und Konsolidierung“ des Frame-Comics führen wird. Das sind jedoch letztlich nur sekundäre Relevanzen.

Denn die eigentliche erkenntnisgenerierende Kraft entwickelt der Frame-Comic neben seiner spezifischen Qualitäten als Sekundärdokument über die methodisch-methodologischen Implikationen, die mit diesem Sekundärdokument unweigerlich verbunden sind. Der Frame-Comic zwingt durch die motivierte Aussparung in der Dokumentation und der damit verbundenen Verdichtung der Handlungsdynamik dazu, diese quasi aus jedem Einzelframe zu rekonstruieren. Er verlangt somit, jeden einzelnen Frame als eigenständiges Dokument eigenwertig nicht nur in seiner faktischen Abbildqualität – in seinem Jetzt –, sondern gerade auch in seinem notwendigen Vorher und seinem motivierten Nachher zu analysieren. Im gewissen Sinne ist die deskriptive Rekonstruktion dessen, was es in dem Frame zu sehen gibt, eigentlich nur der Startpunkt zum Generieren weiterführender, analyseleitender Fragen:

Was muss, aufgrund des im aktuellen Frame dokumentierten Geschehens, vorher zwingend passiert sein, und was sind – ganz im Sinne der konversationsanalytischen Vorstellungen – die Anschlussimplikationen dessen, was es aktuell zu sehen gibt: Wie wird/kann es jetzt aufgrund welcher Hinweise weitergehen?

Der Frame-Comic ist mit den Anforderungen, die er an seine Analyse stellt, eine gute und – wie ich meine – notwendige methodisch-methodologische Übung. Er setzt die für die konversationsanalytische Rekonstruktion so zentrale Frage „What is next?“ und die darin zum Ausdruck kommende präferenzielle analytische Orientierung an „Progression der Interaktionskonstitution“ (Stivers/Robinson 2006) durch seine Auslassungen und seine Reduktion auf das visuell Wahrnehmbare im eingefrorenen Format selektiver Standbilder motiviert außer Kraft.

Er verleiht dadurch dem einzelnen Frame einen wichtigen Status, den der Analytiker durch methodische Analyseverfahren respektieren muss. Ein Teil des Respekts wird dadurch erbracht, dass bei der analytischen Rekonstruktion die entwickelten Seh-Arten immer auch als reflexiver Ausdruck der Kompetenzgrundlagen des Analytikers thematisch gemacht werden.

Die zentrale Bedeutung des Frame-Comics besteht nicht nur darin, adäquate empirische Grundlage für bestimmte Fragestellungen zu sein. Der Frame-Comic ermöglicht zudem aufgrund seiner Konstitutionslogik und seiner zentralen Eigenschaften die Ent-Dynamisierung des Primärdokuments und er-

öffnet so einen nur auf den ersten Blick erschwerten methodischen Zugang zu relevanten Aspekten der Raumnutzung und Interaktionskonstitution.

Und – last but not least – stiftet der Frame-Comic dazu an, sich die konversationsanalytische und objektiv-hermeneutische Maxime „order at all points“ im Sinne folgender Frage problematisch zu machen: Ist es möglich bzw. nicht eigentlich naheliegend, die bisherige Relevanz und Gültigkeit der analytischen Maxime „order at all points“ ihrer Bindung an die sequenzielle Analyse des interaktiven Vollzug der Ordnung zu entheben?

Wenn die Fallstruktur tatsächlich ein Prozess ist, der sich in permanenter Reproduktion befindet, dann kann man bezogen auf den Frame-Comic und auf der Grundlage der bislang mit ihm gemachten Erfahrungen in folgender Richtung weiterdenken:

Muss man bei der Rekonstruktion einer Fall- bzw. Ordnungsstruktur notwendigerweise ihre gesamte interaktive Reproduktion sequenzanalytisch aufarbeiten? Reicht es nicht vielmehr aus, aussagekräftige und sequenziell geordnete *points* (extrahiert aus Videoaufnahmen) zu analysieren, um der *order* auf die Spur zu kommen?

10. Literatur

- Auer, Peter/Couper-Kuhlen, Elizabeth/Müller, Frank Ernst (1999): *Language in time. The rhythm and tempo of spoken interaction*. Oxford: Oxford University Press.
- Bergmann, Jörg R. (1982): Schweigephasen im Gespräch. Aspekte ihrer interaktiven Organisation. In: Soeffner, Hans-Georg (Hg.): *Beiträge zu einer empirischen Sprachsoziologie*. Tübingen: Narr, S. 143-184.
- Bergmann, Jörg R. (1985): Flüchtigkeit und methodische Fixierung sozialer Wirklichkeit. Aufzeichnungen als Daten der interpretativen Soziologie. In: Bonß, Wolfgang/Hartmann, Heinz (Hg.): *Entzauberte Wissenschaft. Zur Relativität und Geltung soziologischer Forschung*. Göttingen: Schwartz, S. 299-320.
- Bezemer, Jeff/Mavers, Diane (2011): Multimodal transcription as academic practice: a social semiotic perspective. In: *International Journal of Social Research Methodology* 14, 3, S. 191-207.
- Birdwhistell, Ray L. (1970): *Kinesics and context. Essays on body motion communication*. Philadelphia: University of Pennsylvania Press.
- Deppermann, Arnulf (1999): *Gespräche analysieren*. Opladen: Leske und Budrich.
- Deppermann, Arnulf (2008): Verstehen im Gespräch. In: Eichinger, Ludwig/Kämper, Heidrun (Hg.): *Sprache – Kognition – Kultur*. (= Jahrbuch des Instituts für Deutsche Sprache 2007). Berlin/New York: de Gruyter, S. 225-261
- Deppermann, Arnulf/Schmitt, Reinhold (2009): Verstehensdokumentationen: Zur Phänomenologie von Verstehen in der Interaktion. In: *Deutsche Sprache* 36, S. 220-245.

- Erickson, Frederick (2004): *Talk and social theory: ecologies of speaking and listening in everyday life*. Cambridge: Polity Press.
- Goffman, Erving (1983): The interaction order. In: *American Sociological Review* 48, S. 1-17.
- Goodwin, Charles (1981): *Conversational organization: interaction between speakers and hearers*. London: Academic Press.
- Goodwin, Charles (1994). Professional vision. In: *American Anthropologist* 96, 3, S. 606-633.
- Heidtmann, Daniela (2009): *Multimodalität der Kooperation im Lehr-Lern-Diskurs. Wie Ideen für Filme entstehen. (= Studien zur Deutschen Sprache 50)*. Tübingen: Narr.
- Heilman, Samuel C. (1979): Communication and interaction: A parallel in the theoretical outlooks of Erving Goffman and Ray Birdwhistell. In: *Communication* 4, S. 221-234.
- Jefferson, Gail (2004): Glossary of transcript symbols with an introduction. In: Lerner, Gene (Hg.): *Conversation analysis. Studies from the first generation*. Amsterdam: Benjamins, S. 13-31.
- Kallmeyer, Werner (1996): Einleitung. Was ist Gesprächsrhetorik? In: Kallmeyer (Hg.), S. 7-18.
- Kallmeyer, Werner (Hg.) (1996): *Gesprächsrhetorik. Rhetorische Verfahren im Gesprächsprozess. (= Studien zur deutschen Sprache 4)*. Tübingen: Narr.
- Kallmeyer, Werner/Schmitt, Reinhold (1996): Forcieren oder: Die verschärfte Gangart. Zur Analyse von Kooperationsformen im Gespräch. In: Kallmeyer (Hg.), S. 20-118.
- Kendon, Adam (1990): Some context for context analysis: a view of the origins of structural studies of face-to-face interaction. In: Kendon, Adam (Hg.): *Conducting interaction – Patterns of behavior in focused encounters*. Cambridge: Cambridge University Press, S. 15-49.
- Laurier, Eric (2014a): The graphic transcript: poaching comic book grammar for inscribing the visual, spatial and temporal aspects of action. In: *Geography Compass* 8, 4, S. 235-248.
- Laurier, Eric (2014b): Dissolving the dog: the home made video. In: *Cultural Geographies*, S. 1-12.
- Mondada, Lorenza (2015): The facilitator's task of formulating citizens' proposals in political meetings: Orchestrating multiple embodied orientations to recipients. In: *Gesprächsforschung – Onlinezeitschrift zur verbalen Interaktion* 16, S. 1-62. www.gespraechsforschung-ozs.de/fileadmin/dateien/heft2015/ga-mondada.pdf (Stand: Juni/2016).
- Norris, Sigrid (2004): *Analyzing multimodal interaction*. London: Routledge Falmer.
- Pustički, Nikolina/Schmitt, Reinhold (2015): Zu einer amüsanten Unterbrechung während einer argumentativen Auseinandersetzung. In: Dausendschön-Gay, Ulrich/Gülich, Elisabeth/Krafft, Ulrich (Hg.): *Ko-Konstruktionen in der Interaktion. Die gemeinsame Arbeit an Äußerungen und anderen sozialen Ereignissen*. Bielefeld: transcript, S. 183-208.

- Putzier, Eva-Maria (2011): Das Chemieexperiment: Inszenierung im naturwissenschaftlichen Unterricht. In: Schmitt, Reinhold (Hg.): Unterricht ist Interaktion! Analysen zur De-facto-Didaktik. (= *amades*. Arbeitspapiere und Materialien zur deutschen Sprache 41). Mannheim: Institut für Deutsche Sprache, S. 69-107.
- Sacks, Harvey/Schegloff, Emanuel A./Jefferson, Gail (1974): A simplest systematics for the organization of turn-taking for conversation. In: *Language* 50, S. 696-735.
- Schefflen, Albert E. (1964): The significance of posture in communication systems. In: *Psychiatry* 27, S. 316-313.
- Schefflen, Albert E. (1972): *Body language and social order: Communication as behavioral control*. Englewood Cliffs, NJ: Prentice-Hall.
- Schmitt, Reinhold (2001): Von der Videoaufzeichnung zum Konzept „Interaktives Führungshandeln“: Methodische Probleme einer inhaltlich orientierten Gesprächsanalyse. In: *Gesprächsforschung – Onlinezeitschrift zur verbalen Interaktion* 2, S. 141-192. www.gespraechsforschung-ozs.de/fileadmin/dateien/heft2001/ga-schmitt.pdf (Stand: Juni/2016).
- Schmitt, Reinhold (2004): Die Gesprächspause: „Verbale Auszeiten“ aus multimodaler Perspektive. In: *Deutsche Sprache* 32, 1, S. 56-84.
- Schmitt, Reinhold (2007a): Von der Konversationsanalyse zur Analyse multimodaler Interaktion. In: Kämper, Heidrun/Eichinger, Ludwig M. (Hg.): *Sprach-Perspektiven. Germanistische Linguistik und das Institut für Deutsche Sprache*. (= Studien zur Deutschen Sprache 40). Tübingen: Narr, S. 395-417.
- Schmitt, Reinhold (2007b): Theoretische und methodische Implikationen der Analyse multimodaler Interaktion. In: Holly, Werner/Paul, Ingwer (Hg.): *Medialität und Sprache*. (= Mitteilungen des Deutschen Germanistenverbandes 54, 1). Bielefeld: Aisthesis, S. 26-53.
- Schmitt, Reinhold (2010): Verfahren der Verstehensdokumentation am Filmset: Antizipatorische Initiativen und probeweise Konzeptrealisierung. In: Deppermann, Arnulf/Reitemeier, Ulrich/Schmitt, Reinhold/Spranz-Fogasy, Thomas (Hg.): *Verstehen in professionellen Handlungsfeldern*. (= Studien zur Deutschen Sprache 52). Tübingen: Narr, S. 210-362.
- Schmitt, Reinhold (2012a): Gehen als situierte Praktik: „Gemeinsam gehen“ und „hinter jemandem herlaufen“. In: *Gesprächsforschung Online* 13, S. 1-44. www.gespraechsforschung-ozs.de/fileadmin/dateien/heft2012/ga-schmitt.pdf (Stand: Juni/2016).
- Schmitt, Reinhold (2012b): Räumliche Grundlagen interaktiver Beteiligung. Das Konzept „Interaktionsensemble“. In: Hausendorf, Heiko/Mondada, Lorenza/Schmitt, Reinhold (Hg.): *Raum als interaktive Ressource*. (= Studien zur Deutschen Sprache 56). Tübingen: Narr, S. 37-87.
- Schmitt, Reinhold (2012c): Störung und Reparatur eines religiösen Ritus: Die erloschene Osterkerze. In: Klein, Wolfgang/Habscheid, Stephan (Hg.): *Dinge und Maschinen in der Kommunikation*. (= Sonderheft der Zeitschrift für Literaturwissenschaft und Linguistik 168), S. 62-91.

- Schmitt, Reinhold (2013a): Zum „Be-Greifen“ relevanter Aspekte der räumlichen Umgebung. In: Shamne, Nikolai Leonidovich (Hg.): Raum in der Sprache. Raum der Sprache. Raum der Interaktionen. Festschrift zum 30-jährigen Bestehen des Instituts für Philologie und interkulturelle Kommunikation der Universität Wolgograd. Wolgograd: Verlag der Universität Wolgograd, S. 13-27.
- Schmitt, Reinhold (2013b): Körperlich-räumliche Aspekte der Interaktion. (= Studien zur Deutschen Sprache 68). Tübingen: Narr.
- Schmitt, Reinhold (i.Vorb.): Methodische Verfahren der multimodalen Interaktionsanalyse: „Visuelle Erstanalyse“ und „Frame-Comic“. Mannheim.
- Schmitt, Reinhold/Dausendschön-Gay, Ulrich (2015): Freiraum schaffen im Klassenzimmer: Fallbasierte methodologische Überlegungen zur Raumanalyse. (= Arbeitspapiere des UFSP Sprache und Raum (SpurR) 4). Zürich. www.spur.uzh.ch/research/publications/SpurR_Arbeitspapier_Nr04_150711.pdf (Stand: Juni/2016).
- Schmitt, Reinhold/Deppermann, Arnulf (2007): Monitoring und Koordination als Voraussetzungen der multimodalen Konstitution von Interaktionsräumen. In: Schmitt, Reinhold (Hg.): Koordination. Analysen zur multimodalen Interaktion. (= Studien zur Deutschen Sprache 38). Tübingen: Narr, S. 95-128.
- Schmitt, Reinhold/Knöbl, Ralf (2013): „Recipient design“ aus multimodaler Sicht. In: Deutsche Sprache 41, 2, S. 242-276.
- Schmitt, Reinhold/Knöbl, Ralf (2014): Recipient design: Zur multimodalen Repräsentation des Anderen im eigenen Verhalten. Mannheim: Institut für Deutsche Sprache. <http://ids-pub.bsz-bw.de/frontdoor/index/index/docId/3239> (Stand: Juni/2016).
- Selting, Margret et al. (1998): Gesprächsanalytisches Transkriptionssystem (GAT): In: Linguistische Berichte 173, S. 91-122.
- Selting, Margret et al. (2009): Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). In: Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion 10, S. 353-402. www.gespraechsforschung-ozs.de/fileadmin/dateien/heft2009/px-gat2.pdf (Stand: Juni/2016).
- Stivers, Tanya/Robinson, Jeffrey, D. (2006): A preference for progressivity in interaction. In: Language in Society 35, S. 367-392.
- Stukenbrock, Anja (2009): Herausforderungen der multimodalen Transkription: Methodische und theoretische Überlegungen aus der wissenschaftlichen Praxis. In: Birkner, Karin/Stukenbrock, Anja (Hg.): Die Arbeit mit Transkripten in Fortbildung, Lehre und Forschung. Mannheim: Verlag für Gesprächsforschung, S. 144-169. www.verlag-gespraechsforschung.de/2009/pdf/transkripte.pdf (Stand: Juni/2016).
- vom Lehn, Dirk (2014): Transkription und ethnomethodologische Videoanalyse. In: Moritz, Christine (Hg.): Transkription von Video- und Filmdaten in der qualitativen Sozialforschung. Multidisziplinäre Annäherung an einen komplexen Datentypus. Wiesbaden: Springer VS, S. 391-406.

Gliederung

1.	Vorbemerkung.....	189
2.	Frame-Comic? Eine erste Annäherung	189
3.	Wie analysiert man eine Gesprächspause?.....	192
4.	Visuelle Erstanalyse	196
5.	Zur Begriffsbildung	197
6.	Frame-Comic als Interaktionsdokument	200
7.	Visuelles Segmentieren.....	201
7.1	Interaktionstheoretische Voraussetzungen	202
7.2	Methodisches Vorgehen.....	203
8.	Das Verhältnis von Transkript und Frame-Comic	206
8.1	Vollständigkeit der Abbildung.....	207
8.1.1	Spezifik/Relevanz der Ausdrucksressource	209
8.1.2	Kontinuität und Kontiguität.....	210
8.1.3	Extraktionsweite und Seh-Arten-Potenzial	211
8.1.4	Zunehmende Detaillierung.....	213
8.2	Sequenzialität (der Abbildung)	214
8.3	Repräsentation von Zeitlichkeit	215
9.	Potenzial und Grenzen des Frame-Comics	218
10.	Literatur	220