

Musikvideos systematisch analysieren

Anwendung der Analysesoftware MAXQDA im Rahmen
der Audiovisuellen Grounded-Theory-Methodologie

Analyzing Music Videos systematically. Application of the MAXQDA Analysis Software within the Framework of the Audiovisual Grounded Theory Methodology

Zusammenfassung

Der Beitrag stellt die praktische Übertragung der Grounded-Theory-Methodologie (GTM) auf die Analyse von Musikvideos in Form einer Audiovisuellen Grounded-Theory-Methodologie (AVGTM) dar. Dabei wird nicht nur das Verfahren mit einem Schwerpunkt auf der „Segmentierung“ und „Kodierung“ illustriert, sondern im Detail ausbuchstabiert, wie die Analyseschritte unter Nutzung der Software MAXQDA (2020, Version Plus Portable) umgesetzt wurden. Der Text gibt damit Antworten zu zwei Fragen, die zuletzt stärker vernehmbar waren und prospektiv noch an Relevanz gewinnen werden: (1) Wie lassen sich methodisch kontrollierte Analysen von audiovisuellen Daten durchführen und was ist dabei medienspezifisch und methodisch-methodologisch zu beachten? (2) Wie kann die Software in welchen Phasen die Analyse unterstützen?

Schlagwörter: Grounded-Theory-Methodologie, Audiovisuelle Grounded-Theory-Methodologie, Musikvideos, CAQDAS, MAXQDA

Abstract

This paper presents a practical transfer of the Grounded Theory Methodology (GTM) to the analysis of music videos called „Audio Visual Grounded Theory Methodology“ (AVGTM). We illustrate the analytical procedures in general (with a focus on "segmentation" and "coding"), but also underscore the concrete way of a software based access to the videos (MAXQDA 2020, Version Plus Portable). The text thus provides answers to two questions that have recently become more significant and will prospectively become even more relevant: (1) How can methodologically controlled analyses of audiovisual data be carried out and what has to be considered in terms of media-specific and methodological issues? (2) Which advantages does the software offer in which phases of the analysis?

Keywords: Grounded Theory Methodology, Audio Visual Grounded Theory Methodology, Music Videos, CAQDAS, MAXQDA

1 Einleitung

Seit nun mehr zwei Jahrzehnten werden in den Sozialwissenschaften zunehmend visuelle Daten berücksichtigt und entsprechend wurden und werden Ansätze entwickelt, die methodische Vorschläge für die Auswertung beinhalten. Mittlerweile liegen einige sehr elaborierte Verfahren vor, wie die Dokumentarische Bild- und Videointerpretation (Bohnsack 2009), die Wissenssoziologische Videohermeneutik

(Raab 2008; Raab/Joller/Stanisavljevic 2016) oder die Soziologische Film- und Fernsehanalyse (Peltzer/Keppler 2015). Darüber hinaus zeichnen sich in den letzten Jahren vielfältige Entwicklungen einer qualitativen Videoanalyse ab (Moritz/Corsten 2018). Zu diesen gehören neben der Neukonzeption weiterer Verfahren auch systematische Vorschläge, in denen vorliegende Forschungsansätze mit Blick auf die Auswertung von (audio-)visuellen Daten ausgeweitet werden. Einen dieser Vorschläge haben wir ausgehend vom DFG-geförderten Projekt „Musikvideos, Szenemedien & Social Media – Zur Aushandlung von Rassismus im deutschsprachigen HipHop“¹ für die Grounded-Theory-Methodologie (kurz: GTM) unterbreitet (Dietrich/Mey 2018), indem wir die für Textanalysen tradierten Verfahrenslogiken in Form einer „Audiovisuellen Grounded-Theory-Methodologie“ (kurz: AVGTM; Dietrich/Mey 2019, 2020) spezifiziert haben.

Im Zuge der verschiedenen theoretisch-methodologischen Ausarbeitungen, die sich mit dem Etikett einer „Visuellen Soziologie“ (Raab/Breckner 2016) versehen lassen, sind zunehmend nicht nur sehr konkrete Vorgehensweisen in der Arbeit mit (audio-)visuellen Daten erörtert, sondern auch konstant Fragen nach dem Einsatz von spezieller Software gestellt worden. Mittlerweile wurde seitens der Entwickler*innen auch auf den Trend zu „Visualität“ reagiert und Tools entwickelt, die über die primär für die Textanalyse vorgesehenen Features hinausreichen (Evers et al. 2011).

Vor dem Hintergrund dieser Entwicklungen stellen wir im Folgenden zwei miteinander verknüpfte Aspekte ins Zentrum unserer Darstellung: *Erstens*, die Art und Weise, wie sich mithilfe der AVGTM audiovisuelle Daten *generell* interpretieren lassen – interessant ist dies, weil die GTM zwar ein international genutztes Verfahren ist, aber gemessen an der Relevanz visueller Daten relativ wenig voll ausgearbeitete Analyseverfahren zu diesen Daten vorliegen (initial Konecki 2011, vgl. auch Kautt 2017). Unsere Ausführungen zu den Prozeduren der Analyse im Rahmen der AVGTM (Abschnitt 2 und 3) sind dann v.a. auf die Fragen der Segmentierung und Kodierung konzentriert. Sie bereiten inhaltlich vor, was *zweitens* behandelt werden soll – nämlich wie die AVGTM ganz *konkret* unter Nutzung der Auswertungssoftware MAXQDA 2020 (Version Plus Portable) (Rädiker/Kuckartz 2019) umsetzbar ist (Abschnitt 4). Mit unserer Darstellung wollen wir die Arbeitsschritte einer methodisch kontrollierten AV-Analyse möglichst im Detail vorstellen und im Zuge dessen eine Einschätzung zur punktuellen Anwendung von Software geben – nicht nur aus den erwähnten Gründen, sondern auch, weil audiovisuelle Daten in Zeiten der „Mediatisierung“ (Krotz 2014) und der Ausdifferenzierung immer neuer bildorientierter Apps für Laptops und Smartphones eher noch relevanter werden dürften.

2 Grundschrirte: Segmentier- und Kodierarbeit in der AVGTM

Die Ausgangsüberlegung zur Entwicklung einer Audiovisuellen Grounded-Theory-Methodologie (AVGTM) bestand darin, gegenüber den vorliegenden Ansätzen der interpretativen Sozialforschung im deutschsprachigen Raum, die sich durch eine z.T. sehr feingliedrige und minutiös ausbuchstabierte Analysetechnik – inklusive

entsprechender Partituren und Datenaufbereitungen – auszeichnen, einen Ansatz anzubieten, der eine systematische, aber etwas flexiblere Vorgehensweise eröffnet und dabei größere Datenmengen bearbeitbar macht. Bei der Entwicklung der AVGTM wurde sodann darauf geachtet, die einzelnen Arbeitsschritte in eine ganzheitliche „GTM-Verfahrenslogik“ (Mey/Mruck 2009) – inklusive der Samplebildung und Fallauswahl – zu adaptieren und dabei dem anspruchsvollen Datentypus mit seinem komplexen Zusammenspiel von Text, Ton und Bewegtbild methodisch-methodologisch Rechnung zu tragen. Im Zentrum der AVGTM steht eine Analysearbeit, die die Schritte Fallauswahl, Segmentierung, Kodierung und Kategorienbildung umfasst, wobei nachfolgend vor allem auf die Kodierung abgehoben wird, weil hier die Software am intensivsten genutzt wird. Alle weiteren Schritte werden aber dergestalt mitbehandelt, dass mindestens ein erster Eindruck von dem Verfahren insgesamt möglich sein dürfte (für eine ausführlichere Darstellung siehe Dietrich/Mey 2018, 2019, 2020).

Vor Beginn der konkreten Analysearbeit geht es zunächst um eine *begründete Materialbestimmung*, d.h. die Selektion der Fälle und Dateneinheiten vor dem Hintergrund der Forschungsfrage. Ein „Fall“ ist im Rahmen einer Musikvideostudie (wie unserer) ein Musikvideo und hier gilt es dann zunächst die Voraussetzung für die Interpretation und damit auch die Kodierung zu schaffen. Dies wird möglich, wenn der initiale Fall *segmentiert* wird, d.h. eine *Einteilung in Sinneinheiten* erfolgt.² Die Festlegung der sogenannten Segmente erfolgt allerdings nicht aufgrund zuvor festgelegter Kriterien, sondern *nach Maßgabe der interpretativen Kontrollierbarkeit*. Hierbei ist stets zu erwägen, ob im Falle eines auserkorenen Segments vor dem Hintergrund der Forschungsfrage eine schlüssige Kodierung erwartbar ist und die gewählte Einheit einen beherrschbaren Komplexitätsgrad aufweist. Anschließend können die Segmente im Sinne der GTM-Prozeduren kodiert werden, etwa den Vorschlägen von Strauss und Corbin (1996) folgend offen/axial/selektiv oder sensu Charmaz (2014) initial und fokussiert. Bei der eröffnenden Analysearbeit, in unserem Fall als *offenes Kodieren* realisiert, gilt es, die gebildeten Segmente nacheinander durch generative Fragen (auch „W-Fragen“ genannt) „aufzubrechen“, um zu einer konzeptuellen Erschließung und Sensibilisierung für den untersuchungsrelevanten Gegenstand zu gelangen. Insgesamt verläuft die Analyse von AV-Daten in der GTM-Vorgehensweise zunächst kleinteilig und mikroanalytisch: Die Segmente, die aufgrund der angedeuteten Ermessensspielräume (Stichwort: interpretative Kontrollierbarkeit) durchaus variieren können und damit einzelne Einstellungen, aber auch längere narrative Einheiten umfassen können (zu den Unterschieden zwischen Einstellung, Szene und Sequenz vgl. Peltzer/Keppler 2015, S. 40–41), werden also zunächst im Detail untersucht. Im Vordergrund steht dabei, möglichst offen vorzugehen und theoretische Annahmen, die mit der Forschungsfrage assoziiert sind, reflektiert handzuhaben (Mruck/Mey 2019). Dies meint theoretische Begriffe und Perspektiven als lediglich „sensibilisierende Konzepte“ im Sinne Blumers (1954) mitzuführen, sodass zunächst verschiedene Lesarten und Perspektiven am Material entwickelt werden können, die dann in die Vergabe von Codes zu den Segmenten resultieren. Kodierung im Sinne der GTM zielt immer auf eine interpretatorische Konzeptualisierung der Daten. Entsprechend sollten die Codes nicht deskriptiv angelegt sein, sondern das identifizierte Konzept derart bezeichnen, dass sie als Basis für die späteren Kategorien (die auf Konzepte höherer Ordnung verweisen) funktionieren. Für das Gelingen der Kodierung ist dann eine *enge Verzahnung von Kodierung und Memowriting* vorgesehen. Gerade bei der Analyse von AV-Daten gilt es, die relevanten, die Kodevergabe leitenden, Interpre-

tationsideen in konzeptuellen/theoretischen Memos festzuhalten und forschungsprozessbezogene Überlegungen in Planungs- und Methodenmemos anzulegen. In der weiteren Datenanalyse (sei es fallintern oder bei einem weiteren Fall) kann dann zunehmend großzügiger und ggf. auch schneller vorgegangen werden, insofern durch die zunächst kleinteilige Analyse vorbereitet sinnhafte Zusammenhänge zielgerichteter erkannt und herausgearbeitet werden können. Diese fortschreitende Analysearbeit erlaubt es, zunehmend von den Datendetails zu abstrahieren und eine Verdichtung leisten zu können. In diesem Zusammenhang kommt dem Memowriting wiederum eine herausgehobene Bedeutung zu: In den Memos wird nicht nur expliziert, welche Ideen und Bezugspunkte zur Vergabe welches Codes geführt haben, sondern auch wie und auf welche Weise die Codes, die in eine Liste einzutragen sind, zu einer *Kategorie* verdichtet wurden.

Liegen die Kategorien zu dem Video dann einmal vor, dann bietet das von Strauss und Corbin (1996) vorgeschlagene *Kodierparadigma* einen heuristischen Rahmen für eine systematischere Modellierung ihres Zusammenhangs. Das rekonstruierte Untersuchungsphänomen kann dann hinsichtlich seiner Kontexte/Bedingungen/Strategien/Konsequenzen erfasst werden. Auch hierbei setzt sich das Memowriting als kontinuierliches Fortschreiben der Analyseergebnisse fort: Ein *Auswertungsmemo*, das mit Details aus den konzeptuellen/theoretischen Memos angereichert werden sollte, erlaubt es dann, eine „schlüssige Geschichte“ zu erzählen. Jene schlüssige Geschichte bildet dann die datenfundierte Theorie zur jeweiligen Fragestellung, wobei gerade die Ausformulierung der mit dem Kodierparadigma verbundenen Modellierung im Auswertungsmemo erfahrungsgemäß viel Potenzial zur Konturierung und Präzision der Darstellung bietet.

Bei den hier skizzierten Grundschritten der AVGTM sind allerdings – wie im Folgenden markiert – gerade mit Blick auf die Segmentierung und die Kodevergabe noch einige besondere Aspekte zu unterstreichen, die insbesondere mit der Medialität der AV-Produkte zusammenhängen und für deren Untersuchung relevant sind.

3 Besonderheiten bei der Segmentierung und Kodierung der AV-Daten

Bei AV-Daten – in unserem Forschungskontext sind dies Musikvideos, die wir als Artefakte betrachten, in denen sich relevante Konzepte zur Thematisierung von „Rassismus“ in der HipHop-Szene dokumentieren – gilt es für eine *umfassende Analyse* zu berücksichtigen, dass diese sowohl das *visuelle* Geschehen (inkl. film-sprachlicher Mittel) als auch die *auditive Ebene* – in unserem Falle ist dies der Songtext (Lyrics) sowie die Musik (Ton) – abdeckt. Anzumerken ist hierbei, dass wir mit Blick auf die Kontrollierbarkeit des komplexen Zusammenspiels von Bewegtbild, Text und Ton die Ebenen analytisch trennen, also erst die visuelle Ebene untersuchen und anschließend die text-ton-basierte. Sinnvoll ist dies, weil im Falle der Musikvideos die visuelle Ebene dem Song oft eine gänzlich neue semantische Ebene hinzufügt und sie stark narrativ angelegt ist. Die Separation soll insofern auch ermöglichen, den jeweiligen inhaltlichen Beitrag der Ebenen für die Kon-

struktion des im Zentrum des Forschungsinteresses stehenden Phänomens aufzeigen zu können.

Im Gegensatz zu den meisten Verfahren, die auf Mikroanalysen abheben, liegt bei der Anwendung der AVGTM für die Auswertung der visuellen wie text-ton-basierten Ebene keine a priori Festlegung auf kleine Untersuchungseinheiten vor. Insofern wird anders als etwa bei der Soziologischen Film- und Fernsehanalyse (Peltzer/Keppler 2015) auf eine im Vorfeld vorgenommene Unterteilung in Einstellungen, Szenen (im sogenannten Einstellungsprotokoll, das für analytisch bedeutsame Schlüsselszenen vorgesehen ist) oder Sequenzen (im obligatorischen sogenannten Sequenzprotokoll zur Überblickbildung) verzichtet. Ebenso wenig ist eine technisch a priori festgelegte Einteilung des Produkts in identische Zeiteinheiten (wie etwa im Programm Feldpartitur angelegt) vorgesehen. Vielmehr sind für die Segmentierung – auf visueller wie auf Text-Tone-Ebene – Grundüberlegungen leitend, die der GTM-Verfahrenslogik entsprechen: Dazu gehört die Gegenstandssensitivität: Dies betrifft die Frage, wie der Fall beschaffen ist, indem reflektiert wird, welche formalen Mittel konstitutiv sind und damit zusammenhängend, in welcher Form die Segmentierung des Datenmaterials für die jeweils konkret verfolgte Forschungsfrage umzusetzen ist. Insofern können zur Bestimmung der Segmente sowohl inhaltliche wie formale Elemente herangezogen werden: Ob im Falle der visuellen Analyse die Einteilung des Videos in Sinneinheiten schlussendlich anhand formeller Kriterien (z.B. Einstellungen, Szenen) oder vornehmlich inhaltsbezogene erfolgt (z.B. alle Passagen des Musikvideos, in denen es um körperliche Gewalt geht) hängt letztlich von begründeten Forschungsentscheidungen mit Blick auf Erkenntnisinteressen und Datenbeschaffenheit ab und *nicht* genuin vom Typus des AV-Datums. Um die visuelle Segmentierung der Analyse zugänglich zu machen, wird in jedem Fall ein sogenanntes *Segmentprotokoll* angelegt (das sich vom *Sequenzprotokoll* in der klassischen Filmanalyse dadurch unterscheidet, dass es nicht rein deskriptiven Zwecken verschrieben ist): Das Segmentprotokoll dient der Dokumentation der identifizierten Sinneinheiten, es fungiert aber darüber hinaus als „Ort“, an dem die zentralen Interpretationsschritte in Form von Kodememos hinterlegt sind. Die Kodememos unterstützen die Nachvollziehbarkeit der Interpretation der Segmente, sie werden so angelegt, dass sichergestellt ist, dass die Segmente nach konstanten Beobachtungsschwerpunkten untersucht werden. Derartige Beobachtungsschwerpunkte, die wir im Projektkontext mit dem Begriff „Rubriken“ belegt haben, ergeben sich bei der *Analyse der identifizierten Segmente*. Konstante Rubriken in den Kodememos können z.B. Einstellungsgrößen sein, die Mise en Scène (Bildaufbau) oder auch so etwas wie die optische Stilisierung der Figuren („Kostüm“). Begründet darzulegen ist dabei, welche Elemente die interessierenden Handlungsabläufe auf basale Art mitkonstruieren und welche aus analytisch-programmatischen Gründen eher vernachlässigbar sind. Ein Beispiel: Bei der Komparation der HipHop-Videos wurde recht schnell ersichtlich, dass die spezifischen Montagetechniken für den Rhythmus des Videos (als Teil der Gattung „Musikvideo“) charakteristische Funktionen übernehmen, sie aber keinen zentralen Fokus für die spezifische Analyse der Rassismuskonstruktionen in den Segmenten darstellen. Für die inhaltliche Bearbeitung sind andere Analysefoki bedeutsamer, die sich besser in thematische Rubriken fassen ließen. So zeigte sich in der Analyse einzelner Segmente der ersten von uns bearbeiteten Videos, dass der Modus der Rassismusthematisierung v.a. über die Darstellung von Interaktionen gekennzeichnet ist. Entsprechend legten wir „Interaktion“ als eine zentrale Rubrik fest, um diesen Aspekt bei der Kodevergabe zu jedem visuellen Segment im Blick zu

haben. Neben „Interaktion“ bildeten sich „Geschehen“ und „formale Auffälligkeiten“ als drei Rubriken im Sinne von Ordnungskriterien heraus. *Zusammenfassend* lassen sich also zwei Überlegungen für die Segmentierung und dazugehörige Kodememos festhalten: (1) Die Bestimmung der zu analysierenden Segmente *erfolgt fallabhängig*, entsprechend können inhaltliche und/oder formale Kriterien genutzt werden, (2) die Ergebnisse der *Analyse* sind in einem auf das Segment bezogenen Kodememo (im Segmentprotokoll) zu dokumentieren.

Abbildung 1 verdeutlicht am Beispiel des in MAXQDA angefertigte Memos zum Rap-Video „K4L“ der Künstlerin Ebow den Aufbau und die Funktion des Segmentprotokolls.

The screenshot displays the MAXQDA software interface. On the left, there are two panels: 'Liste der Dokumente' (List of Documents) and 'Liste der Codes' (List of Codes). The 'Liste der Dokumente' panel shows a folder named 'Dokumente' containing a file 'Ebow-K4L_prod_by_walter_p99_arketra'. The 'Liste der Codes' panel shows a list of codes including #T6, #T5, #T4, #T3, #T2, #T1, #V14 - Ende, #V13 - s/w Ebow & Menschen III, #V12 - farbig Fotostüdyo II mit Rap-Gestik, #V11 - farbig s/w abwechselnd: Blaulicht und Wut, #V10 - s/w Vervielfachung Ebows, #V9 - s/w Ebow & unterschiedlichste Menschen, #V8 - s/w Müsli auf Sofa, #V7 - s/w Ebow am Boden/Polizeigewalt, #V6 - farbig "Fotostüdyo Mirza", #V5 - s/w Ebow & Gruppenzugehörigkeiten, #V4 - s/w Zeitungen, Ebow & Miri, #V3 - s/w Ebow, #V2 - s/w Titel und Regisseur, and #V1 - s/w drei "Standbilder".

The right panel shows the 'Segmentprotokoll' (Segment Protocol) for the segment '#V8 - s/w Müsli auf Sofa (24.02.21 13:23, elenamancuso)'. The protocol is structured into three main sections: 'Geschehen' (Event), 'Interaktion' (Interaction), and 'formale Auffälligkeiten' (Formal Features). The 'Geschehen' section describes the scene: 'drei junge Frauen (Ebow in der Mitte) sitzen in Pyjama/Jogginghose auf Sofa, essen Müsli und schauen in die Kamera - es wirkt, als würden sie in einen Fernseher schauen (den man aber als Zuschauer:in nie sieht) (zunächst Halbnähe, dann Halbtotal Einstellung), die Frau links schüttet sich Müsli nach ohne den Blick vom "Fernseher" zu nehmen -> Detailaufnahme Müslischale, dann Sofadecke: Müsli fällt (unbemerkt) auf Sofa'. The 'Interaktion' section states: 'die drei Frauen schauen gelangweilt/unbeeindruckt und dennoch gebannt, fassungslos, fast ohnmächtig in die Kamera ("den Fernseher"); keine Interaktion untereinander'. The 'formale Auffälligkeiten' section notes: 'Kamera nimmt die Perspektive des fiktiven Fernsehschäfers ein, der dem Sofa gegenüberstehen müsste (wird über Tonebene klarer wenn aus Off Interviewstimmen zu hören sind), die drei Frauen wirken mädchenhaft (über Kleidung, Posen) und erinnern an Jugendliche, die nach der Schule Müsli zu Mittag essen, weil niemand zuhause ist, der 'ein ordentliches Mittagessen' zubereitet (zu weit gespannte Interpretation?), Situation wirkt dadurch wie aus dem Alltag'. Below these sections, there is a 'Kode' (Code) section with the text: '#V8: Inszenierung eines konsternierten und gleichzeitig alltäglichen Ausgeliefertseins nicht deutsch gelesener Menschen gegenüber Berichterstattungen aus dem Fernsehen (in weißem Raum)'. At the bottom right, there is a timeline with a vertical axis labeled '1', '2', '3' and a horizontal axis labeled '0'. The timeline shows a sequence of events: 'Sie kennen den Style', 'Kennen die Lines', 'Kennen das Life', 'Kanak for Life', 'Kanak for Life', 'Jedes Klischee bestätigt', and 'In jeder Bildzeitung verewigt, ey'.

Abb.1: Auszug aus dem Segmentprotokoll (Kodememo) zum Musikvideo „K4L“ von Ebow, 2019.

Für eine präzise Kodierung empfiehlt es sich also, die analytisch bedeutsamen Beobachtungsschwerpunkte in Form von Rubriken innerhalb des Kodememos anzulegen. Anzumerken ist dabei, dass die Rubriken teilweise idealtypisch verfasst sind, sie dienen der Strukturierung des analytischen Blicks – v.a., weil eine Analyse, die im strengsten Sinne die Rubriken trennt, nicht möglich ist (so ist kaum ein „Geschehen“ oder eine „Interaktion“ ohne „formale Auffälligkeiten“ beschreibbar, insofern immer wieder auch Details zur Art, wie etwa die Interaktion zu sehen ist, integriert werden müssen, z.B. wenn ein gewaltvoller Übergriff „aus der Perspektive des Opfers“ gezeigt wird). Angezielt ist bei der Analyse grundsätzlich, jedes Segment entlang der Rubriken derart „abzuarbeiten“, dass durch das einheitliche Verfahren kontrolliert erstellte Codes vergeben werden können. Für unverzichtbar bei

diesem analytischen „Abarbeiten“ der Segmente halten auch wir allerdings eine knappe Zusammenfassung dessen, was im gewählten Segment passiert (dies entspricht eben der Rubrik „Geschehen“ und an *dieser* Stelle folgt das *Segment*protokoll dann dem in der Film- und Fernsehanalyse favorisierten deskriptiven Sequenzprotokoll).

Auf die beschriebene Weise wird die visuelle Performance des Videos dann durchanalysiert – schlussendlich liegt ein Segmentprotokoll vor, das alle identifizierten Segmente umfasst und dabei genauso viele Codes aufweist. Letztere sind das Resultat der Beschäftigung mit den identischen Beobachtungsschwerpunkten, den Rubriken in den Kodememos.

Die *Segmentierung und Kodierung des Visuellen* ist aber nur die eine Ebene, die andere Ebene betrifft *Text und Ton* (beim HipHop-Video: Lyrics und Beat). Da Text und Ton gekoppelt untersucht werden sollen (und dies in unserem Fall auch dem „Erleben“ eines Songs am nächsten kommt), gilt es hier zuvorderst zu klären, ob der Text oder der Ton relevantere Einsichten verspricht: Dies ist abhängig von der disziplinären Ausrichtung und der Forschungsfrage. Da unsere Perspektive eine sozialwissenschaftliche und keine musikwissenschaftliche ist und in der GTM sprachlich realisierte Konstruktionen ohnehin den primären Analysegegenstand bilden, werden bei den AV-Produkten zunächst immer die Textsegmente kodiert (die „Lyrics“), was zur Folge hat, dass das zu kodierende musikalische Segment jenes ist, auf das sich der „Sprechakt“ (hier: der Rap-Part) erstreckt. Dies meint konkret, die identifizierten Textsegmente zunächst – je nach herangezogener GTM-Variante offen (Strauss/Corbin 1996) bzw. initial (Charmaz 2014) – zu kodieren. Die Lyrics (wie sie beispielsweise bei Plattformen wie Genius vorliegen) werden also in Sinneinheiten unterteilt, sodass für jedes Segment ein Kode (im Projektkontext „Textkode“ genannt) vergeben werden kann. Erst danach wird die musikalische Stelle kodiert („Musikkode“), auf die sich die Textpassage bezieht. Hierbei geht es dann weniger um die minutiöse „Sezierung“ des Instrumentals als vielmehr um die Interpretation der Stimmung oder des Referenzmodus, den die instrumentalen Elemente erzeugen (was besonders im Falle von HipHop heißen kann, intertextuellen Verweisen nachzugehen, siehe dazu Mikos 2003). Die Kodevergabe für das musikalische Segment sollte dann so erfolgen, dass deutlich wird, welche musikalische Qualität stimmungsbezogen charakteristisch für das Segment ist oder aber aufscheint, welche soziokulturelle Referenz hier (z.B. durch ästhetische Stilmittel wie Sounds aus einer Großstadtkulisse) markiert wird. Der finale Schritt zur Vergabe eines übergreifenden Codes („Gesamtkode“) für das jeweilige Segment ist es dann, Text- und Musikkode miteinander zu konfrontieren, sodass eben ein die Text- und Musikebene abdeckender Kode vergeben werden kann. Ratsam ist es hierbei, den jeweiligen Kode – wo es die Beschaffenheit des Segments zulässt – so zu erstellen, dass er vor dem Hintergrund der Forschungsfrage inhaltliche *und* formale Konstruktionen beinhaltet, sodass idealerweise darin angezeigt wird, wie der (textbasierte) Inhalt sich performativ umsetzt.³ Exemplarisch lässt sich ein Kode anführen, den wir zu einem Text-Ton-Segment in „Nach wie vor“ (2020) von K.I.Z.-Rapper Tarek vergeben haben: Der Kode „Hymnisch-choral umgesetzte Stilisierung von Selbstbehauptung jenseits biografischer Krisen“ markiert den Versuch, den musikalischen Umsetzungsmodus (nämlich hymnenhaft) des Textinhalts (der krisenunabhängigen Selbstbehauptung) herauszuheben und in eine prägnante Formulierung zu bringen.

Schlussendlich liegen sowohl auf der *visuellen Ebene als auch der text-ton-basierten Ebene* Kodelisten vor, auf Basis derer separat durch axiales Kodieren *Kate-*

gorien gebildet werden.⁴ Die Kategorien bilden dann die Fundamente für die Formulierung der GT. Um eine systematische und schlüssige Geschichte erzählen und diese in Form eines Auswertungsmemos entfalten zu können, bietet es sich dann an, noch das sogenannte paradigmatische Modell hinzuzunehmen, um die Relationen der Kategorien und insbesondere die Kernkategorie zu erschließen. Abbildung zwei markiert nochmal die wesentlichen Schritte zu Überblickszwecken:

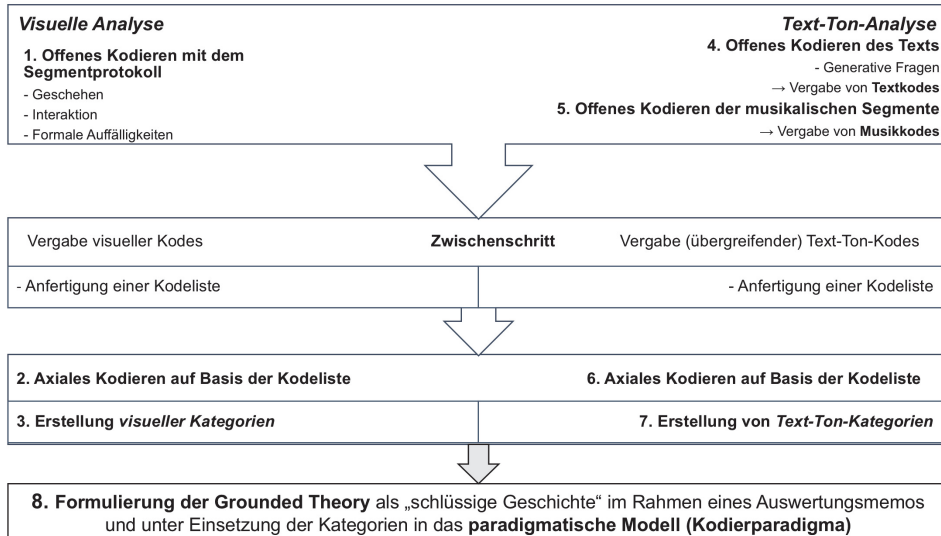


Abb. 2: Zusammenfassung der Analysearbeit mit der AVGTM

4 Computergestützte Auswertung

Nach Skizzierung der grundsätzlichen methodischen Überlegungen mit Blick auf die Spezifika des Audiovisuellen stellen wir im Folgenden nun konkreter dar, wie die AVGTM-Analyse unter Einsatz der MAXQDA-Software umgesetzt wird. Da sie vorwiegend im Zuge des offenen Kodierens für die Generierung der Codes genutzt wurde, legen wir den Fokus der Darstellung auf diesen Bereich, indem wir zunächst die visuelle Analyse konkretisieren und anschließend die Text-Ton-Analyse. Dass wir uns softwarebezogen für MAXQDA entschieden, hängt mit einigen Optionen zusammen, die dieses Tool zwar nicht exklusiv anbietet, aber u.E. gut umsetzt: Dazu zählt die Möglichkeit, visuelle Segmente optisch zu markieren und wiederholt abzuspielen (was im Grunde alle Videoanalysen verlangen) sowie das Angebot, interpretatorische, theoretische oder konzeptuelle Überlegungen in Memos zu dokumentieren, was gerade für GTM-Studien, die durch konstantes „Memoing“ gekennzeichnet sind, bedeutsam ist (Mey/Mruck 2011).

4.1 Visuelle Analyse

Die Entscheidung, die Analyse der visuellen Ebene an den Anfang zu setzen, bietet die Möglichkeit, erste Eindrücke aus den Bildern, explizit *ohne* Text und musikalische Untermalung, wirken zu lassen. Dadurch soll verhindert werden, die Bilder zu sehr „textgefärbt“ zu interpretieren (dies setzt eine „analytische Distanz“ voraus, da das Video bereits mindestens einmal im Zuge der Fallauswahl mit Text und Musik gesichtet wurde). Anders als etwa beim Vorgehen in der Soziologischen Film- und Fernsehanalyse werden die Bilddaten nicht per se durch eine schriftsprachliche Übersetzung (Transkription) bearbeitet, in der alles Visuelle minutiös beschrieben bzw. gar als Protokoll erstellt wird. Da mit der AVGTM eine Konzeptualisierung – und nicht Deskription – geleistet werden soll, sind Codes umfassend zu formulieren und entsprechende Kodennotizen in Memos anzulegen, um die Interpretationsschritte zur Vergabe eines Codes zu dokumentieren.

Bei der praktischen Umsetzung der Segmentierung und Kodierung ist zunächst erforderlich, das Video softwarebezogen zugänglich zu machen: Beim Import der zuvor im mp4-Format heruntergeladenen Videodatei in MAXQDA öffnet sich automatisch der MAXQDA „Media-Browser“. Es bietet sich an, diesen auf komplette Bildschirmgröße zu vergrößern oder auf einem zweiten Bildschirm zu platzieren und direkt mit der visuellen Analyse zu beginnen. Es sollte zunächst darauf geachtet werden, den Ton auszustellen (z.B. über den „Control Panel“, über den auch die Abspielgeschwindigkeit angepasst werden kann). Um sich für ein inhaltliches Segment entscheiden zu können und die *Segmentierung* umzusetzen, kann das Video je nach Bedarf (gesteuert über die F4-Taste) abgespielt oder pausiert und zudem mithilfe der Pfeiltasten im Zehntelsekundenbereich kleinteilig gesichtet werden. Das ausgewählte Segment wird dann durch Gedrückthalten der linken Maustaste über den entsprechenden Abschnitt hinweg markiert. Dieser wird nun automatisch blau hinterlegt und kann über Pfeile an seinen Enden jeweils noch präzisiert und nachjustiert werden. Alternativ kann die zu kodierende Stelle während des Abspielens über die Tasten F7 („Setze Clipanfang“) und F8 („Setze Clipende“) festgelegt werden. Das Kodieren der vorgenommenen Markierung und somit die Umsetzung der eigentlichen *Kodierarbeit* erfolgt, indem durch Drücken der rechten Maustaste im markierten Bereich sich ein Fenster öffnet. Über die oberste Auswahloption „Codieren mit neuem Code“⁵ gelangt man zum Fenster „Neuen Code einfügen“. In diesem Fenster ist dann der Code einzutragen. Da wir von Beginn an konzeptuell kodieren – und damit ganz im Sinne von Konopásek (2008) das in der GTM softwareunterstützte Kodieren nicht technisch, sondern als Interpretieren verstehen – und dabei insbesondere auf die Konstruktion von Rassismus oder Rassismuskritik zielen – sind die von uns generierten Codes komplex(er) formuliert. Sie lauten etwa (wie im Falle des bereits angeführten Tarek-Videos) „Bildsprachlich realisierte subversive Aneignung weiß dominierter Populärkultur zur Konstruktion des schwarzen Helden als gewaltbereitem Gegner weißer Populisten“.⁶ Ein solcher Code ist im beschriebenen Fenster bei MAXQDA in der Zeile „Code“ aufgrund der Beschränkung der Zeichenzahl auf 63 dann aber nicht eintragbar. Wir umgehen dies, indem wir in die „Code“-Zeile statt einen Code lediglich das Kürzel „#V1“ (für „Visuelles Segment 1“, „#V2“ für „Visuelles Segment 2“ usw.) eintragen, den Code selbst führen wir in dem unter der „Code-Zeile“ auftauchenden „Code-Memo“ an, da hier keine Zeichenbegrenzung vorliegt und das Segment in extenso unter Berücksichtigung der als wichtig erachteten Rubriken *Geschehen*, *Interaktion* und *formalen Auf-*

fälligkeiten bearbeitet werden kann. Der im Memo (dargestellt als kleiner gelber Notizzettel rechts in der „Liste der Codes“) vergebene Code wird sodann kopiert und unter „Eigenschaften“ (Rechtsklick auf entsprechenden Code in der Liste) als „Code-Alias“ eingefügt. Dieses Procedere wird für alle Codes wiederholt. MAXQDA bietet zudem die Option, verschiedene Farben für die Segmente zu verwenden, wir bleiben zumeist bei einer Farbe für alle visuellen Codes. Wenn dementsprechend der komplette Zeitstrahl unter dem Video „eingefärbt“ ist, kann der Media-Browser zunächst geschlossen werden, um das von MAXQDA automatisch erstellte und im Feld links unten der Arbeitsoberfläche auftauchende „Codesystem“ zu vervollständigen und spätestens an dieser Stelle den eigentlichen Code aus dem „Code-Memo“ in das „Code-Alias“-Feld einzutragen.⁷ Der „Code-Alias“ bietet bis zu 255 Zeichen und stellte sich für die ausführlich formulierten, interpretatorisch komplexen Codes in unserem Forschungsprojekt als kompatible Alternative zu den in MAXQDA beschränkten Kodennamen heraus. Über den Reiter „Codes“ – „Code-Alias-Tabelle“ können alle oder speziell ausgewählte Codes tabellarisch aufgelistet, in Excel oder Word exportiert und als Grundlage für die weitere Kodierarbeit und Kategorienbildung verwendet werden. Am Ende der visuellen Analyse sollte auch außerhalb von MAXQDA eine *Kodeliste* in einem Textverarbeitungsprogramm vorliegen, mit der weitergearbeitet werden kann (siehe den Folgeabschnitt „Abschließende Datenanalyse“).

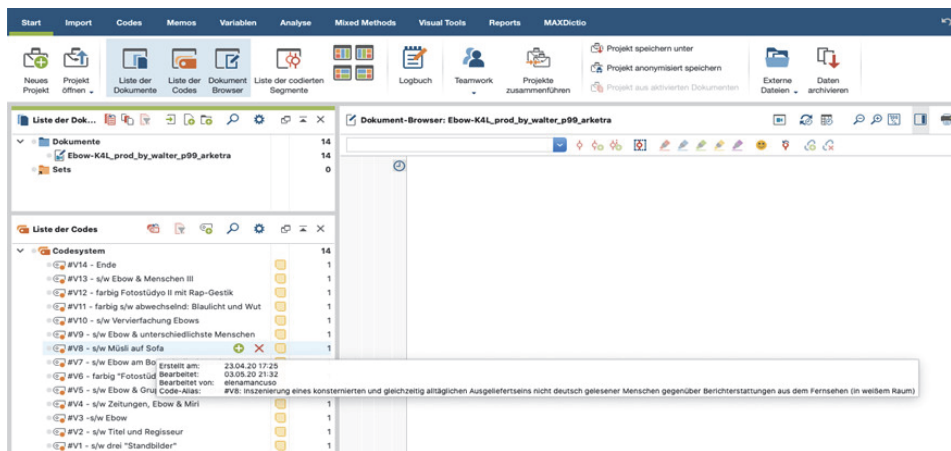


Abb. 3: „Liste der Codes“ und „Code-Alias“ als Grundlage für die zu exportierende Kodeliste

4.2 Text-/Tonanalyse

Nach der visuellen Ebene erfolgt im nächsten Arbeitsschritt die Analyse der Text-/Tonebene (hier: Lyrics und Beat), die zunächst separat für Text und Ton vorgenommen wird. Um diese artifizielle Trennung „zeitnah“ zu kompensieren und das Zusammenspiel der Ebenen zu berücksichtigen, erfolgt die Konfrontation des Textkodes mit den Toncodes, sodass ein übergreifender Code für das jeweilige Segment gebildet wird. Die Analyse zielt also auf einen übergeordneten Text-/Toncode („Gesamtkode“), der gewissermaßen die interakustische Ebene des jeweiligen Text-Ton-

Segments darstellt. Praktisch bedeutet dies in MAXQDA zuerst die Funktion „Videodatei transkribieren“ auszuwählen und die Lyrics in das Fenster „Dokument-Browser“ einzufügen. Da die Songtexte i.d.R. im Netz verfügbar sind, entfallen hier Transkriptionen der gesprochenen/gesungenen Sprache.

Zur *Segmentierung*: Diese erfolgt durch das Markieren eines interpretativ kontrollierbaren Textausschnitts. Für das Segment wird dann über Rechtsklick „Codieren mit neuem Code“ im Bereich „Code-Memo“ ein *Textkode* vergeben. Wie bei der Kodierung der visuellen Ebene tragen wir auch aufgrund der Zeichenlimitation hier lediglich „#T[Zahl]“ in das „Code“-Feld ein und auch der weitere Ablauf – Kodierung und Memowriting simultan im selben Fenster – verläuft technisch gesehen identisch.

Bevor dann die *Tonebene* (Musik) bearbeitet wird, müssen in einem Zwischenschritt die gebildeten Textsegmente mit der Videodatei „verknüpft“ werden. Es muss abgeglichen werden, an welcher Stelle im Video welche Textzeilen auftauchen, da diese bisher quasi separat analysiert wurden und, durch das Einfügen der Lyrics aus dem Internet, nicht im eigentlichen Sinne transkribiert wurde. Um die entsprechenden, bereits kodierten Textstellen im Zeitstrahl des Videos sichtbar zu machen, lässt man das Video abspielen und am Ende der entsprechenden Textpassage im Video stoppen. Parallel dazu muss im Text durch das Drücken der Enter-Taste das Ende eines Textsegments deutlich gemacht werden. MAXQDA fügt dadurch an der jeweiligen Stelle eine Zeitmarke ein zeigt diese durch eine kleine Uhr an der dazugehörigen Passage an. Durch Klick auf diese Uhr im Transkriptionstext kann dann die zugehörige Stelle im Video abgespielt werden – die Kodierung erfolgt jedoch erst im nächsten Schritt. Zunächst wird der abgespielte und auf dem Zeitstrahl des Videos automatisch blau markierte Abschnitt über Rechtsklick „Codieren mit neuem Code“ als übergeordneter Text/Ton-Kode („#TT[Nummer]“) festgelegt. Im „Code-Memo“ können dann bereits die Überschriften „Textkode“, „Tonkode:“ und „Gesamtkode:“ angelegt werden. Hier lassen sich nun die bereits vergebenen Textkodes einfügen.

Das entsprechende Textsegment wird anschließend über Klick auf die Uhr so oft wie nötig angehört und im Memo ein „Tonkode“ vergeben (z.B. „Markierung einer Zäsur durch Beatpause“). Im Idealfall entsteht dann über die Konfrontation der beiden Codes (Text und Ton) im „Code-Memo“ ein aussagekräftiger *Gesamtkode*.⁸ Die letzte Aktion, die auf Ebene der Text/Tonanalyse in MAXQDA realisiert wird, ist es, über Rechtsklick auf den Text/Tonkode (#TT[Zahl] – Einstellungen) in der von MAXQDA angelegten „Liste der Codes“ den Gesamtkode wiederum beim „Code-Alias“ einzutragen. Schlussendlich werden alle Gesamtkodes über „Codes > Code-Alias-Tabelle“ in eine Word-Kodeliste exportiert.

4.3 Abschließende Datenanalyse

Auch wenn die übergeordnete Analyse – in der Logik der GTM sensu Strauss/Corbin also das axiale und selektive Kodieren – und das darin eingebettete Memowriting, also das extensive Überarbeiten aller Memos, zu den sukzessive gebildeten Kategorien und deren Relationierung, generell in MAXQDA erfolgen kann, erfolgten die weiteren Arbeitsschritte unserer Projektarbeit ohne Software-Unterstützung. Denn aufgrund der von uns vergebenen Codes, die die Limitation von 63 Zeichen überschritten, waren die in MAXQDA angebotenen Zusammenführungs- oder

Auflistungsoptionen nicht nutzbar. Entsprechend wurden die beiden Kodelisten – einmal aus der visuellen Analyse und einmal aus der Text-Ton-Analyse – exportiert. Aufgabe der weiteren Auswertungsschritte ist es, die vorliegenden Analyseergebnisse weiter zu bearbeiten und die Codes zu Kategorien zu verdichten. Die bereichsspezifische Separation in visuelle und Text-Ton-Ebene wird in dieser Phase beibehalten, um aus den Codes gehaltvolle (konzeptuelle) Kategorien zu entwickeln.⁹ Eine Zusammenführung der Ebenen (die notwendig ist, insofern es sich bei Musikvideos um eben *audiovisuelle* Produkte handelt) erfolgt in unserem Ablaufschema tatsächlich erst zum Abschluss der Analysearbeit – formal ließe sich sagen, erst im Zuge der *Formulierung der GT*. Dieser finale Arbeitsschritt erfordert es, ein Narrativ zu generieren, das entlang der erstellten Kategorien und ihrer Relationen entfaltet wird. Insofern wird in der Phase der abschließenden Datenanalyse vor allem das konzeptuelle Arbeiten als „Schreibarbeit“ umgesetzt.

Spätestens nach Abschluss der Analyse *aller* berücksichtigten Videos (also im Idealfall, wenn die theoretische Sättigung vorliegt), im Grunde aber auch schon im Verlauf der Auswertungsarbeit, wird – zunächst probeweise, später gezielt – die Fallebene verlassen. Dies meint die erstellten Auswertungsmemos fallübergreifend zusammenzuführen.¹⁰ Im Zuge dieser Verdichtung erfolgt – wie bereits auf der Fallebene – wiederum eine Systematisierung der erstellten Kategorien in Anlehnung an das Kodierparadigma mit seiner Kontext(e)-Bedingung(en)-Strategie(n)-Konsequenz(en)-Matrix (sensu Strauss/Corbin 1996). Dabei werden die visuellen und text-ton-basierten Kategorien modelliert, um zu sondieren, wie genau sich diese zueinander verhalten, um darüber zu a) einer Identifizierung der Kernkategorie („Schlüsselphänomen“) sowie b) den Relationierungen der Kategorien zu gelangen (vgl. dazu eine exemplarische Einzelfallstudie von Dietrich 2020).

5 Fazit

Aufgrund der besonderen Herausforderungen, die GTM für die Analyse von visuellen Daten zu modifizieren, war es in einem ersten Schritt nötig, die einzelnen Verfahrensschritte aufeinander abzustimmen, um dem Datentypus aufgrund seiner Medialität (komplexes Zusammenspiel von Text, Ton, Bewegtbild) für eine kontrollierte und nachvollziehbare Vorgehensweise zu „übersetzen“. Dabei erwies sich die Software (mit Ausnahme der Zeichenbegrenzung von Codenamen) gerade bei der am Anfang durchgeführten „mikroskopischen“ Analyse im Zuge des offenen Kodierens als hilfreich, um das „dichte“ Material zu erschließen. Diese Feststellung macht interessanterweise auf eine Art Verschiebung aufmerksam: Während es bei der GTM-Auswertung von verbalen Daten und daraus hervorgegangenen Transkripten durchaus sinnvoll sein kann, die eröffnende Kodierung nicht unmittelbar in der Auswertungssoftware zu starten, sondern etwas verzögert, wenn erste Grundlinien für das konzeptuelle Arbeiten (vorläufige Kategorien) vorliegen (Mey 2019, S. 70), so erweist es sich bei den komplexen audiovisuellen Daten als äußerst sinnvoll, den für die initiale offene Kodierung notwendigen Segmentierungsprozess direkt softwarebasiert zu bewerkstelligen (s. dazu auch Friese 2019). Ein direkter Start der Analysearbeit mit dem Tool – sowohl für textuelle wie audiovisuelle Daten – setzt voraus, die GTM-Arbeitsschritte des (offenen) Kodierens und Memowri-

tings gemäß einem kontinuierlichen „stop coding – start memoing“ konsequent zu praktizieren. Die Software bietet hierzu eine gute Organisation und Handhabung, insbesondere aufgrund der übersichtlichen Gestaltung der Benutzeroberfläche und der Option, den Mediaplayer flexibel verschieben zu können. Auf diese Weise lässt sich gut zwischen den unterschiedlichen Schritten (wiederholte Ansicht eines Segments, Memoschreiben, Kodevergabe usw.) navigieren und dadurch ein reibungsloses Arbeiten „am Material“ erreichen.

Bei der Datenanalyse hat die Software vor allem auch dazu beigetragen, die zunächst formalen Überlegungen zur Segmentbildung und damit verbundene Explikation der Analyseeinheiten zu konkretisieren. Eine sehr reflektierte Bestimmung der zu kodierenden Segmente steht dabei im Dienste, eine kontrollierte Vorgehensweise der Dateninterpretation zu gewährleisten. Insofern hat uns die Verwendung der Software auch von Beginn an bei der Methodenentwicklung und der Präzision der Arbeitsschritte der AVGTM inspiriert.

Für künftige Studien gilt daher auszuloten – sowohl konkret in jedem Forschungsprojekt, zum einen mit Blick auf die Fragestellung, zum anderen angesichts des einbezogenen Materials –, wann und in welchem Umfang ein Software-basiertes Vorgehen hilfreich ist, d.h. die zu leistenden Arbeitsschritte sinnvoll unterstützt werden können. Dazu bieten die technischen Werkzeuge vielfältige Optionen für das Datenmanagement sowie eine Reihe von Features (Such-Optionen, Farbmarkierungen etc.). Für die Interpretationsarbeit eröffnet sich darüber hinaus die Möglichkeit, dass unterschiedliche, in den Forschungsprozess involvierte Personen gemeinsam software-basiert an demselben Datum arbeiten, Arbeitsschritte der anderen nachvollziehen und sich gegenseitig ergänzen. Ungeachtet dieser Vorteile bleibt künftig genereller zu diskutieren, wie die einzelnen Software-Pakete (neben MAXQDA also etwa ATLAS.ti oder Nvivo, aber auch jene, die spezifischer für Bild-/Videodaten entwickelt wurden, wie ELAN, Feldpartitur oder Videograph) unterschiedlich nutzbar sind und ggf. vor dem Hintergrund spezifischer Projektdesigns weiterentwickelt werden können. Im Zuge dessen wäre eine Diskussion über das Spannungsverhältnis von „Methode mit Technik – Technik mit Methode“ zu intensivieren.¹¹ Dass diese lohnenswert erscheint, ist aus unserer Sicht unstrittig, denn mit den Auswertungstools liegen zahlreiche und hilfreiche Angebote vor, die den Nutzer*innen Spielräume für einen analytischen Zugriff eröffnen, der sich an etwas ausrichtet, was wir für durchaus zentral halten: gegenstandssensitiv das Verfahren anpassen und strukturieren zu können.

Danksagung

Wir bedanken uns für Hinweise bei den beiden anonymen Gutachtenden und für zusätzliche Anmerkungen bei Paul Eisewicht sowie Udo Kuckartz und Stefan Rädiker.

Literatur

- Blumer, H. (1954): What is wrong with social theory? In: *American Sociological Review*, 18. Jg., H. 1, S. 3–10. <https://doi.org/10.2307/2088165>
- Bohnsack, R. (2009): Qualitative Bild- und Videointerpretation. Die dokumentarische Methode. Opladen. https://doi.org/10.1007/978-3-8349-9441-7_20
- Charmaz, K. (2014): *Constructing grounded theory*. 2. Auflage London. <https://doi.org/10.1002/9781405165518.wbeosg070.pub2>
- Dietrich, M. (2020): 'Samy Deluxe' Adriano (2018). Eine Analyse von Rassismus(kritik)konstruktionen aus Perspektive der Grounded-Theory-Methodologie. In: Höllein, D./Woitkowski, F./Lehnert, N. (Hrsg.): *Rap – Text – Analyse*. Deutschsprachiger Rap seit 2000. 20 Einzeltextanalysen. Bielefeld, S. 113–125. <https://doi.org/10.14361/9783839446287-009>
- Dietrich, M./Mey, G. (2018): Grounding visuals. Annotationen zur Analyse audiovisueller Daten mit der Grounded-Theory-Methodologie. In: Moritz, C./Corsten, M. (Hrsg.): *Handbuch Qualitative Videoanalyse*. Wiesbaden, S. 135–152. https://doi.org/10.1007/978-3-658-15894-1_8
- Dietrich, M./Mey, G. (2019): Visuelle Jugendkulturforschung: Trends und Entwicklungen. In: *Diskurs Kindheits- und Jugendforschung*, 14. Jg., H. 3, S. 293–307. <https://doi.org/10.3224/diskurs.v14i3.04>
- Dietrich, M./Mey, G. (2020): Perspektiven einer Audiovisuellen Grounded-Theory-Methodologie und ihr Potenzial im Forschungsfeld „digitale Jugendkulturen“. In: *Psychosozial*, 43. Jg., H. 2, S. 70–83. <https://doi.org/10.30820/0171-3434-2020-2-70>
- Evers, J./Mruck, K./Silver, C./Peeters, B. (Hrsg.) (2011): The KWALON experiment: Discussions on qualitative data analysis software by developers and users. In: *Forum Qualitative Sozialforschung*, 12. Jg., H. 1. <http://www.qualitative-research.net/index.php/fqs/issue/view/36> (25. Februar 2021)
- Friese, S. (2019): *Qualitative Data Analysis with ATLAS.ti*. London.
- Kautt, Y. (2017): Grounded Theory als Methodologie und Methode der Analyse visueller Kommunikation. In: *Forum Qualitative Sozialforschung*, 18. Jg., H. 3, Art. 8. <http://dx.doi.org/10.17169/fqs-18.3.2859>
- Konecki, K. (2011): Visual grounded theory. A methodological outline and examples from empirical work. In: *Revija za Sociologiju*, 41. Jg., H. 2, S. 131–160. <https://doi.org/10.5613/rzs.41.2.1>
- Konopásek, Z. (2008): Making thinking visible with Atlas.ti: Computer assisted qualitative analysis as textual practices. *Forum Qualitative Sozialforschung*, 9. Jg., H. 2, Art. 12. <https://doi.org/10.17169/fqs-9.2.420>
- Krotz F. (2014): Einleitung: Projektübergreifende Konzepte und theoretische Bezüge der Untersuchung mediatisierter Welten. In: Krotz, F./Despotović, C./Kruse, M.M. (Hrsg.): *Die Mediatisierung sozialer Welten*. Medien, Kultur, Kommunikation. Wiesbaden, S. 7–32. https://doi.org/10.1007/978-3-658-04077-2_1
- Mey, G. (2019): „Wir sind keine Kodierautomaten“ – Positionen und Potenziale der Grounded-Theory-Methodologie. Ein Interview. In: *Forum Kritische Psychologie – Neue Folge*, 22. Jg., H. 2, S. 66–80.
- Mey, G./Mruck, K. (2009): Methodologie und Methodik der Grounded Theory. In: Kempf, W./Kiefer, M. (Hrsg.): *Forschungsmethoden der Psychologie*. Zwischen naturwissenschaftlichem Experiment und sozialwissenschaftlicher Hermeneutik. Band 3: Psychologie als Natur- und Kulturwissenschaft. Die soziale Konstruktion der Wirklichkeit. Berlin, S. 100–152.
- Mey, G./Mruck, K. (2011): Grounded-Theory-Methodologie. Entwicklung, Stand, Perspektiven. In: Mey, G./Mruck, K. (Hrsg.): *Grounded Theory Reader*. 2. Auflage Wiesbaden, S. 11–48. <https://doi.org/10.1007/978-3-531-93318-4>

- Mikos, L. (2003): Interpolation und Sampling. Kulturelles Gedächtnis und Intertextualität. In: Androutsopoulos, J. (Hrsg.): Globale Kultur – lokale Praktiken. Bielefeld, S. 64–85. <https://doi.org/10.14361/9783839401149-004>
- Moritz, C./Corsten M. (Hrsg.) (2018): Handbuch Qualitative Videoanalyse. Wiesbaden. <https://doi.org/10.1007/978-3-658-15894-1>
- Mruck, K./Mey, G. (2019): Grounded theory and reflexivity in the process of qualitative research. In: Bryant, A./Charmaz, K. (Hrsg.): The Sage handbook of current developments in grounded theory. London, S. 470–496. <https://doi.org/10.4135/9781526485656.n25>
- Peltzer, A./Keppler, A. (2015): Die soziologische Film- und Fernsehanalyse. Eine Einführung. Oldenbourg. <https://doi.org/10.1515/9783110413779>
- Raab, J. (2008): Visuelle Wissenssoziologie. Theoretische Konzeption und materiale Analysen. Konstanz. <https://doi.org/10.17192/ep2009.4.631>
- Raab, J./Breckner, R. (Hrsg.) (2016): Materiale Visuelle Soziologie. In: Zeitschrift für Qualitative Forschung, 17. Jg., H. 1+2, S. 5–167.
- Raab, J./Joller, S./Stanisavljevic, M. (2016): ‚Stelzereien‘. Eine Fallanalyse zur wissenssoziologischen Videohermeneutik. In: Burzan, N./Hitzler, R./Kirschner, H. (Hrsg.): Materiale Analysen. Methodenfragen in Projekten. Wiesbaden, S. 234–263. https://doi.org/10.1007/978-3-658-12614-8_13
- Rädiker, S./Kuckartz, U. (2019): Analyse qualitativer Daten mit MAXQDA: Text, Audio und Video. Wiesbaden. <https://doi.org/10.1007/978-3-658-22095-2>
- Strauss, A./Corbin, J. (1996): Grounded Theory: Grundlagen Qualitativer Sozialforschung. Weinheim.

Anmerkungen

- 1 <http://gepris.dfg.de/gepris/projekt/403605416>
- 2 Anzumerken ist hierbei, dass sich im Falle eines Musikvideos die Einteilung aufgrund der geringen Spiellänge immer auf das gesamte Video bezieht; anders im Falle von vergleichsweise langen Spielfilmen, bei denen unter Reflexion der Forschungsfrage eine Auswahl analytisch bedeutsamer Segmente erfolgen kann.
- 3 Allgemeine Regeln zur angemessenen Berücksichtigung von „Sound“ (seien es Klänge, Geräusche etc., aber eben auch Musik) zu unterbreiten, ist schwierig. Insofern bildet dies eine sozialwissenschaftliche Herausforderung, die künftig stärker in den unterschiedlichen Methoden (auch im Dialog mit der Musikwissenschaft) bearbeitet werden sollte. Die bisherigen Ausführungen hierzu (z.B. Peltzer/Keppler 2015, S. 182) fallen recht pragmatisch aus.
- 4 Da wir uns projektintern dafür interessieren, was die Bildebene dem text-ton-basierten Song konzeptuell hinzufügt oder wie sie ihn stützt, belassen wir es bei den separaten Kategorien und setzen die Zusammenführung der Ebenen narrativ im Rahmen der Formulierung der GT im Kodememo um. Grundsätzlich denkbar ist es aber auch, die Kategorien nochmals alle Ebenen übergreifend zu elaborieren, sodass Bild-Text-Ton-Kategorien vorliegen. Auch bei dieser Entscheidung ist die Forschungsfrage (sowie die Beschaffenheit und die Komplexitätsgrad der Kategorien) das dominante Kriterium.
- 5 In MAXQDA (und dazugehörigen Referenzen, z.B. Rädiker/Kuckartz 2019) wird die englische Schreibweise „Code“ verwendet, in der überwiegenden deutschsprachigen Literatur zur GTM (z.B. Mey/Mruck 2011) vornehmlich „Kode“ genutzt.
- 6 Aus unserer Perspektive sind Kodes oft präziser in ihrem Hinweis auf Konzepte, wenn sie etwas ausführlicher formuliert sind – tatsächlich müssen häufig sprachliche Wendungen genutzt werden, die im Sinne eines „doing“ den Konstruktions- und/oder Abgren-

zungsprozess einer Artikulation verdeutlichen: Kommen Formulierungen wie „Konstruktion von XY unter Abgrenzung von XZ“ zum Einsatz, dann sind 63 Zeichen recht schnell ausgeschöpft.

- 7 Praktikabler erscheint es, dies gebündelt nach Abschluss einer Kodiersession vorzunehmen, um ein unnötiges „Hin- und Her-Switchen“ zwischen Media-Browser und Arbeitsoberfläche zu vermeiden.
- 8 Um die Leistung der Tonebene (Musik) abzubilden, hat es sich zudem im Projektverlauf etabliert, für die übergreifenden Text-Tonkode Suffixe (wie „Melancholisierung“) und/oder Adjektiv-Kombinationen (wie „düster unterlegte“) zu vergeben.
- 9 Die Zusammenführung der Ebenen kann grundsätzlich und in anderen Forschungszusammenhängen auch schon bei der Kategorienbildung erfolgen. Projektintern interessieren wir uns wie angedeutet für den spezifischen Beitrag der visuellen Ebenen und der Text-Ton-Ebene bei der Konstruktion von rassismusbezogenen Themen. Soll allerdings schon auf Kategorie-Ebene die Trennung aufgehoben werden, dann gilt es die Voraussetzungen dafür zu prüfen. Dies meint zu sondieren, ob Kategorien gebildet werden können, die dann nicht zu überladen sind, d.h. so gut formulierbar sind, dass die auf den verschiedenen medialen Ebenen rekonstruierten Sinnangebote noch adäquat wiedergegeben werden. Besteht der Eindruck, dass die Konstruktionen auf den verschiedenen Ebenen angemessen abgebildet werden können, dann kann eine entsprechende Konfrontation der visuellen Kategorien im Sinne des axialen Kodierens vorgenommen werden.
- 10 Im Grunde genommen sind schon alle Memos, die über den ersten Fall hinausreichen fallübergreifend angelegt, insofern hier das GTM-Prinzip der „constant comparison method“ (Mey/Mruck 2011) beachtet werden sollte. Dies inkludiert nicht nur den Einzelfall als solchen, sondern auch seine Charakteristik im Vergleich mit den anderen untersuchten Fällen herauszuarbeiten und ggf. prospektive Analyseperspektiven zu eröffnen.
- 11 Für eine erste Auseinandersetzung siehe etwa das Symposium „Methode mit Technik – Technik mit Methode“ auf dem 5. Berliner Methodentreffen Qualitative Forschung; Videodokumentation unter: <https://www.youtube.com/watch?v=X2vi2eAZFCY> (26. Mai 2021).