



universität  
wien

# DISSERTATION

Titel der Dissertation

## Die Videointerpretation von Fernsehshows und Musikvideos

Ausgewählte Fallbeispiele zur methodologischen Erweiterung der dokumentarischen Methode

verfasst von

Mag. phil. Mag. rer. soc. oec. Stefan Hampf

angestrebter akademischer Grad

Doktor der Psychologie (Dr. phil.)

Wien, im Mai 2013

Studienkennzahl lt. Studienblatt:	A 092 298
Dissertationsgebiet lt. Studienblatt:	Psychologie
Betreuerin / Betreuer:	a.o. Univ.-Prof. Dr. Thomas Slunecko



# ***DIE VIDEOINTERPRETATION VON FERNSEHSHOWS UND MUSIKVIDEOS***

Ausgewählte Fallbeispiele zur methodologischen  
Erweiterung der dokumentarischen Methode

*Stefan Hampf, Mai 2013  
Dissertation, Universität Wien  
Betreuer: a.o. Univ.-Prof. Dr. Thomas Slunecko*

*Coverfoto: Fotomosaik eines Fotogramms der Fernsehshow „Istanbul Total“ (2004)*

Quelle des Videos: ProSieben, <http://tvtotal.prosieben.de/tvtotal/videos/player/?contentId=14144>

Das Coverfoto zeigt ein Standbild des Moderators Stefan Raab der Fernsehshow „Istanbul Total“ (Fotogramm 79, **Abbildung 6**) in 50%iger Überblendung mit 300 weiteren Standbildern desselben Videos. Die Standbilder wurden jeweils im Sekundentakt aus dem Video entnommen und anschließend mithilfe der Software „Mosaik“ der Firma Meanterm ([www.meanterm.com](http://www.meanterm.com)) verkleinert in 20 Spalten und 15 Zeilen angeordnet. Der Algorithmus der Software sorgt dafür, dass die resultierende Gesamtkomposition der Bilder hinsichtlich Farbe und Form dem Standbild des Moderators möglichst nahe kommt (Softwareeinstellungen: Normal 20x15, Source Opacity 50%, Mix 6/10).

Das Coverfoto wurde durch das Filmplakat der Hollywoodmediensatire „Truman Show“ (1998) inspiriert. Link: <http://www.imdb.com/title/tt0120382/>

*Für Nataliya*



## **Zusammenfassung**

Mit dieser Arbeit wird ein vorrangig methodologisches Erkenntnisinteresse verfolgt: Wie bewährt sich das Verfahren der dokumentarischen Video- und Filminterpretation in der Forschungspraxis und welche Möglichkeiten bestehen zu seiner Erweiterung? Dazu werden exemplarische Videointerpretationen von Fernsehshows und Musikvideos durchgeführt. In der Gegenüberstellung dieser beiden Gattungen von Videomaterial tritt ein deutlicher qualitativer Unterschied zwischen Bildmischung (Fernsehshows) und Montage (Musikvideos) hervor. Das formal bestimmende Gestaltungselement von Fernsehshows ist die Einstellung, was den Personen vor der Kamera noch gewisse Gestaltungsmöglichkeiten gibt. Demgegenüber ist im Falle von Musikvideos durch die Montage die vollständige Kontrolle sämtlicher Leistungen der Personen vor und hinter der Kamera sichergestellt. Wie die Untersuchung zeigt, werden zur effektiven Rekonstruktion der Montage zusätzliche Instrumente benötigt. Ein zentraler Ansatzpunkt dafür sind Videotranskripte. Mit ihrer Hilfe können etwa Montagemuster eines Videos oder Films verdeutlicht und mit anderem Fallmaterial verglichen werden. Montagemuster lassen sich aber auch direkt hinsichtlich der in ihnen aufgeworfenen Farbkontraste interpretieren. Dafür wurde im Rahmen der Arbeit die Farbkontrasttheorie Ittens für die dokumentarische Methode nutzbar gemacht. Wie gezeigt werden kann, stellt die Interpretation von Farbkontrasten ein wichtiges Bindeglied zwischen Einstellungen und Montage dar. Farbkontraste dokumentieren sich sowohl im Fotogramm als auch zwischen Fotogrammen und bieten damit die Möglichkeit zur wechselseitigen Validierung des ikonischen Sinngehalts. Zur Entwicklung weiterer Instrumente in dieser Arbeit zählt zum einen die Übertragung des Imdahlschen Verfahrens der Kompositionsvariation auf die Variation von Einstellungen, Montage und Gesten bzw. Gebärden. Zum anderen wird die bildgebundene komparative Analyse als Interpretationsinstrument vorgestellt. Auf der Grundlage der formalen Komposition eines Bildes lassen sich mittels spezieller ikonischer Suchalgorithmen im Internet ähnliche Bilder finden. Die dabei entstehenden minimalen und maximalen Kontraste eignen sich wiederum als Ansatzpunkte für die weitere Bild- und Videointerpretation.

Schlagwörter: Videointerpretation, dokumentarische Methode, Kulturpsychologie, Medienpsychologie, Montage, Farben

## **Abstract**

The focus of this thesis is primarily methodological and aims to answer the following questions: how successful is the process of documentary video and film interpretation in practical research and which opportunities are available to expand this process? In addition, video interpretations of television programmes and music videos serve as practical examples of the methods both currently used and proposed in this thesis. As a result of the comparison of these two genres of video material, a considerable difference in quality emerges between the techniques of compositing (used in television programmes) and montage (in music videos). The defining design element of television programmes is the camera angle, which provides the people in front of the camera with a certain scope for creativity. In contrast, in the case of music videos the entire scope of efforts and performance by all those involved, both in front of and behind the camera, is controlled as a result of the montage technique. As the analysis shows, additional instruments are required for the effective reconstruction of montage. A central means to approach this analysis is provided by video transcripts. With them, patterns of montage in a video or film can, for example, be ascertained and then compared with other, related material. However, patterns of montage can also be analysed directly in terms of the colour contrasts identified during the interpretation process; Johannes Itten's theory of colour contrast is applied to this end in the course of this thesis in combination with the documentary method. As will be demonstrated, the interpretation of colour contrasts constitutes an important connection between camera angles and montage. Colour contrasts can be documented both in one photogram and between photograms; consequently, they offer the possibility of mutual validation of the iconic meaning. Further methods of film and video interpretation are developed in this thesis: the application of Imdahl's method of composition variation to the variation of camera angles, montage and gestures, as well as the comparative analysis of images, are both proposed as further instruments to be adopted in the interpretation of films and videos. Using the basis of the formal composition of an image, similar depictions can be found on the internet with the help of specialised iconic search algorithms. The resulting minimal and maximal contrasts can, in turn, be applied as further search criteria for the picture and video interpretation.

**Key words:** video interpretation, documentary method, cultural psychology, media psychology, montage, colours

# Inhaltsverzeichnis

<b>Einleitung</b> .....	<b>13</b>
<b>1 Videomaterial und Methodologie der dokumentarischen Methode</b> .....	<b>19</b>
<b>1.1 Fernsehshows und Musikvideos als öffentliche Alltagsdokumente</b> .....	<b>19</b>
<b>1.2 Ikonik: Die Eigensinnigkeit der Bilder</b> .....	<b>20</b>
1.2.1 Wiedererkennendes und sehendes Sehen .....	20
1.2.2 Planimetrische Komposition .....	21
1.2.3 Sehendes Sehen und gegenstandslose Kunst.....	22
<b>2 Montage</b> .....	<b>27</b>
<b>2.1 Begriffsdefinitionen</b> .....	<b>28</b>
<b>2.2 Der methodologische Stellenwert der Montage</b> .....	<b>29</b>
<b>2.3 Zur wechselseitigen Abhängigkeit von Montage und Fotogrammen</b> .....	<b>30</b>
<b>2.4 Zur empirischen Rekonstruktion der Montage</b> .....	<b>31</b>
2.4.1 Montagevariation und Gattungsanalyse.....	31
2.4.2 Farben und Farbkontraste .....	32
<b>3 Transkription</b> .....	<b>37</b>
<b>3.1 Zum Verhältnis von Videotranskripten und Fallmaterial</b> .....	<b>37</b>
<b>3.2 Das Videotranskriptionssystem MoViQ</b> .....	<b>38</b>
3.2.1 Gütekriterien der Videotranskription .....	39
3.2.1.1 Praktikabilität.....	39
3.2.1.2 Ausbaufähigkeit und Flexibilität.....	40
3.2.1.3 Erlernbarkeit.....	40
3.2.1.4 Lesbarkeit .....	41
3.2.1.5 Die eindeutige Zuordenbarkeit von akustischen und visuellen Ereignissen.....	41
3.2.2 Herausforderungen und Potentiale der Videotranskription.....	42
3.2.3 Fallbeispiel .....	43
3.2.4 Aufbau des Videotranskripts.....	45
3.2.4.1 Transkriptkopf.....	46
3.2.4.2 MoViQ-Partitur .....	46
3.2.4.3 Bildgröße.....	48
<b>3.3 Videotranskripte als Grundlage des Samplings</b> .....	<b>48</b>
3.3.1 Sequenzen.....	48
3.3.2 Einstellungen und Fotogramme.....	49
<b>3.4 Videotranskripte als Grundlage der Interpretation</b> .....	<b>50</b>
3.4.1 Fokussierte Gesten bzw. Gebärden .....	50
3.4.2 Videotranskription und Montage.....	52
<b>3.5 Anonymisierung</b> .....	<b>54</b>
<b>3.6 Zusammenfassung</b> .....	<b>56</b>
<b>4 Auswahl des Fallmaterials</b> .....	<b>57</b>
<b>5 „Istanbul Total“: Eingangspassage</b> .....	<b>61</b>
<b>5.1 Auswahl und Interpretation von Videosequenzen</b> .....	<b>62</b>
<b>5.2 Exemplarische Videointerpretationen</b> .....	<b>63</b>
5.2.1 Montagevariation und Gattungsanalyse.....	63
5.2.1.1 Gattungsimmanenter Vergleich.....	72
5.2.1.2 Gattungsübergreifender Vergleich.....	74
5.2.2 Einstellungsvariation.....	77
5.2.2.1 Einstellung und Fotogramm .....	77
5.2.2.2 Exkurs: Text- und bildgebundene Suche im Internet .....	78
5.2.2.3 Fallbeispiel.....	80
5.2.3 Exkurs: Gesten- bzw. Gebärdenvariation.....	86
5.2.4 Reflektierende Gesamtinterpretation und Zusammenfassung.....	91

<b>6</b>	<b>„Istanbul Total“: Die Beyaz-Show .....</b>	<b>97</b>
6.1	Auswahl der Passage .....	97
6.2	Videotranskript.....	98
6.3	<b>Interpretation in der Bilddimension .....</b>	<b>102</b>
6.3.1	Formulierende Interpretation: Sequenzen, Einstellungswechsel und Montage .....	103
6.3.2	Auswahl der Fotogramme .....	106
6.3.3	Formulierende und reflektierende Interpretation der Fotogramme .....	107
6.3.3.1	Fotogramm 1: Stefan Raab auf der Bühne .....	107
6.3.3.2	Fotogramm 2: Beyaz, Assistentin und Stefan Raab.....	113
6.3.3.3	Fotogramm 3: Songcontest Song .....	118
6.3.4	Reflektierende Interpretation von Einstellungswechseln und Montage .....	121
6.3.5	Fokussierungen: Steigerungen, Verdichtungen und Diskontinuitäten.....	122
6.3.5.1	Fokussierungen im Bereich der Performanz der abbildenden Bildproduzent/inn/en.	122
6.3.5.2	Fokussierungen im Bereich der Performanz der abgebildeten Bildproduzent/inn/en	123
6.4	<b>Interpretation in der Dimension von Text und Ton .....</b>	<b>124</b>
6.4.1	Text-Transkript.....	124
6.4.2	Formulierende Interpretation des Textes: Thematische Gliederung.....	125
6.4.3	Reflektierende Interpretation in der Dimension des Textes und Tones.....	125
6.5	<b>Reflektierende Gesamtinterpretation.....</b>	<b>128</b>
6.5.1	Der Gebrauchswert von Faksimiles und Pendants.....	129
6.5.2	Instrumentalisierung des Türkischen zur Selbstinszenierung .....	129
6.5.3	Die Immunisierung des Moderators von „Istanbul Total“ .....	130
6.5.4	Monostrukturierung in „Istanbul Total“ und der Beyaz-Show.....	130
<b>7</b>	<b>Zur Videointerpretation von Musikvideos .....</b>	<b>135</b>
7.1	Zwei Versionen von „Araba“ .....	137
7.2	Die medialen Eigenschaften von Musikvideos .....	139
7.3	<b>Methodologische Herausforderungen an die Video- und Filminterpretation.....</b>	<b>141</b>
7.3.1	Die Rekonstruktion der Bilddimension von Musikvideos .....	142
7.3.2	Die Produktionsverhältnisse von Musikvideos und Fernsehshows .....	143
7.3.3	Farbkontraste, Montagemuster und Räumlichkeit.....	143
7.4	<b>Zur Auswahl der Sequenzen: Repräsentanz und Fokussierung .....</b>	<b>144</b>
<b>8</b>	<b>„Araba“ (1996): Die Version für den türkischen Markt .....</b>	<b>145</b>
8.1	Auswahl der Sequenzen .....	145
8.2	Videotranskription der Eingangspassage nach dem System MoViQ.....	145
8.3	<b>Formulierende Interpretation der Sequenzen.....</b>	<b>147</b>
8.3.1	Vorikonografische Interpretation der Sequenzen bzw. Schnitte .....	148
8.3.2	Ikonografische Interpretation bzw. Auswahl der Sequenzen.....	148
8.3.2.1	Sequenz „Ferrari“ .....	148
8.3.2.2	Sequenz „Ruderer“ .....	148
8.3.2.3	Sequenz „Sänger“ .....	149
8.3.2.4	Sequenz „Fischer“ .....	149
8.3.2.5	Sequenz „Schiffsbug“ .....	149
8.3.2.6	Sequenz „Filmcountdown“ .....	149
8.3.2.7	Sequenz „Hund“ .....	149
8.4	<b>Auswahl der Fotogramme.....</b>	<b>149</b>
8.5	<b>Interpretation der Fotogramme .....</b>	<b>150</b>
8.5.1	HS/ES I: „Ferrari“ .....	150
8.5.1.1	Formulierende Interpretation .....	150
8.5.1.2	Reflektierende Interpretation.....	151
8.5.2	HS/ES II: „Ruderer“ .....	155
8.5.2.1	Formulierende Interpretation .....	155
8.5.2.2	Reflektierende Interpretation.....	156
8.5.3	HS/ES III: „Sänger“.....	159
8.5.3.1	Formulierende Interpretation .....	159
8.5.3.2	Reflektierende Interpretation.....	160

<b>8.6</b>	<b>Zur Relationierung der Fotogramme im Rahmen der Montage .....</b>	<b>161</b>
8.6.1	Reflektierende Interpretation von Einstellungswechsel und Montage.....	161
8.6.2	Reflektierende Interpretation von Farbkontrasten und Montagemustern .....	163
<b>8.7</b>	<b>Ikonologische bzw. ikonische Interpretation .....</b>	<b>169</b>
8.7.1	Exklusivität und Durchschnittlichkeit.....	169
8.7.2	Teilhabe und Distanzierung.....	170
8.7.3	Gesicht und Körperlichkeit .....	171
8.7.4	Akteure und deren Extensionen .....	173
8.7.5	Warten auf die Stellvertreter .....	173
8.7.6	Die Pose des Geschichtenerzählers .....	174
8.7.7	Die Übergegensätzlichkeit von Authentizität und Inszenierung.....	176
<b>9</b>	<b>„Araba“ (2004): Die Version für den internationalen Markt .....</b>	<b>181</b>
<b>9.1</b>	<b>Auswahl des Fallmaterials zur komparativen Analyse .....</b>	<b>181</b>
<b>9.2</b>	<b>Die Rekonstruktion der Montage der Musikvideos „Araba“ 1996 und 2004.....</b>	<b>184</b>
9.2.1	Montage als Ausdruck von Farbkontrasten.....	185
9.2.1.1	Farbe-an-sich-Kontrast.....	186
9.2.1.2	Hell-Dunkel-Kontrast.....	186
9.2.1.3	Kalt-Warm-Kontrast.....	187
9.2.1.4	Komplementär-Kontrast .....	187
9.2.1.5	Simultan-Kontrast.....	188
9.2.1.6	Qualitäts-Kontrast.....	188
9.2.1.7	Quantitäts-Kontrast.....	188
9.2.2	Vorikonografische und ikonografische Interpretation .....	189
9.2.3	Szenische Choreografie und Montage .....	190
9.2.3.1	„Araba“ für den türkischen Markt .....	190
9.2.3.2	„Araba“ für den internationalen Markt .....	191
<b>9.3</b>	<b>Ikonologische bzw. ikonische Interpretation .....</b>	<b>193</b>
9.3.1	Farbkontraste .....	194
9.3.2	Szenische Choreografie im Verhältnis zu Farbkontrasten .....	195
9.3.3	Biografische Brüche von Männern und Frauen .....	195
9.3.4	Reparatur als Naturgesetz.....	196
9.3.5	Reparatur als Angelegenheit individueller Freiheit.....	197
<b>9.4</b>	<b>Zusammenfassung.....</b>	<b>198</b>
<b>10</b>	<b>Abschließende Diskussion.....</b>	<b>205</b>
<b>11</b>	<b>Literaturverzeichnis .....</b>	<b>211</b>
<b>Anhang:</b>	<b>Vorarbeiten zur inhaltlichen Interpretation von „Istanbul Total“ .....</b>	<b>217</b>
	<b>Interpretationsansatz 1: Kommerzieller Verwertungskontext .....</b>	<b>217</b>
	<b>Interpretationsansatz 3: Sexismus und Chauvinismus .....</b>	<b>220</b>
	<b>Interpretationsansatz 4: Islam und Fundamentalismus.....</b>	<b>225</b>
	<b>Interpretationsansatz 5: Gewalt und Gefahr.....</b>	<b>226</b>





**Abbildung 1: Ausschnitt des Komnenus-Mosaiks (1122-1134), Hagia Sophia, Istanbul**

Der Bildausschnitt zeigt die Repräsentation von Irene (Frau des byzantinischen Kaisers Johannes II. Komnenus), die eine versiegelte Schriftrolle in den Händen hält (Qantara, 2013).

Quelle: [http://en.wikipedia.org/wiki/File:Comnenus\\_mosaics\\_Hagia\\_Sophia.jpg](http://en.wikipedia.org/wiki/File:Comnenus_mosaics_Hagia_Sophia.jpg)

## Einleitung

In seinem Buch „Die Kunst der Farbe“ vergleicht der Bauhauskünstler, Kunsttheoretiker und Kunstpädagoge Johannes Itten (2000, 2010) den Anspruch an seine berufliche Tätigkeit mit der eines Wagenbauers<sup>1</sup>. Vermöge seines theoretisch inspirierten Erfahrungswissens, seiner durch Praxis erworbenen Fertigkeiten und anhand von ausgewähltem Fallmaterial möchte er einen möglichst „nützlichen Wagen [...] bauen, der allen eine Hilfe sein soll [...], um rasch und sicher vorwärtszukommen“ (Itten, 2000, S. 11, 2010, S. 7). Zur einführenden Erklärung meiner Dissertation könnte es kein passenderes Gleichnis geben. Auch mir geht es im Wesentlichen darum, dass meine Überlegungen für die Arbeiten anderer brauchbar sind. Im Unterschied zu Itten möchte ich dazu jedoch keinen eigenen Wagen bauen, d.h. keine eigene Methodologie entwickeln. Im Rahmen dieser Doktorarbeit zum Thema Videointerpretation besteht mein Beitrag in der Demonstration von Techniken und Erweiterungsmöglichkeiten für die Methodologie der dokumentarischen Methode. Dabei bewegen mich vor allem Fragen der Leistungsfähigkeit der Video- und Filminterpretation nach dieser Methode.

Anhand ausgewählter Fallbeispiele möchte ich in diesem Zusammenhang zum einen vorhandene methodisch-technische und metatheoretische Grenzen ausloten. Denn nur, wenn man die Grenzen des Verfahrens kennt, lassen sich auch Strategien zu deren Überwindung entwickeln. Zugleich geht es mir zum anderen um die empirische Exploration möglicher Erweiterungen, die zur Erschließung zusätzlicher Erkenntnispotentiale im Rahmen der dokumentarischen Video- und Filminterpretation geeignet sind. Die diesbezüglichen metatheoretischen Überlegungen sind vor allem auf meine Sozialisierung als Kultur- und Medienpsychologie zurückzuführen. Nach der Grundauffassung der Kultur- und Medienpsychologie sind Menschen und Medien nicht als voneinander getrennt anzusehen, sondern bilden füreinander wechselseitige Kontexte, in denen sie sich gemeinsam dynamisch konstituieren (Slunecko, 2008). Die diesbezüglichen praktischen Ansätze stammen aus der unmittelbaren Beschäftigung mit Videomaterial. In der Forschungspraxis wird deutlich, ob sich bestimmte methodologische Konzepte bewähren oder nicht. Dies hat aber nicht nur damit zu tun, dass sich erst in der Anwendung der tatsächliche Erkenntniswert rekonstruierter Ergebnisse beurteilen lässt und etwaige methodische Schwächen deutlich werden. Das direkte Hantieren mit Videomaterial bietet überhaupt erst die Chance, den Rahmen

---

<sup>1</sup> Itten beruft sich dabei auf ein Gleichnis aus der „Veda“, dem ältesten Teil der indischen Literatur (ca. 1500–2000 v. Chr.). „Veda“ bedeutet auf Sanskrit auch „Wissen“ und auf Hindi „Wissenschaft“.

vorgegebener Arbeitsschritte zu verlassen und sich damit zeitweise auch völlig außerhalb wissenschaftlicher Zweckmäßigkeit zu bewegen.

Zusammenfassend sind für mich sowohl das Aufdecken konkreter methodischer Grenzen in der Forschungspraxis der dokumentarischen Video- und Filminterpretation als auch die metatheoretische Reflexion ihrer Forschungslogik in einem interdisziplinären Diskurs als wesentliche Bausteine einer lebendigen Auseinandersetzung mit diesem Verfahren anzusehen. Zugleich möchte ich nicht vergessen darauf hinzuweisen, wie relevant das freie Experimentieren mit Fallmaterial zur Entwicklung kreativer Problemlösungen ist. Damit meine ich nicht nur die zeitweise Einklammerung des Geltungscharakters des Verfahrens, sondern auch den bewussten Rollenwechsel vom/von der forscherschen Betrachter/in zum/zur aktiven Gestalter/in von Videomaterial. Dieser kleine Schritt stellt für viele Psycholog/inn/en und Sozialforscher/innen eine nicht zu unterschätzende Schwelle dar<sup>2</sup>. Aus meiner Sicht bietet jedoch gerade diese Art der De- und Rekontextualisierung wertvolle Impulse zur Entwicklung empirischer Suchstrategien sowie eines tieferen Verständnisses visueller Medialität. Ich habe mich deshalb in dieser Arbeit bemüht entsprechende Fallbeispiele anzuführen, welche die Erkenntnispotentiale dieses Vorgehens verdeutlichen sollen.

Zur Auswahl des Fallmaterials ist in diesem Zusammenhang festzuhalten, dass sie sowohl nach methodologischen als auch inhaltlichen Überlegungen getroffen wurde. Aus methodologischer Sicht wurden Fernsehshows und Musikvideos ausgewählt, weil beides hochgradig erfolgreiche Produkte des professionell-kommerziellen Unterhaltungssektors sind, in denen sich die Leistungen großer und vielschichtiger Produktionskollektive dokumentieren. Zugleich haben populäre Fernseh- bzw. Videoformate ein hohes Zuschauerpotential. Laut Quotenmeter (2004) erreichte die Fernsehshow „Istanbul Total“ (Kap. 5 und 6) am 10. Mai 2004, d.h. bereits am ersten Sendetag, alleine in der Deutschland 2,13 Mio. Zuseher/inne/n, was überdurchschnittlichen 10,8 Prozent Marktanteil entspricht. Der Song des untersuchten Musikvideos „Araba“ (Kap. 7) war 1996 in der Türkei so populär, dass er acht Jahre später als „Neuverfilmung“ (Kap. 9) nochmals auf den internationalen Markt kam. Zugleich lassen sich im Fallmaterial von Fernsehshows und Musikvideos auch unterschiedliche Herstellungsbedingungen

---

<sup>2</sup> Paradoxerweise besteht auf Seiten der Forscher/innen selten echte Scheu vor den Inhalten visueller Medien, selbst wenn diese Inhalte beklemmend oder erschütternd sind. Es kann hingegen durchaus irritieren, Bild- oder Videomaterial nur seiner Form nach mit fremden Bildkontexten in Verbindung zu bringen.

nachweisen, was die Auswertung nach dem Verfahren der dokumentarischen Video- und Filminterpretation vor methodologische Herausforderungen stellt.

Wenngleich der konkrete Inhalt der einzelnen Videos für die Betrachtungsweise dieser Arbeit damit an sich von sekundärer Bedeutung ist, habe ich in meiner Präsentation dennoch auf eine gewisse thematische Einheit des Fallmaterials geachtet. Der Hintergrund dieser Vorgehensweise ist zum einen, den inhaltlichen Fokus soweit konstant zu halten, dass über die Fallbeispiele hinweg ein Vergleich der Interpretationsschritte und Materialarten möglich wird. Zum anderen kann das Fallmaterial aber auch als eine Art Echo des Beginns meiner Dissertation gesehen werden, als mein Erkenntnisinteresse noch stärker von inhaltlichen Überlegungen getrieben war. Die deutsche Fernsehshow „Istanbul Total“ stellte damals mein erstes untersuchtes Videomaterial dar. Ursprünglich forschungsleitend war für mich die Frage nach den kultur- und medienspezifischen Herstellungsbedingungen von Deutschem und Türkischem im Fernsehen<sup>3</sup>. In dieser Doktorarbeit ist die inhaltliche Rekonstruktion des Fallmaterials („Istanbul Total“ sowie andere Videos) jedoch nur so weit von Belang, als dadurch ein Beitrag zum Verständnis der formalen Durchführung der Interpretationsschritte sowie zu den jeweiligen Erweiterungen der Methodologie zu erwarten ist.

Im *ersten Kapitel* dieser Arbeit stelle ich die methodologische Grundposition der Video- und Filminterpretation nach der dokumentarischen Methode sowie zentrale Begriffe vor. Eine detailliertere Begriffsklärung erfolgt sowohl in den theoretischen Folgekapiteln als auch im Zusammenhang mit konkreten Fallbeispielen im empirischen Teil der Arbeit. Im *zweiten Kapitel* widme ich mich in einem theoretischen Zugang der methodologischen Herausforderung der *Montage*. Die Montage wird seitens der Filmwissenschaft als das zentrale Unterscheidungsmerkmal von Videos und Filmen angesehen, hat in der Psychologie und den Sozialwissenschaften jedoch noch nicht diesen Stellenwert erreicht. Im *dritten Kapitel* gehe ich auf das technische Erfordernis sowie die Möglichkeiten der Videotranskription ein. Dabei stelle ich zum einen das Transkriptionssystem MoViQ und zum anderen die verschiedenen Einsatzmöglichkeiten der Videotranskription dar. Diese beinhalten nicht nur neue Möglichkeiten zur Darstellung, sondern auch zur Auswahl und Interpretation von Videomaterial. Im *vierten*

---

<sup>3</sup> Anhand der Eingangspassage von „Istanbul Total“ fertigte ich auch mein erstes Videotranskript (Hampl, 2005, 2006) an. Dieses wird im Kapitel 3 vorgestellt. Im Anhang dieser Arbeit finden sich außerdem noch eine Dokumentation der Vorarbeiten und Interpretationsansätze von damals.

*Kapitel* erläutere und begründe ich die Auswahl des Fallmaterials. Mit dem *fünften Kapitel* beginnt der eigentliche empirische Teil der Arbeit, wobei ich meine methodologischen Überlegungen anhand konkreter Auswertungsschritte von Fallmaterial vorstelle. Im Wesentlichen ist der empirische Teil als mehrstufige komparative Analyse und Methodentriangulation konzipiert. Zu Beginn geht es mir darum, die Leser/innen mit den interpretativen Möglichkeiten des Mediums Video vertraut zu machen. Der Einstieg erfolgt daher zuerst durch die Präsentation einiger konkret am Fallmaterial entwickelter Instrumente. Im Rahmen der Methodentriangulation wird die Interpretation der Eingangspassage der Fernsehshow „Istanbul Total“ (erster Sendetag) anhand einer Montage-, Einstellungs- und Gesten- bzw. Gebärdenvariation durchgeführt. Im Anschluss daran stellt das *sechste Kapitel* sowohl einen minimalen inhaltlichen Kontrast als auch eine methodische Vertiefung dar. Am zweiten Sendetag der Fernsehshow „Istanbul Total“ wird ein Ausschnitt aus der türkischen Beyaz-Show gezeigt. Als „Show in der Show“ wird diese fokussierte Sequenz für die komparative Analyse ausgewählt und nach den von Bohnsack (2009) für die Film- und Videointerpretation beschriebenen Arbeitsschritten sowohl in der Bild- als auch in der Textdimension detailinterpretiert. Im Vergleich zu Fernsehshows wird im *siebten Kapitel* die mediale Gattung Musikvideos vorgestellt. Im *achten Kapitel* wird das Musikvideo „Araba“ interpretiert, das den größten Erfolg in der Karriere des türkischen Sängers Mustafa Sandal darstellt. Am dritten Sendetag der Fernsehshow „Istanbul Total“ hat Mustafa Sandal eine Neuauflage seines Musikvideos für den internationalen Markt vorgestellt. Im *neunten Kapitel* erfolgt daher die komparative Analyse der beiden Musikvideos, von denen das eine für den türkischen und das andere für den internationalen Markt produziert wurde und die sich auf der Bildebene stark unterscheiden. Das *zehnte Kapitel* stellt schließlich die Zusammenfassung der Arbeit dar. Im *elften Kapitel* ist die im Rahmen der Arbeit verwendete Literatur angegeben. Am Ende der Arbeit befinden sich im Anhang noch eine Dokumentation der Vorarbeiten zur Interpretation des Fallmaterials von „Istanbul Total“. Dazu zählen einerseits die Darstellung der Unternehmensverflechtungen hinter der Fernsehshow „Istanbul Total“ sowie weiterführende inhaltliche Interpretationsansätze.



**Abbildung 2: Animal locomotion. Plate 44 (Ausschnitt)**

"BPLDC no.: 08\_11\_000459

Title: Animal locomotion. Plate 44

Volume: Vol. VII. : Males and Females (draped) and Miscellaneous Subjects.

Creator: Muybridge, Eadweard, 1830-1904

Copyright Date: 1887

Extent: 1 photomechanical print : collotype

Genre: Collotypes; Motion study photographs; Book illustrations

Description: Walking, taking off hat

Subject: Men; Walking; Human locomotion

Notes: Plate in: Animal locomotion : an electro-photographic investigation of consecutive phases of animal movements, 1872-1885 / By Eadweard Muybridge. Philadelphia : University of Pennsylvania, 1887, v. 7; Plate descriptions from: Animal locomotion : an electro-photographic investigation of consecutive phases of animal movements : prospectus and catalogue of plates / by Eadweard Muybridge. Philadelphia : Printed by J.B. Lippincott Company, 1887."

Quelle: Eadweard Muybridge (2009)

## **1 Videomaterial und Methodologie der dokumentarischen Methode**

Die Grundidee, sich im Rahmen dieser Arbeit an die Auseinandersetzung mit methodologischen Fragen sowie an mögliche Erweiterungen heranzuwagen, beruht auf der Entstehungsgeschichte der dokumentarischen Video- und Filminterpretation selbst. Das Verfahren existiert seit nunmehr zirka fünf Jahren (Bohnsack, 2009; Erstausgabe: 2008) und ist ursprünglich als Erweiterung der Methode der Bildinterpretation entstanden. Die auf Bohnsack zurückgehende Bildinterpretation nach der dokumentarischen Methode (Bohnsack, 2009) ist derzeit, neben der visuellen Wissenssoziologie (Breckner, 2010; Raab, 2008), einer der vielversprechendsten Interpretationsansätze, der für die Psychologie nutzbar gemacht werden kann. Sie zeichnet sich durch einen hohen Grad an theoretischer Durchdringung des Gegenstands Bild, durch Klarheit und Systematik der Arbeitsschritte sowie durch empirische Prüfinstrumente zur methodischen Kontrolle des Common Sense der Forscher/innen aus<sup>4</sup>. Im Zusammenhang mit dem Verfahren der Video- und Filminterpretation nach der dokumentarischen Methode möchte ich in der Folge auf einige zentrale Begriffe eingehen, die für das Grundverständnis der Methodologie erforderlich sind. Weiterführende Begrifflichkeiten und Konzepte werden im weiteren Verlauf der Arbeit direkt anhand von Fallmaterial erläutert.

### **1.1 Fernsehshows und Musikvideos als öffentliche Alltagsdokumente**

Wie der Titel dieser Arbeit erkennen lässt, beschränkt sich meine Auseinandersetzung mit dem Verfahren der Video- und Filminterpretation auf Kulturdokumente der sogenannten Massenkommunikation: Fernsehshows und Musikvideos. Diese Medien sind insbesondere aufgrund der dahinterliegenden, weitgehend anonymen Produzent/inn/en/kollektive von Interesse. Dabei folge ich der grundsätzlichen Einteilung Bohnsacks (2009, S. 117), der Filme bzw. Videos auf einer ersten Ebene in Erhebungsinstrumente und Alltagsdokumente einteilt. Auf der zweiten Ebene dann lassen sich Alltagsdokumente nochmals in Produkte des privaten und in solche des öffentlichen Bereichs differenzieren. Fernsehshows und Musikvideos sind demnach als Alltagsdokumente des öffentlichen Bereichs anzusehen und können auf empirischem Wege prinzipiell sowohl mittels Produkt- als auch mittels Rezeptionsanalysen

---

<sup>4</sup> Letzteres kann als methodologisches Manko der nomothetischen Psychologie (Valsiner, Molenaar, Lyra, & Nandita, 2009) angesehen werden. Ähnlich wie bislang den Cultural Studies (Bohnsack, 2009, S. 125) gelingt auch ihr keine systematische Unterscheidung von Common Sense und wissenschaftlicher Interpretation.

untersucht werden. Während Rezeptionsanalysen in der Psychologie eine lange Tradition haben, sind Produktanalysen eher selten. Dieses Ungleichgewicht hat unter anderem mit dem bisherigen Fehlen geeigneter Instrumente zur direkten systematischen Rekonstruktion von Medien- und Kulturprodukten zu tun. Die Video- und Filminterpretation nach der dokumentarischen Methode wurde dazu entwickelt, diese Lücke zu schließen.

## **1.2 Ikonik: Die Eigensinnigkeit der Bilder**

Laut Bohnsack (2009, S. 51) liegt die größte Besonderheit bildlicher Medien darin, komplexe Sinngehalte im Modus der Gleichzeitigkeit hervorzubringen. Dabei sind Bilder und Videos in der Lage, auf simultane Weise Widersprüchlichkeiten zu fassen, ohne die resultierenden Spannungen zwischen Differenz und Synthese in eine bestimmte Richtung hin auflösen. Diese spezielle Medieneigenschaft von Bildern hat Imdahl (1996a) als „Übergegensätzlichkeit“ bezeichnet. Just aus dem Zwischenzustand von Entweder *und* Oder generieren Bilder ihre ikonische Energie und Wirkung. Nach Bohnsack (2009) ist die Vermittlung dieses ikonischen Sinns im Medium der Sprache oder des Textes nur annäherungsweise und mit großem Aufwand realisierbar<sup>5</sup>. Bohnsacks programmatische Forderung ist es deshalb, sich im Bereich der Methodenentwicklung der „Eigensinnigkeit von Bildern“ (Bohnsack, 2009, S. 32) anzunehmen. Für die Rekonstruktion der spezifischen Simultanstruktur von Bildern schlägt Bohnsack (2009) nach dem Verfahren der dokumentarischen Methode ein Prozedere vor, das Panofskys (2006) und Imdahls (1996a) kunstwissenschaftliche Bildtheorien auf der Basis ihrer wissenssoziologischen Wurzeln für den Anwendungsbereich der empirischen Sozialforschung nutzbar macht. Aus methodologischer Sicht stützt sich das Verfahren der dokumentarischen Methode, in vergleichbarer Weise wie die objektive Hermeneutik, auf die „Leitdifferenz“ (Bohnsack, 2009, S. 19) von immanentem Sinn und Dokumentsinn.

### **1.2.1 Wiedererkennendes und sehendes Sehen**

Nach Bohnsack (2009, S. 56) ist der erste Arbeitsschritt bei der Interpretation von (Stand-)Bildern im Sinne Panofskys (2006) die vorikonografische bzw. ikonografische

---

<sup>5</sup> Die Struktur von Text und Sprache ist im Wesentlichen eine sequenzielle, wobei Slunecko (2008) mit McLuhan (1992) darauf hinweist, dass der Text gegenüber der oralen Sprache noch ein zusätzliches Maß an Distanzierung mit sich bringt. Bei Slunecko (2008) finden sich jedoch auch Hinweise, wie die Entwicklung der Schriftlichkeit die Bildwahrnehmung kulturhistorisch im Sinne eines stärker „wiedererkennenden Sehens“ vorbereitet.

Interpretation. Dabei sind die wiedererkennbaren Bildbestandteile des Vorder-, Mittel- und Hintergrunds in einfachen Worten zu beschreiben und kulturspezifische Wissensbestände zu explizieren. Zusammenfassend bezeichnet Bohnsack (2009) diesen Interpretationsabschnitt als *formulierende Interpretation*. Das erklärte Ziel dieses Vorgehens ist es, den Inhalt des Bildes, das *Was*, möglichst konkret und intersubjektiv überprüfbar in ein textliches Abbild zu überführen.

Im Zuge der anschließenden *reflektierenden Interpretation* erfordert der folgende Arbeitsschritt die Suspendierung bzw. „Einklammerung“ (Bohnsack, 2007, S. 163; Mannheim, 2003, S. 88) des gerade wiedererkannten Bildinhalts. Im Sinne des „sehenden Sehens“ (Imdahl, 1996a) soll nun der Blick auf die planimetrische Komposition des Bildes gerichtet werden, welche laut Bohnsack (2009, S. 40) die eigentlich bedeutsame Ebene zum Verständnis der Ikonik von Bildern sei.

### **1.2.2 Planimetrische Komposition**

Unter der planimetrischen Komposition eines Bildes wird nach Imdahl (1996a) seine Flächenstruktur verstanden. Bei der Rekonstruktion ist das Bild als Ensemble verschiedener Formen, Farben, „Gewichte“, „Spannungen“ zu verstehen, welche die Blickrichtung leiten und unterschwellig die Aufmerksamkeit der Bildbetrachter/innen auf sich ziehen. Ziel der Rekonstruktion dieser Ebene ist die Entwicklung eines Zugangs zur formalen Grundstruktur des Bildes, welche sich stets in Form einer „Übergegensätzlichkeit“ (Imdahl, 1996a) im Bild dokumentiert.

Die Rekonstruktion der planimetrischen Komposition ist für Ungeübte insofern voraussetzungsreich, als sie, wie bereits zuvor erwähnt, auf dem sehenden Sehen aufbaut. Entsprechend seiner metatheoretischen Fundierung ist das sehende Sehen zwar für den Alltag in hohem Maße handlungsleitend, jedoch nicht ohne weiteres bewusst verfügbar. In der Regel wird im Alltag das wiedererkennende Sehen als die *einzige* Form des Sehens verstanden. Das sehende Sehen kann jedoch durch das Betrachten gegenstandsloser bzw. abstrakter Kunst geübt werden.

### 1.2.3 Sehendes Sehen und gegenstandslose Kunst



Abbildung 3: „Schwarzes Quadrat auf weißem Grund“, Kasimir Malewitsch, 1915

Besonders ergiebig zur Einübung des sehenden Sehens erscheinen die bildnerischen Gestaltungstheorien des *staatlichen Bauhauses*, einer Kunstschule, die 1919 von Walter Gropius gegründet wurde und bis 1933 bestand. Die Künstler/innen des Bauhauses (etwa: Malewitsch, Kandinsky, Itten etc.), ihres Zeichens Vertreter/innen der Avantgarde bzw. klassischen Moderne, waren Zeitgenoss/inn/en von Karl Mannheim und Erwin Panofsky und bestechen sowohl aufgrund ihrer eigenen Kunstproduktion als auch wegen des hohen Grades an theoretischer Durchdringung ihrer künstlerischen Überlegungen. Da die Gestaltungsprinzipien der Bauhauskünstler/innen weitgehend in schriftlich publizierter Form sowie als eigene Bilder vorliegen, lassen sie sich für die rekonstruktive qualitative Bildforschung nutzbar machen. Insbesondere Ittens (2000, 2010) „Kunst der Farbe“ sowie Kandinskys „Punkt und Linie zu Fläche“ (1986)<sup>6</sup> bieten überaus konkrete Anhaltspunkte dafür, wie sich die planimetrische Komposition im Sinne Imdahls rekonstruieren lässt, und erscheinen damit in hohem Maße als Arbeitsschritte der dokumentarischen Bildinterpretation geeignet <sup>7</sup>. Aus metatheoretischer Sicht hat Malewitsch (1994)<sup>8</sup> die Auseinandersetzung mit dem Bild maßgeblich beeinflusst, indem er es von seiner Abbildungsfunktion befreit hat.

---

<sup>6</sup> Erstveröffentlichung: 1926

<sup>7</sup> Im Sinne einer empirisch-theoretischen Durchdringung des Planimetriekonzepts erscheinen auch die Texte der Bauhauskünstler Klee (1990) und Malewitsch (1994) sowie des japanischen Designtheoretikers Kenya (2009) für die rekonstruktive Bildforschung von zentraler Bedeutung.

<sup>8</sup> Erstveröffentlichung: 1927

Eine Entsprechung der Imdahlschen Konzeption der planimetrischen Grundstruktur des Bildes, die durch das „sehende Sehen“ bestimmt sei, findet sich in Kasimir Malewitschs suprematistischem Manifest über „die gegenstandslose Welt“ (Malewitsch, 1994) aus dem Jahre 1922. Für Imdahl (1994, S. 316) erweist sich erst am Bildmaterial der „gegenstandslosen Kunst“ die wahre Überlegenheit seines Konzepts der Ikonik gegenüber Panofskys Ikonografie und Ikonologie. In Panofskys Ansatz lassen sich gegenstandslose Bilder (z. B. Abbildung 3) zwar beschreiben und in eine Stilgeschichte einordnen, jedoch bleibt auf dieser Ebene deren eigentlicher Bildsinn, die Ikonik, welche auf der Planimetrie beruht, verschlossen. Für das wiedererkennende Sehen ist ein schwarzes Quadrat nichts weiter als ein schwarzes Quadrat. Erst die empirische Berücksichtigung des sehenden Sehens ermöglicht es, die formalen Strukturierungsangebote des Bildes genauer zu erfassen und auf diesem Wege die damit verbundenen kulturgeschichtlichen Grundbedingungen des Wiedererkennens zu rekonstruieren. Wissenschaftlich gesehen müssen wir uns letztlich fragen, was dieses Quadrat überhaupt so selbstverständlich hervortreten lässt. Schon eine erste planimetrische Betrachtung zeigt, dass wir dafür die kognitive Leistung erbringen müssen, den weißen Rand auszublenden, der das Quadrat umgibt. Rein planimetrisch betrachtet ist die Leinwand des Bildes in exakt gleichem Ausmaß schwarz, wie sie weiß ist<sup>9</sup>. Durch die Rekonstruktion der Planimetrie erhalten wir folglich Aufschluss über uns selbstverständlich gewordene Sehgewohnheiten – gewissermaßen die Standards der visuellen bzw. „ikonischen Kommunikation“<sup>10</sup> (Przyborski & Slunecko, 2011, 2013). Im Sinne „methodisch kontrollierten Fremdverstehens“ (Schütze, Meinefeld & Weymann, 1973) bietet uns erst dieser Schritt die Möglichkeit, die Alltagswahrnehmung empirisch sauber vom Fallmaterial zu unterscheiden.

Die abstrakte Kunst verweist in fokussierter und direkter Weise auf die spezifische Eigensinnigkeit des Bildes, die durch kein anderes Medium ersetzt werden kann. Im Vergleich zum Kunsthistoriker Imdahl, der seine Bildinterpretationen auf dem

---

<sup>9</sup> Der weiße Bilderahmen ist dabei gar nicht mit eingerechnet.

<sup>10</sup> Wie sich mit Kenya (2009), Krämer (1998), Slunecko (2008) sowie Slunecko und Przyborski (2009) zeigen lässt, kann im vorliegenden Bild (Abbildung 1) nur dann etwas wahrgenommen werden, wenn es gleichzeitig gelingt, das Trägermedium mit zu erfassen und dadurch „unsichtbar zu machen“ (Krämer, 1998). Kenya (2009) erklärt, wie sich die ursprüngliche Trägereigenschaft des Papiers, als Hintergrund darauf gemalter oder geschriebener Information, in der asiatischen und westlichen Welt sukzessive auf die Farbe Weiß übertragen und damit kultur- und mediengeschichtlich zu deren Unsichtbarwerden beigetragen hat. Slunecko und Przyborski (2009) erläutern anhand des Beispiels eines marokkanischen Berberteppichs, in den westliches Millimeterpapier eingewoben ist, die kulturellen Missverständnisse, die entstehen, wenn das konventionalisierte Unsichtbarmachen des Trägermaterials misslingt.

ontologischen Verhältnis von „sehendem Sehen“ und „wiedererkennendem Sehen“ aufbaut, erörtert der Kunstschaffende Malewitsch die Modi der Bildauffassung noch grundsätzlicher auf der Ebene von „Gegenständlichkeit“ und „Gegenstandslosigkeit“. In seinem Sinne bleibt der Zugang zur Ikonik des Bildes jedoch nicht nur dem wiedererkennenden Sehen verwehrt, sondern entzieht sich prinzipiell jedem gegenständlichen Bildverständnis, das rein verstandesmäßig am „Praktischen“, „Wirtschaftlichen“ und „Nützlichen“ orientiert ist<sup>11</sup>.

In Hinblick auf die Entwicklung empirischer Instrumente zur Rekonstruktion des „sehenden Sehens“ (Imdahl, 1996a) bzw. der „Gegenstandslosigkeit“ (Malewitsch, 1994) sind nicht nur Malewitschs Schriften, sondern insbesondere seine Bilder von hoher Relevanz. Laut Imdahl „steckt die Leistung eines jeden Bildes von Rang in der Erschaffung einer solchen stellvertretenden Repräsentationsform, die in ihrer Augenscheinlichkeit nur als das System eines Bildes bestehen und somit selbst durch nichts anderes vertreten werden kann“ (1994, S. 319). Malewitschs wohl bedeutendstes Werk, „Schwarzes Quadrat auf weißem Grund“ (Abbildung 3), ist in diesem Sinne zu Recht zu einer Ikone der „Gegenstandslosigkeit“ geworden. Bildimmanent fordert es das „wiedererkennende Sehen“ sowie den Common Sense heraus und öffnet sich zugleich für das „sehendes, Sehen“, welches das Bild in seiner Bildlichkeit anerkennt<sup>12</sup>.

Zusammenfassend lässt sich von einem zweifachen Nutzen der gegenstandslosen Kunst für die empirische Forschung sprechen: Zum einen schult sie den forschenden Blick auf das Bild, zum anderen bietet sie taugliche Mittel zu dessen Interpretation an. Indem sie sich dem Common Sense entzieht, erlaubt sie das Einüben einer auf die spezifische Medialität des Bildlichen eingestellten Untersuchungshaltung. Indem sie ihre eigene Gestaltungspraxis reflektiert und anleitet, bietet sie wertvolle Ansatzpunkte und Instrumente, die für die empirische Forschung brauchbar sind. In weiterer Folge möchte ich im Hinblick auf die Interpretation der planimetrischen Struktur von Bildern insbesondere auf die Arbeiten von Itten (2000, 2010) eingehen, dessen Farbkontrasttheorie ich als unmittelbar geeignet für die Bild- und Videointerpretation nach der dokumentarischen Methode ansehe.

---

<sup>11</sup> In diesem Sinne können Malewitschs Schriften auch als Kritik am wissenschaftlich-psychologischen Mainstream verstanden werden. Die Forderung nach „Gegenstandslosigkeit“ ist letztlich auch ein Aufruf gegen die „Vergegenständlichung“ bzw. „Reifizierung“ (Przyborski & Slunecko, 2009).

<sup>12</sup> Malewitschs Titel „Schwarzes Quadrat auf weißem Grund“ macht diesen Gedanken sogar explizit zum Thema. Das schwarze Quadrat wird vergegenständlicht. Man hat den Titel des Bildes jedoch erst verstanden, wenn man das Bild „sehend“ erfasst und seine Übergegensätzlichkeit verstanden hat.



**Abbildung 4: One Week Skin (ESPN-Hs), Jason Salavon (2012)**

Quelle: <http://salavon.com/work/one-week-skin/image/405/>

"Archival Inkjet Print  
70" x 53". Ed. 7 + 2 APs.

A suite of large print-works sourcing and abstracting exactly one week of specific pay-TV networks at 15 frames per second. These millions of ordered, discrete images are reformatted using various graphic, sequential, and data-visual approaches.

The three different reformulations (or "skins") of network streams each lay equal claim to accurately representing ("visualizing") the captured TV matter. Perhaps this emphasizes a contradictory sense of both truth and arbitrariness in the formal manifestations." (Salavon, 2012)

## 2 Montage

Neben der Entwicklung verschiedener methodologischer Ansätze zur Bildinterpretation (etwa: Bohnsack, 2009; Breckner, 2010; Przyborski & Slunecko, 2011, 2013) sind in den Sozialwissenschaften während der letzten Jahre auch einige Konzepte zur Interpretation von Filmen und Videos entstanden (etwa: Bohnsack, 2009; Dinkelaker & Herrle, 2009; Raab, 2008). Eines der metatheoretisch und empirisch elaboriertesten Verfahren stellt die Film- und Videointerpretation nach der dokumentarischen Methode dar (Bohnsack, 2009), die sich forschungspraktisch bis heute im Zusammenhang mit zahlreichen Fragestellungen und Fallmaterialien bewähren konnte (Baltruschat, Bohnsack & Pfaff, 2010; etwa: Baltruschat, 2010; Bohnsack, 2009; Hampl, 2010; Sobotka, 2009).

Mit der zunehmenden Etablierung der Methode der Film- und Videointerpretation in den Sozialwissenschaften ist jedoch nicht nur ein Mehr an empirischen Fallstudien zu erwarten, sondern auch eine steigende methodologische Auseinandersetzung, zu der diese Arbeit einladen möchte. Bei näherer Betrachtung der bisherigen Forschungsarbeiten, die mithilfe der Film- und Videointerpretation nach der dokumentarischen Methode entstanden sind, fällt eine besondere Schwerpunktsetzung auf die Interpretation von *Fotogrammen*, d. h. von Standbildern, auf. Die Interpretation von Fotogrammen ist im Rahmen der dokumentarischen Methode metatheoretisch gut begründet und stützt sich sowohl auf die *wechselseitige Validierung* (Bohnsack, 2009, S. 155) mit anderen ausgewählten Standbildern eines Films bzw. Videos als auch auf die Interpretation des Sprechtextes. Somit lassen sich durch die Relationierung der Fotogramme in der dokumentarischen Methode die Ergebnisse auf ihre Gültigkeit hin überprüfen. „Homologien“ oder „Strukturidentitäten“ bestehen laut Bohnsack (2009, S. 165) jedoch auch in den „Relationen der Einstellungen *zueinander*“ – in der Montage. Erst durch den systematischen Vergleich der Interpretationen auf der Ebene der Fotogramme *und* der Montage werden daher die Potentiale der dokumentarischen Methode – im Sinne einer „Relationierung der beiden Relationen“ (Bohnsack, 2009, S. 165) – wirklich ausgeschöpft. Dieser Aspekt wurde in der Forschungspraxis und bei der Methodologieentwicklung bisher eher vernachlässigt.

Indem ich mich in dieser Arbeit den Phänomenen in der Bilddimension von Videos und Filmen genauer zuwende, komme auch ich nicht umhin, gewisse Punkte im Rahmen meiner metatheoretischen Betrachtung einzuklammern. Dabei handelt es sich um die Tondimension von Videos und Filmen sowie die wechselseitige Referenz von Ton- und

Bilddimension. Die diesbezüglich laut Bohnsack (2009) erforderlichen Arbeitsschritte zur Interpretation des Ton- bzw. Textraums werden in Kapitel 6 zwar anhand des empirischen Fallmaterials einer Fernsehshow genauer vorgestellt. Eine systematische Reflexion der Bedeutung der Tondimension für die Montage kann jedoch im Rahmen dieser Arbeit nicht geleistet werden. Mein vorrangiges Erkenntnisinteresse ist die Klärung der Verhältnisse zwischen Bildern und bewegten Bildern in der Bilddimension. Erst auf Basis dieser Erkenntnisse kann meiner Ansicht nach die methodologische Miteinbeziehung der Tondimension in differenzierter Weise erfolgen.

## 2.1 Begriffsdefinitionen

Der Fokus dieses Kapitels liegt auf der Rekonstruktion der *Montage* bzw. den damit zusammenhängenden Leistungen der Bildproduzent/inn/en. Der Begriff der Montage lässt sich von seiner sprachlichen Entstehungs- und filmwissenschaftlichen Verwendungsgeschichte her erläutern. Ursprünglich stammt das Wort Montage aus dem Französischen und bezeichnet prinzipiell jegliche Art des mechanischen *Zusammenbauens*. Die technische Aufladung des Begriffs wurde mit dem Aufkommen der Filmtechnologie auf das *Zusammenkleben* einzelner Filmstreifen übertragen. Im deutschen Sprachraum wird synonym zur Montage das Wort *Schnitt* verwendet. Auch dieser Begriff betont die technische Seite der Filmherstellung, wobei damit genau genommen das Gegenteil von „Montage“, nämlich das *Auseinanderschneiden* von Filmstreifen, bezeichnet wird.

In dieser Arbeit findet eine weiter gefasste Definition von Montage Verwendung, welche die rein mechanische Betrachtung überschreitet. Diese Definition stammt aus der Film- und Fernsehwissenschaft (etwa: Mikos, 2003) und begegnet uns auch bei der Übersetzung von „Montage“ ins Englische: Der Begriff *editing*<sup>13</sup> leitet sich aus der *Bearbeitung* und *Herausgabe* von Texten ab und schließt neben der mechanischen auch die künstlerische *Gestaltung* von Filmmaterial mit ein. Im Hinblick auf die empirische Untersuchung von Filmen und Videos entspricht diese Auffassung den Überlegungen von Bohnsack (2009, S. 159) und Mikos (2003), welche die Montage als Vorgang der mechanischen *und* ästhetischen Gesamtkomposition ansehen. Aus methodologischer Sicht wird somit unter Montage nicht nur das technische Zusammenfügen einzelner Einstellungen zu Sequenzen verstanden, sondern auch die effektive „Herstellung

---

<sup>13</sup> Analog zum Deutschen ist im Englischen neben „editing“ auch das Wort „cutting“ gebräuchlich.

narrativer und ästhetischer Strukturen durch diesen technischen Vorgang“ (Mikos, 2003, S. 207).

Als *Einstellung* bezeichnet Bohnsack (2009, S. 159) kontinuierlich aufgenommene Filmstücke, welche ihrerseits aus zeitlich geordneten Einzelbildfolgen bestehen. In Abgrenzung zur Montage ist laut Bohnsack (2009, S. 159) eine Einstellung „so lange gegeben, wie kein Schnitt erfolgt“. Während sich die Einstellung im Rahmen der Standbildfotografie rein durch die Wahl des *Kamerastandpunkts* sowie des Bildausschnitts bzw. der *Einstellungsgröße* definiert, kommen bei Film und Video die *Kamerabewegungen* hinzu (Bohnsack, 2009, S. 160). Dazu zählen sowohl Kamerafahrten und -schwenks als auch Vergrößerungen bzw. Verkleinerungen des Bildausschnitts.

Die Begriffe *Szene* und *Sequenz* sind in ihrer filmwissenschaftlichen Verwendung nicht einheitlich voneinander abgegrenzt. Laut Bohnsack (2009, S. 160) handelt es sich in beiden Fällen um filmische Einheiten, die aus der Verbindung „einer oder mehrerer Einstellungen“ entstehen. Als Unterscheidungsmerkmal kann nach Monaco (2009) die Länge dieser Einheiten dienen: „Sequenzen sind sicherlich länger als Szenen [...]“ (Monaco, 2009, S. 171). Bohnsack (2009, S. 161) schlägt für die sozialwissenschaftliche Analyse jedoch eine inhaltliche Differenzierung vor, die sich aus der Differenzierung abbildender und abgebildeter Bildproduzent/inn/en ergibt. Bei genauerer Betrachtung orientiere sich der Begriff der Szene eher an den Personen oder Objekten *vor* der Kamera, also den Leistungen der abgebildeten Bildproduzent/inn/en, während sich Einstellung und Montage auf die Leistungen der abbildenden Bildproduzent/inn/en, also der Personen *hinter* der Kamera, beziehen. Damit ist im Sinne Bohnsacks (2009) der Begriff der Sequenz dafür geeignet, die erwähnten filmischen Einheiten in sich aufzunehmen: „Beides zusammen – also der Zusammenhang von *Szene* einerseits und *Einstellung* und *Montage* andererseits – konstituieren eine *Sequenz*“ (Bohnsack, 2009, S. 161).

## **2.2 Der methodologische Stellenwert der Montage**

Für das Verfahren der Film- und Videointerpretation nach der dokumentarischen Methode ist die Rekonstruktion der Montage nach Bohnsack (2009, S. 158) insbesondere aus zwei Gründen von Interesse: Zum einen stellt die Montage, vom Standpunkt der zeitgenössischen (etwa: Mikos, 2003; Wiedemann, 2005) sowie zum Teil auch der klassischen Filmwissenschaft (etwa: Pudowkin, 1999) aus gesehen, das

wesentlichste Alleinstellungsmerkmal des Mediums Film bzw. Video dar<sup>14</sup>. Laut Aussage des Filmregisseurs Stanley Kubrick ist „[d]ie Montage [...] der einzige Vorgang, bei dem der Film keine Anleihen bei anderen Künsten“ macht (Kubrick, zitiert bei Schumm, 2004, S. 167). Zum anderen dokumentiert sich, wie bereits erwähnt, in der Montage die exklusiven Leistungen der abbildenden Bildproduzent/inn/en, was wesentliche Konsequenzen für die Interpretation mit sich bringt: Während bei der Untersuchung von Fotogrammen stets ebenso die Leistungen der abgebildeten Bildproduzent/inn/en, d. h. der Personen *vor* der Kamera, berücksichtigt werden müssen, stellt die Montage „eindeutig das Produkt der abbildenden Bildproduzent/inn/en dar“ und eröffnet „uns den Zugang zu deren *modus operandi*“ (Bohnsack, 2009, S. 162).

In diesem Zusammenhang kann aus der Perspektive der Forschungspraxis argumentiert werden, dass der Stellenwert der Rekonstruktion der Montage mit der Art des verwendeten Fallmaterials zusammenhängt. Für die Interpretation von Filmen und Videos, die von den Forscher/inne/n selbst zu Studienzwecken erstellt wurden und bei denen somit die Leistungen der abgebildeten Bildproduzent/inn/en im Vordergrund stehen, spielt die Rekonstruktion der Montage in der Regel eine sekundäre Rolle<sup>15</sup>. Für die Interpretation von solchen Filmen oder Videos, die bereits als Alltagsdokumente vorliegen – insbesondere, wenn es sich dabei um kommerzielle Produkte der kollektiven Handlungspraxis arbeitsteiliger Spezialistenteams handelt – kommt der Montage hingegen eine Schlüsselfunktion zu.

### **2.3 Zur wechselseitigen Abhängigkeit von Montage und Fotogrammen**

Der vorige Abschnitt hat deutlich gemacht, welche prinzipiellen erkenntnistheoretischen Nachteile sich aus der sozialwissenschaftlichen Vernachlässigung der Montage ergeben. Konkrete Probleme erwachsen daraus auch für die Forschungspraxis – etwa im Zusammenhang mit Film- oder Videomaterial, das durch aufwändige Schnittfolgen und viele Einstellungen bestimmt ist. Oft lässt sich in solchen Fällen schwer entscheiden, aus welchen Sequenzen und Einstellungen sich das Video zusammensetzt und welche Fotogramme als für die Zusammensetzung repräsentativ ausgewählt werden sollen. Auch wenn die Auswahl der Fotogramme gelingt, ist der Stellenwert der

---

<sup>14</sup> Laut Flusser (1997, S. 91) gibt es einen wesentlichen Unterschied zwischen Fotografen und Kameramännern/-frauen: „[D]er Fotograf und sein Apparat erzeugen die Fotografie, während der Kameramann mit seiner Kamera nur das Rohmaterial für die spätere Erzeugung des Films liefert“. „[D]er eigentliche Erzeuger des Films“ ist laut Flusser „der Mann mit der Schere und dem Klebstoff“ (S. 92).

<sup>15</sup> Dies schließt nicht aus, dass auch in diesem Fall die Rekonstruktion der Montage das Erkenntnisinteresse von Forschungsprojekten darstellen kann.

Interpretationsergebnisse insofern eingeschränkt, als diese nicht im Zusammenhang der Gesamtkomposition von Filmen oder Videos mit der Montage relationiert werden können.

Bohnsack (2009) beschreibt ebenso am Beispiel von „Istanbul Total“, wie sich anhand des Einstellungswechsels und der Montage die hierarchische Ordnung des Videomaterials in Form von Haupt-, Unter- und eingelagerten Sequenzen rekonstruieren lässt. Nach den Prinzipien der „Repräsentanz“ und „Fokussierung“ (Bohnsack, 2009, S.201) lassen sich aus den identifizierten Teilsequenzen Fotogramme für die Interpretation auswählen. Wie bereits erwähnt, stellt die Auswahl für die dokumentarische Methode an sich kein elementares Problem dar. Nach metatheoretischer Grundauffassung muss sich der Dokumentsinn – wenn auch in unterschiedlich starker Fokussierung – prinzipiell in allen Aspekten des Fallmaterials widerspiegeln. Der Auswahlprozess von Fotogrammen auf dem Wege der Ableitung aus Sequenzen kann meiner Erfahrung nach jedoch bei komplexerem Videomaterial massiv erschwert oder sogar verunmöglicht werden. So zeichnen sich etwa Musikvideos oder Werbefilme durch eine so hohe Anzahl an Schnitten und Einstellungen aus, dass keine hierarchische Festlegung von Teilsequenzen, über welche die Fotogramme ausgewählt werden könnten, gelingt. In diesem Fall wäre es vorteilhaft, einen direkten empirischen Zugang zu den Montagemustern entwickeln zu können, aus dem sich dann wiederum – im Sinne der Validierung einer rekonstruierten Fokussierungsmetapher – Suchstrategien für die Fotogramme generieren ließen. Wie zahlreiche Arbeiten zeigen (etwa: Baltruschat, 2010; Bohnsack, 2009; Sobotka, 2009), ist im Sinne einer wechselseitigen Validierung auch der Vergleich von Fotogrammen mit der Ton- und Sprachebene von Filmen bzw. Videos möglich. Aus den hier dargelegten Gründen halte ich es jedoch für die dokumentarische Methode für unverzichtbar, die Methodenentwicklung zur empirischen Rekonstruktion der Montage gezielt voranzutreiben. Dazu möchte ich in der Folge einige Vorschläge machen.

## **2.4 Zur empirischen Rekonstruktion der Montage**

### **2.4.1 Montagevariation und Gattungsanalyse**

Zur Entwicklung eines tieferen Verständnisses der Gesamtkomposition von Bildern schlägt Imdahl (1994) deren experimentelle Veränderung durch die Forscher/innen vor. Im Rahmen der Film- und Videointerpretation lässt sich das von Imdahl (1994) entwickelte und von Bohnsack (2001a, S. 77) als „Kompositionsvariation“ bezeichnete

Verfahren auch für die Interpretation von *Fotogrammen*, d. h. von Standbildern, anwenden. Die Übertragung des Konzepts auf die Interpretation der Montage ist anspruchsvoller, da sich Einstellungslängen, -wechsel und -reihenfolgen von Seiten der Forscher/innen nicht so ohne weiteres modifizieren lassen<sup>16</sup>. Es ist jedoch möglich, das Verfahren der Kompositionsvariation als Sonderform des theoretischen Samplings zur komparativen Analyse aufzufassen und so für die Interpretation der Montage nutzbar zu machen. Unter Kompositionsvariation der Montage bzw. *Montagevariation* (Kap. 5.2.1) kann zusammenfassend der Vergleich von Montagemustern verstanden werden, die sich in Videotranskripten dokumentieren. Die aus der Montagevariation resultierenden Möglichkeiten einer gattungsimmanenten und gattungsübergreifenden Interpretation der Montage stelle ich im Rahmen dieser Arbeit vor. Auf dem Wege der Montagevariation lassen sich im Sinne von Keppler (2006) sowohl Film- und Videomaterial zur komparativen Analyse auswählen als auch die Gattungszugehörigkeit eines bestimmten Fallmaterials empirisch bestimmen. Abschließend lässt sich aus der Rekonstruktion von Einstellungswechseln und Farbkontrasten ein ausgesprochen direkter Zugang zur empirischen Interpretation der Montage entwickeln.

#### **2.4.2 Farben und Farbkontraste**

Farben sind insofern für die Bild- und Videointerpretation von zentraler Bedeutung, als sie den Gegenstand, d. h. die (bewegten) Bilder, überhaupt erst in Erscheinung treten lassen<sup>17</sup>. Im Zusammenhang mit der Film- und Videointerpretation eröffnet diese globale Eigenschaft von Farben auch einen empirischen Zugang zum formalen Verständnis von Fotogrammen und Montage. Mit der Rekonstruktion von Farben ist die sozialwissenschaftliche Forschung aber auch vor ein empirisches Problem gestellt. Entgegen der Alltagsauffassung sind Farben – übrigens analog zur Kontextabhängigkeit von Texten – nämlich in bedeutsamer Weise von ihrer Umgebung abhängig und damit nicht ohne weiteres individuell bestimmbar. Johannes Itten (2000, 2010) ist dieser Problematik bei der Gestaltung von Bildern begegnet. In Abkehr von einem isolierten Farbverständnis entwickelte er eine Farbtheorie, die explizit die Verhältnisse von Farben betont. Wie die empirischen Fallbeispiele dieser Arbeit zeigen (Kap. 7 und 9), lässt sich die Gestaltungstheorie der „sieben Farbkontraste“ (2000, 2010) als

---

<sup>16</sup> Zur Verschiebung von Personen in Filmen bzw. Videos oder zur Modifikation der Montage sind fortgeschrittene Videoschnittkenntnisse erforderlich.

<sup>17</sup> Ittens (2000, 2010) Definition von Farben schließt auch die „Nichtfarben“ Schwarz und Weiß ein. Diese kleine aber wesentliche Präzisierung ist für die Untersuchung von Schwarz-Weiß-Filmen von Bedeutung.

Rekonstruktionstheorie im Rahmen der dokumentarischen Methode nutzen. Für die empirische Forschung ist von Vorteil, dass Itten (2000, 2010) sich – im Gegensatz zu Imdahls (2003) kunsttheoretischen Reflexionen – sehr pragmatisch mit dem Verhältnis zwischen physikalisch-chemischer *Farbwirklichkeit* und psychophysischer *Farbwirkung* befasst hat. Dabei sind mit Ittens Theorie der Farbkontraste nicht nur zahlreiche analytische und kunsthistorische Fallbeispiele, sondern auch anschauliche Gestaltungsprinzipien verbunden, die für die sozialwissenschaftliche Forschungspraxis brauchbar sind.

Zur Rekonstruktion der Montage ist es möglich, bei den farblichen Relationen von Fotogrammen und Einstellungen anzusetzen, die sich in Videotranskripten als spezifische Montagemuster dokumentieren: Farbwechsel verstärken Schnitte und betonen den Einstellungswechsel, indem sie den visuellen Kontrast zwischen den verbundenen Einstellungen erhöhen<sup>18</sup>. Werden Farben über Einstellungen hinweg beim Schnitt beibehalten, tritt dieser formal im Video bzw. Film stärker in den Hintergrund, wodurch eine stärkere Verbindung zwischen den Einstellungen entsteht (vgl. Kap. 3.4.2). Ein geringerer Unterschied zwischen den Fotogrammen dokumentiert sich auch im Videotranskript. Dieses wirkt dadurch weniger mosaikartig und insgesamt homogener. In weiterer Folge spielen Farbkontraste im Zusammenhang mit der Montage auch für die Herstellung der *Diegese* (Souriau, 1997), d.h. des „Filmraums“, eine große Rolle. In professionellen Film- und Videoproduktionen werden durch die Verbindung farblich homologer Einstellungen Räume<sup>19</sup> und Atmosphären geschaffen bzw. durch die Verbindung farblich heterogener Einstellungen Ortswechsel (Kap. 8.6.1) angezeigt. Aus der empirischen Rekonstruktion der Verhältnisse von Einstellungs- und Farbwechseln sind im Rahmen der Montage somit auch Zugänge zur raumzeitlichen Ordnung von Filmen und Videos zu erwarten.

Auch die Bildinterpretation bzw. die Interpretation von Fotogrammen kann von einer Berücksichtigung der Farbenlehre Ittens (2000, 2010) profitieren. Da Farben und Farbkontraste darüber entscheiden, ob Linien und Formen im Bild in Erscheinung treten, ermöglicht die Interpretation der Farben und Farbkontraste auch die empirische Verständigung über die Konstitution von Linien und Formen im Bild. Das wesentlichste Argument für die Anwendung des empirischen Instrumentariums der Farbkontraste auf

---

<sup>18</sup> Diese Eigenschaft der Farben beim Schnitt wird auch von Computerprogrammen zur automatischen Schnitterkennung genutzt.

<sup>19</sup> Nicht umsonst gibt es im Deutschen den Begriff „Farbraum“.

die Bild- und Videointerpretation ist schließlich die daraus entstehende Möglichkeit zur bereits mehrfach angesprochenen wechselseitigen Validierung von Montage und Fotogrammen. In diesem Sinne hat dieses Kapitel hoffentlich deutlich gemacht, welche Optionen die Montage für die sozialwissenschaftliche Film- und Videointerpretation bereit hält.

Information card with text, including the title "The Leonard Cohen Library" and other details.

The Leonard Cohen Library



**Abbildung 5: Cinema Redux At Moma's Action! Design Over Time, 20.03.10**

Museum of Modern Art, New York City

Quelle: <http://brendandawes.com/blog/cinema-redux-at-momas-action-design-over-time>

“CINEMA REDUX™

Creating a visual fingerprint of an entire movie

Created in 2004 and acquired for the MoMA permanent collection in 2008, Cinema Redux creates a single visual distillation of an entire movie; each row represents one minute of film time, comprised of 60 frames, each taken at one second intervals. The result is a unique fingerprint of an entire movie, born from taking many moments spread across time and bringing all of them together in one single moment to create something new.

Film students, academics and obsessives with time on their hands may use Dawes's grids to postulate new theories about the language of film.

— John Walters, The Guardian

Cinema Redux has been featured in two MoMA exhibitions in New York: Design and the Elastic Mind in 2008 and Action! Design Over Time in 2010.

Cinema Redux is a registered trademark of Brendan Dawes“ (Dawes, 2004)

### 3 Transkription

In vielen Methodenbüchern wird der *Transkription* von Fallmaterial meist wenig oder gar kein Platz eingeräumt. Wenn das Thema Transkription in Publikationen Erwähnung findet, dann meist nur im Zusammenhang mit *Datensicherung* (etwa: Przyborski & Wohlrab-Sahr, 2009, S.160) bzw. *Dokumentation* erhobener Daten (etwa: Bortz & Döring, 2002, S.311). Dabei kommt der Transkription gerade in der Forschungspraxis qualitativer Verfahren eine zentrale Bedeutung zu. Effektiv beansprucht die Erstellung von Transkripten im Rahmen qualitativer Forschungsprojekte einen großen Teil der Arbeitszeit<sup>20</sup>. Im Rahmen von Forschungswerkstätten sind Transkripte von Fallmaterial in der Regel die vorrangige Arbeits- und Kommunikationsgrundlage. Diese im Umgang mit Texttranskripten selbstverständliche Zugangsweise erscheint im Zusammenhang mit Videotranskripten nach wie vor erklärungsbedürftig. Dieses Kapitel geht deshalb den grundsätzlichen Fragen nach, was Videotranskripte zeigen und wie sie sich für die Abbildung, das Sampling und die Interpretation von Fallmaterial einsetzen lassen.

#### 3.1 Zum Verhältnis von Videotranskripten und Fallmaterial

Dass Videotranskription ein geeignetes Mittel darstellen kann, um Video- bzw. Filmmaterial in Publikationen abzubilden, ist im Rahmen der Video- und Filminterpretation nach der dokumentarischen Methode bereits mehrfach demonstriert und betont worden (etwa: Baltruschat, 2010; Bohnsack, 2009; Hampl, 2010; Przyborski & Wohlrab-Sahr, 2009; Sobotka, 2009). Damit Videotranskripte aber auch *anstelle* von Videos bzw. Filmen als Sampling- und Interpretationsinstrumente zum Einsatz kommen können, bedarf es zuvor der methodologischen Bestimmung der gegenseitigen Referenz von Video bzw. Film und Transkript. Nach Bohnsacks Auffassung (2009, S.171), mit Bezug auf Keppler (2006, S.107), bietet die Erstellung von Transkripten an sich bereits „einen Teil der Analyse“, jedoch sind dieser Auffassung nach Videotranskripte noch nicht als Akte der Interpretation, sondern lediglich der medienimmanenten Übertragung anzusehen. Von Übertragung spricht Bohnsack, da durch das Transkript noch keine „begrifflich-theoretische *Explikation*“ (Mannheim, 2003, S.272) der Themen (im Sinne einer formulierenden Interpretation) sowie der

---

<sup>20</sup> Abhängig von Ausgangsmedium (z. B.: Audio- oder Videoaufzeichnung) und Detaillierungsgrad (z. B.: bloßer Sprechtext oder Kennzeichnung von Tonhöhen) kann die Transkription sogar umfangreicher als Erhebung und Auswertung zusammen ausfallen.

Orientierungsstruktur (im Sinne einer reflektierenden Interpretation) vorgenommen wird. Durch ein geeignetes Videotranskriptionssystem – wie etwa MoViQ (Hampl, 2010; Przyborski & Wohlrab-Sahr, 2009), das ich in der Folge noch näher erläutern werde – wird sichergestellt, dass der Grundmodus des Ausgangsmediums durch die Transkription weitgehend beibehalten wird. Dabei ist wesentlich, dass „die visuelle Ebene durch Visuelles und die verbale Ebene durch Textförmiges repräsentiert wird“ (Bohnsack, 2009, S. 171). Letztlich können Videotranskripte daher auch in Analogie zu Texttranskripten verstanden werden, die sich als Übertragungsmedien von Interviews und Gruppendiskussionen in der Forschungspraxis bewährt haben und in dieser Funktion allgemein akzeptiert werden.

Zusammenfassend stellt die reine „Wiedergabe von Datenmaterialien“ in Form von Videotranskripten folglich keine Interpretation dar, solange sie „keine Explikation implizierter Bedeutungsgehalte vornimmt“ (Bohnsack, 2009, S. 170). Innerhalb dieser klar umrissenen Grenzen ist somit die Verwendung von Videotranskripten gegenüber Film- bzw. Videomaterial als adäquates empirisches Kommunikations- und Entscheidungsinstrument für darauf aufbauende Sampling- und Interpretationsprozesse anzusehen.

### **3.2 Das Videotranskriptionssystem MoViQ**

Das Transkriptionssystem MoViQ („Movies and Videos in Qualitative Social Research“) (Hampl, 2010; Przyborski & Wohlrab-Sahr, 2009) wurde mit der Motivation entwickelt, bewegte Bilder möglichst verlustfrei und forschungsökonomisch in Einzelbildserien sowie in dazu synchrone (Sprech-)Textzeilen umzusetzen. Zur Erhöhung der Lesbarkeit und zur schnelleren Erfassbarkeit von Videotranskripten wurde dabei besonderer Wert auf die formale Ähnlichkeit zu bekannten visuellen Darstellungsformen wie Comicstrips oder Fotostories gelegt. Des Weiteren wurde bei der Konzeption – in Anlehnung an die aus der Musik bekannte Partiturschreibweise – auf die Einhaltung eines festen *Bildtakts* (Hampl, 2010, S. 56) bzw. *Zeitrhythmus* (Bohnsack, 2009, S. 171) geachtet. Deleuze (1996) sieht die festen Intervalle zwischen den Bildern<sup>21</sup> sowie die fixe Koppelung von Bild und Ton als grundlegende Medieneigenschaften von sowohl Film als auch Video an. Aus methodologischer Sicht können die Relationen zwischen den Einzelbildern erst durch das bildliche Festhalten von Einzelbildserien systematisch rekonstruiert werden

---

<sup>21</sup> Ursprünglich wurde ein Bildtakt von 16, 18 oder 24 Bildern pro Sekunde durch die mechanischen Apparaturen der Aufnahme und Wiedergabe von Filmen gewährleistet. Aber auch im digitalen Zeitalter werden Filme mit fixen Bildraten von 24 bzw. 25 Bildern pro Sekunde aufgenommen und projiziert.

(Bohnsack, 2009, S. 172; Hampl, 2010, S. 54). Das konstitutive Verhältnis zwischen visuellen und auditiven Filmanteilen lässt sich mittels Videotranskripten durch die synchrone Abbildung von Bild- und Tonspur bestimmen (Bohnsack, 2009, S. 171). Da die Erstellung von Videotranskripten ohne Softwareunterstützung mit hohem Zeit- und Arbeitsaufwand verbunden ist, habe ich im Rahmen meiner Dissertation – zur technischen Unterstützung des Transkriptionsprozesses nach dem System MoViQ – die Software MoviScript (Hampl, 2008) entwickelt sowie ein Internetforum mit knapp 250 registrierten Nutzern<sup>22</sup> aufgebaut. MoviScript ist bisher in zahlreichen Publikationen zum Einsatz gekommen (etwa: Baltruschat, 2010; Bohnsack, 2009; Hampl, 2010; Sobotka, 2009).

### **3.2.1 Gütekriterien der Videotranskription**

Zur Beurteilung der Qualität von Videotranskriptionssystemen haben Przyborski und Wohlrab-Sahr (2009, S. 164) insgesamt fünf Gütekriterien formuliert. Diese möchte ich nun anhand meiner praktischen Erfahrungen mit dem Videotranskriptionssystem MoViQ erläutern.

#### **3.2.1.1 Praktikabilität**

Ein Transkriptionssystem ist dann praktikabel, wenn es sich in der Forschungspraxis leicht einsetzen lässt. Was nützt das eleganteste Transkriptionssystem, wenn seine Anwendbarkeit an technischen Hürden scheitert? Die Erstellung von Videotranskripten ist weit aufwändiger als jene von Texttranskripten. Damit stellt das Gütekriterium der Praktikabilität die Videotranskription prinzipiell vor hohe Herausforderungen. Denn der Herstellungsprozess von Videotranskripten besteht aus einer nicht unbeträchtlichen Zahl an Arbeitsschritten, die zum Teil fortgeschrittene Computer- und Videoschnittkenntnisse erfordern. Auch die Softwareanforderungen sind hoch: Während für die Anfertigung von Texttranskripten ein einfaches Textverarbeitungsprogramm ausreicht, benötigt man für Videotranskripte einen beachtlichen Fächer an Software zur Lösung unterschiedlicher Teilaufgaben. Im komplexen Zusammenspiel der Komponenten können auch zahlreiche Inkompatibilitäten auftreten, die von den Anwender/inne/n erst gelöst werden müssen, um am Ende ein erstes Rohtranskript zu erhalten. Aus diesem Grunde habe ich mich im Zuge meiner Forschungsarbeit der Programmierung der erwähnten Videotranskriptionssoftware MoviScript gewidmet.

---

<sup>22</sup> MoviScript ist auf dieser Internetseite ([www.moviscript.net](http://www.moviscript.net)) kostenlos abrufbar.

Zwar lassen sich MoViQ-Transkripte auch manuell erstellen, forschungspraktisch wirklich praktikabel ist die Anwendung des Transkriptionssystems jedoch nur, wenn sich die Umwandlung der Videodaten in Videotranskripte zumindest teilweise automatisieren lässt.

### **3.2.1.2 Ausbaufähigkeit und Flexibilität**

Ein Transkriptionssystem ist ausbaufähig und flexibel gegenüber dem Datenmaterial, wenn es prinzipiell erlaubt, sämtliche visuelle und auditive Ereignisse abzubilden. Solange es um die Erfassung von Gesprächsdaten geht, können Videotranskripte nach dem System MoViQ diesem Anspruch in umfassender Weise und in derselben Form wie Texttranskripte gerecht werden. Hinsichtlich der Wiedergabe von Ereignissen in der Bilddimension unterliegen Videotranskripte jedoch grundsätzlichen Beschränkungen. So lassen sich aus Platzgründen beispielsweise nicht alle 25 Teilbilder, die ein Video pro Sekunde enthält, im Transkript auch abbilden. Es ist jedoch als wesentliche Stärke des Transkriptionssystems MoViQ anzusehen, dass diese Möglichkeit prinzipiell besteht. Während Texttranskripte im Wesentlichen ein relativ stabiles Aussehen haben, muss man sich vermutlich hinsichtlich der Arbeit mit Videotranskripten an eine größere Bandbreite an Darstellungsformen gewöhnen. Dies hat zum einen mit der Flexibilität der Partiturschreibweise zu tun, die je nach Erkenntnisinteresse und Material das Hinzufügen beliebig vieler Spuren ermöglicht. Zum anderen kann das Fallmaterial jedoch auch erkenntnisgeleitet in variablem Detaillierungsgrad dargestellt werden. Zusammengefasst sprechen diese Punkte für den Einsatz des Transkriptionssystems MoViQ.

### **3.2.1.3 Erlernbarkeit**

Ein Transkriptionssystem ist leicht erlernbar, wenn es ohne lange Einarbeitungszeit und Lektüre von Forscher/inne/n angewendet werden kann. Diese Bedingung erfüllt das Videotranskriptionssystem MoViQ jedenfalls durch seine schlichte Art der Darstellung, die an die Abbildungslogik von Fotostories und Comicstrips anschließt. Eine Hürde in punkto Erlernbarkeit stellen jedoch die technischen Voraussetzungen dar, die zur Erstellung von Videotranskripten erforderlich sind. Der Umgang mit Computerprogrammen und digitaler Videoverarbeitung ist nicht voraussetzungsfrei. Im Wesentlichen berührt dieser Punkt jedoch dieselben Aspekte, die bereits im Zusammenhang mit dem Gütekriterium der Praktikabilität genannt wurden.

#### **3.2.1.4 Lesbarkeit**

Ein Transkriptionssystem ist gut „lesbar“, wenn sich die damit erstellten Transkripte rasch und reibungslos erfassen lassen. Dieses Gütekriterium ist beispielsweise hinsichtlich der intersubjektiven Kommunikation im Rahmen von Forschungswerkstätten relevant. Wenn ein/e Teilnehmer/in ein Videotranskript zur Besprechung von Fallmaterial vorbereitet, so ist es wesentlich, dass alle anderen Beteiligten diesem auch folgen können. Wie die Erfahrung zeigt, sind Videotranskripte, die nach dem System MoViQ erstellt wurden, intuitiv gut erfassbar und leicht lesbar. Dies hat nicht zuletzt mit der zuvor erwähnten Abbildungslogik zu tun, die an kulturbedingte Seh- und Lesegewohnheiten (*media literacy*) anschließt.

#### **3.2.1.5 Die eindeutige Zuordenbarkeit von akustischen und visuellen Ereignissen**

Ein Transkriptionssystem gewährleistet die eindeutige Zuordnung von akustischen und visuellen Ereignissen, wenn es deren synchrone Darstellung im Transkript erlaubt. Das Videotranskriptionssystem MoViQ erfüllt dieses Kriterium prinzipiell durch konsequente Orientierung an der Partiturschreibweise. Ähnlich wie beim Gütekriterium der Ausbaufähigkeit und Flexibilität besteht jedoch nicht die Möglichkeit, diese eindeutige Zuordnung im Videotranskript generell herzustellen. Vielmehr hängt der Detaillierungsgrad der Zuordnung von der Wahl des jeweiligen Betrachtungsausschnitts ab, der seinerseits vom Erkenntnisinteresse abhängt. Üblicherweise ist im Rahmen der Video- und Filminterpretation nach der dokumentarischen Methode (wie auch das in diesem Beitrag behandelte Fallbeispiel zeigt) eine direkte Bild-Text-Zuordnung von einer Sekunde ausreichend. Für die Betrachtung kürzerer Intervalle ist es erforderlich, die Anzahl der Bilder pro Sekunde zu erhöhen. Dadurch vergrößert sich zum einen aber zugleich auch der Platzbedarf des betreffenden Videotranskripts. Zum anderen kann eine zu feine Unterteilung die Lesbarkeit des Transkripts verringern.

Zusammenfassend lässt sich an dieser Stelle festhalten, dass das Videotranskriptionssystem MoViQ grundsätzlich den genannten fünf Gütekriterien für Transkriptionssysteme entspricht. Abstriche sind beim derzeitigen Stand der Technik insbesondere hinsichtlich seiner Praktikabilität zu machen. Diesbezüglich kann die Entwicklung von Videotranskriptionssoftware, wie beispielsweise MoviScript, einen Beitrag zur Überwindung technischer Hürden leisten. Hinsichtlich Ausbaufähigkeit, Flexibilität und Zuordenbarkeit von akustischen und visuellen Ereignissen stellt uns die Videotranskription vor ein klassisches Optimierungsproblem: Verbesserungen in einer

Dimension führen zu Verschlechterungen in einer anderen. Diese Problematik lässt sich nur dadurch bewältigen, dass Videotranskripte in noch viel stärkerem Maße als Texttranskripte an die Erfordernisse des Fallmaterials und des Erkenntnisinteresses angepasst werden. Die besonderen Potentiale des Videotranskriptionssystems MoViQ sind diesbezüglich seine hohe Flexibilität sowie seine leichte Erlernbarkeit und Lesbarkeit.

### **3.2.2 Herausforderungen und Potentiale der Videotranskription**

Nachdem ich im vorhergehenden Abschnitt auf die technischen Hürden der Videotranskription eingegangen bin, möchte ich mich nunmehr ihren Potentialen widmen. Zumindest aus zwei Gründen erachte ich den Einsatz von Videotranskripten als wesentliche Grundlage der Videointerpretation. Der erste und triftigste Grund ist, dass der Einsatz von Transkripten in bedeutsamem Maße die *Nachvollziehbarkeit* der Interpretationen erhöht. In ähnlicher Weise wie bei der Verwendung von Texttranskripten im Rahmen der Interpretation von Interviews und Gruppendiskussionen erleichtern Videotranskripte die erste Orientierung im Fallmaterial sowie die intersubjektive Verständigung über das Fallmaterial im Kontext von Forschungswerkstätten. Die Interpretation in der Bilddimension ist methodisch weit anspruchsvoller als in der Textdimension. Indem Videotranskripte das bewegte Bild im Modus von Bildserien darzustellen vermögen, können sie als unverzichtbarer Referenzrahmen für die darauf aufbauende Interpretation dieser Formalstruktur angesehen werden.

In diesem Zusammenhang – und das wäre der zweite Grund – stellt der Einsatz von Videotranskripten ein unverzichtbares Instrument der *Methodenentwicklung* dar. Im Vergleich zu Texten und Bildern sind Filme und Videos in hohem Maße durch die Verhältnisse zwischen verschiedenen Mediaspekten charakterisiert. In der Montage dokumentiert sich etwa das Verhältnis zwischen Bild und Ton. Jedoch ist gerade die Montage ohne Videotranskripte systematisch nur schwer erfassbar und rekonstruierbar. Zwar lassen sich Videos oder Filme zu jedem beliebigen Zeitpunkt anhalten oder zur nochmaligen Vorführung zurückspulen. Es ist jedoch nicht möglich zugleich Bild- und Ton in ihrem Verhältnis zueinander festzuhalten. Wird das Video angehalten, so bleibt es stumm; läuft es ab, so verändert es sich ständig. Die einzige Lösung dieser paradoxen Konstellation ist es, Bild- und Tondimension von einander zu trennen und auf jeweils medienspezifische Weise zu transformieren. Dabei wird das Videobild in die geordnete

Reihe seiner Standbilder zerlegt, während die gesprochene Sprache in gedruckten Text umgewandelt wird. Durch die (Wieder-)Zusammenfügung der transformierten Teile kann im Modus der Simultaneität ein neues Bild entstehen, das nun sowohl die Bewegungen im Bild wie auch in der Sprache darzustellen vermag – das Videotranskript.

### 3.2.3 Fallbeispiel

Abbildung 6: Videotranskript der Eingangspassage (Ausschnitt) der TV-Sendung „Istanbul Total“ (2004)

Projekt „Das Verhältnis von Türkischem und Deutschem im Fernsehen“

Passage (oder Sequenz):	Ausschnitt Eingangspassage
Film (oder Video):	Istanbul Total – Unser Außenstudio
Dateiname oder URL:	<a href="http://tvtotal.prosieben.de/tvtotal/videos/player/?contentId=14144">http://tvtotal.prosieben.de/tvtotal/videos/player/?contentId=14144</a>
Datum der Aufnahme:	10.05.2004
Time Code:	1:35–2:24 min.
Dauer:	1:50 min.
Transkription:	Stefan Hampl
Korrektur:	Stefan Hampl

TC:	95 sec.	96 sec.	97 sec.	98 sec.	99 sec.
					
Am:	Christof Daum ist vorzeitig Meister geworden hier in der Türkei Fenerbahçe liegt übrigens drüben auf der asia-				
Musik:					
Geräusch:					
Sequenz:					

TC:	100 sec.	101 sec.	102 sec.	103 sec.	104 sec.
					
Am:	tischen Seite ähh da sind diese Anhöhen zu sehen da hinter dem ersten Hügel				
Musik:					
Geräusch:					
Sequenz:					

TC:	105 sec.	106 sec.	107 sec.	108 sec.	109 sec.
					
Am:	ist das Fenerbahçe-Stadion und ähh		da war natürlich gestern was los hier in der Stadt das können Sie sich gar nicht		
Musik:					
Geräusch:					
Sequenz:	Eingelagerte Sequenz (ES)				

TC:	110 sec.	111 sec.	112 sec.	113 sec.	114 sec.
					
Am:	vorstellen überall üü üüü es wurde gehupt und gefeiert ich dachte im ersten Moment ich bin doch schon seit				
Musik:					
Geräusch:					
Sequenz:					

TC:	115 sec.	116 sec.	117 sec.	118 sec.	119 sec.
					
Am:	ein paar Tagen da warum erst heute die Freude ja? aber es handelte sich äh				
Musik:					
Geräusch:					
Sequenz:					

TC:	120 sec.	121 sec.	122 sec.	123 sec.	124 sec.
					
Am:	um die Fans von Fenerbahçe die dort gefeiert haben und äh hier ist einiges los ich weiß nicht				
Musik:					
Geräusch:					
Sequenz:					

TC:	125 sec.	126 sec.	127 sec.	128 sec.	129 sec.
					
Am:	wer schon mal in Istanbul war von Ihnen ich kann Ihnen nur empfehlen fahren Sie hier hin wenn sie hier kriegen sie einfach alles				
Musik:					
Geräusch:					
Sequenz:					

TC:	130 sec.	131 sec.	132 sec.	133 sec.	134 sec.
					
Am:	überall Händler auf der Straße die die Jacke aufmachen und so sagen [sag mal brauchst du 12 Punkte für Max ja?				
Musik:					
Geräusch:					
Sequenz:					

[Am spricht mit türkischem Akzent]

TC:	135 sec.	136 sec.	137 sec.	138 sec.	139 sec.
					
Am:	isch habe dabei hier] und		ähh	fantastisch wir ham ne Schiffahrt gemacht entlang des	
Musik:					
Geräusch:	[Am spricht mit türkischem Akzent]				
Sequenz:					

TC:	140 sec.	141 sec.	142 sec.	143 sec.	144 sec.
					
Am:	Bosporus eine kleine Tour hinten wollt ma durch die Meerenge und sind leider stecken geblieben				
Musik:					
Geräusch:					
	U n t e r s e q u e n z ( U S )				

### 3.2.4 Aufbau des Videotranskripts

Wie der Blick auf das exemplarische Videotranskript (Abbildung 6) verdeutlicht, werden darin kaum Aussagen über die technische Herstellung des zugrundeliegenden Fallmaterials gemacht (bspw. zur Kameraeinstellung). Das hat damit zu tun, dass diese Angaben letztlich für die sozialwissenschaftliche Interpretation im Rahmen der dokumentarischen Methode nicht ausschlaggebend sind. Laut Bohnsack (2009, S. 198) interessieren weniger die technischen Mittel, die zur Herstellung eines Films oder Videos als (Kultur-)Produkt beigetragen haben. Von vorrangigem Interesse sind hingegen das „gestaltete Produkt und der sich in ihm dokumentierende Habitus“ (Bohnsack, 2009, S. 198).

### **3.2.4.1 Transkriptkopf**

Zu Beginn des Videotranskripts werden, nach Przyborski und Wohlrab-Sahr (2009, S. 165), im sogenannten Transkript- oder „Transkriptionskopf“ grundlegende Angaben zu Herkunft und Einordnung der Transkription gemacht. Der Kopf eines Videotranskripts sollte im Wesentlichen folgende Informationen enthalten:

1. Projektbezeichnung
2. Name oder Kennzahl der Passage bzw. Sequenz
3. Name bzw. Bezeichnung des Falles / Dateiname oder URL des Fallmaterials
4. Datum der Aufnahme
5. Time Code, der angibt, wo sich die Stelle im Videomaterial befindet
6. Dauer der Passage
7. Name des Transkribenten/der Transkribentin
8. Name des Korrekturlesers/der Korrekturleserin

Das Transkript stammt aus der Sendung „Istanbul Total“. Die Projektbezeichnung könnte lauten „das Verhältnis von Türkischem und Deutschem im Fernsehen“. Die vorliegende Transkriptpassage bzw. -sequenz habe ich als „Ausschnitt der Eingangspassage“ bezeichnet, da es sich bei ihr um einen Teil der längeren Eingangspassage des untersuchten Videos handelt. Das Video trägt den Namen „*Istanbul Total – Unser Außenstudio*“. Diese Bezeichnung geht auf den Fernsehsender Pro Sieben als Kollektiv abbildender Bildproduzent/inn/en zurück. Zur genaueren Identifizierung des Fallmaterials kann entweder der Name oder die Bezeichnung des Falles angegeben werden. Eine alternative Möglichkeit stellt die Angabe des Dateinamens bzw. der Internetadresse (URL) des Videos dar. Bei Videos, die als Medien der Massenkommunikation fungieren, erleichtert die Angabe der URL das spätere Auffinden des Fallmaterials. Der Punkt „Datum der Aufnahme“ enthält die Information über den Entstehungszeitpunkt des Films oder Videos. Im Feld „Time Code“ sind die konkreten Zeitmarkierungen angegeben, zwischen denen das Fallmaterial der Passage bzw. Sequenz transkribiert wurde. Darunter ist die resultierende „Dauer“ der transkribierten Passage bzw. Sequenz angegeben. Im vorliegenden Fall wurden zwischen den Time Codes 95 und 144 Sekunden insgesamt 110 Sekunden (1:50 Minuten) transkribiert. Als Ersteller des Videotranskripts habe ich im Feld „Transkription“ meinen Namen eingetragen. Die Korrektur wurde ebenfalls von mir vorgenommen.

### **3.2.4.2 MoViQ-Partitur**

Im Anschluss an den Transkriptkopf folgt das eigentliche Videotranskript, das im Sinne des Gütekriteriums der „Ausbaufähigkeit“ (Przyborski & Wohlrab-Sahr, 2009, S. 164)

wie eine Musikpartitur gestaltet ist. Prinzipiell kann diese Partitur aus einer oder mehreren Bildspuren sowie aus einer oder mehreren Ton-, Text- oder Kodierungsspuren bestehen. Das vorliegende Videotranskript besteht aus einer Bildspur und drei Tonspuren. In diesem Zusammenhang lässt sich am konkreten Fallmaterial gleich ein erstes formales Bestimmungsmerkmal des Mediums Video festmachen: Das Bild ist nie leer. In der Bilddimension ist immer etwas zu sehen<sup>23</sup>, während die Ausprägungen auf der Tonebene (Sprache, Musik, Geräusche) zeitweise auch fehlen können. Im betrachteten Videotranskript kommen beispielsweise keine Musik und nur wenige Geräusche vor. Der Sprecher (Am) hingegen ist fast ständig am Wort, was am transkribierten Sprechtext deutlich wird. In zusätzlichen Kodierungsspuren können seitens der Interpret/inn/en bestimmte Transkriptabschnitte gekennzeichnet werden. Im vorliegenden Fall wurde im Videotranskript die Zeile „Sequenz“ hinzugefügt, um hervorzuheben, wie sich das Fallmaterial in *Haupt-, Unter- und eingelagerte Sequenzen*<sup>24</sup> gliedert. Abhängig vom Erkenntnisinteresse ist im Sinne der Ausbaufähigkeit des Transkriptionssystems auch das Hinzufügen von weiteren Spuren möglich<sup>25</sup>.

Neben dem Gütekriterium der „Ausbaufähigkeit“ spricht noch ein weiteres Kriterium für die Partiturschreibweise nach dem System MoViQ: Die resultierende Abbildungslogik ermöglicht nämlich auch eine *eindeutige Zuordnung von visuellen und akustischen Ereignissen* (Przyborski & Wohlrab-Sahr, 2009, S.169), was als Grundvoraussetzung für die Interpretation des Verhältnisses von Bild und (Sprech-)Text bzw. Ton anzusehen ist. Des Weiteren stellt die feste Koppelung zwischen Bild- und Tonebene im Videotranskript – in Kombination mit der Einhaltung eines festen Bildtakts – die wesentliche formale Grundlage für die fallinterne und fallexterne komparative Analyse dar. Indem die Zeitintervalle zwischen den Bildern und den Tonereignissen konstant gehalten werden, treten Gemeinsamkeiten und Unterschiede etwa auf der Ebene der Montage deutlicher hervor. Im vorliegenden Videotranskript (Abbildung 6) ermöglicht der konstante Bildtakt von einem Bild pro Sekunde u. a. die Feststellung, dass der Moderator – gegenüber anderen Sequenzen – die meiste Zeit hindurch im Bild zu sehen ist. Damit ist im Videotranskript auf einen Blick erkennbar, wie der Moderator

---

<sup>23</sup> Dies ist *ex negativo* selbst dann der Fall, wenn das Bild schwarz ist.

<sup>24</sup> Die detaillierte Differenzierung dieser drei Sequenzarten erfolgt im Kapitel 3.3.

<sup>25</sup> Baltruschat (2010, S.30) kennzeichnet bspw. durch Pfeile im Transkript die Richtung der Kamerabewegung.

durch seine bloße Präsenz im Video Einfluss erlangt<sup>26</sup>. Die Einstellung von einem Bild pro Sekunde orientiert sich am Gütekriterium der „Lesbarkeit“ (Przyborski & Wohlrab-Sahr, 2009, S. 164) und hat sich für die ausgewogene Abbildung von Videos und Filmen bewährt. Welche genaue Zeiteinteilung für ein Videotranskript sinnvoll ist, hängt aber letztlich vom Erkenntnisinteresse ab. Stehen bspw. die Leistungen der abgebildeten Bildproduzent/inn/en stärker im Fokus bzw. interessieren die Herstellungsbedingungen spezifischer körperlicher Gesten bzw. Gebärden, so sollten kleinere Zeitintervalle gewählt werden. Bereits bei einem Bildtakt von einer halben Sekunde werden Körperbewegungen weit detailreicher aufgelöst und sind damit besser beobachtbar. Der einzige Nachteil dabei ist, dass feiner aufgelöste<sup>27</sup> Videotranskripte mehr Platz in Publikationen beanspruchen und der Text teilweise gesperrt bzw. zerdehnt gedruckt werden muss, damit er weiterhin synchron zum Bild ist.

### **3.2.4.3 Bildgröße**

MoViQ-Videotranskripte bestehen üblicherweise aus ca. fünf Bildern pro Zeile. Dieser Wert hat sich für Videomaterial im Format 4:3 als zweckmäßig erwiesen, um ein Maximum an Fotogrammen abbilden und zugleich den Inhalt der Fotogramme noch erkennen zu können. Abhängig von der Beschaffenheit des Fallmaterials kann es jedoch erforderlich sein, die Bildanzahl und -größe zu modifizieren. So kann es etwa für Spielfilmmaterial, dessen Bildformat meist breiter ist, zweckmäßig sein, die Bilder pro Zeile auf drei bis vier zu reduzieren. Dies kann auch bei Videos, die zu Forschungszwecken erstellt wurden, erforderlich sein, um die Leistungen der abgebildeten Bildproduzent/inn/en (z.B. Gebärden und gestische bzw. mimische Interaktionen) in ihrem Ablauf besser beobachten zu können (Abbildung 10). Zumindest drei Bilder pro Bildzeile sind aus Abbildungsgründen jedoch zu empfehlen, da ansonsten die Sequenzhaftigkeit des Videos durch die Abbildung verloren geht.

## **3.3 Videotranskripte als Grundlage des Samplings**

### **3.3.1 Sequenzen**

Im Falle von Videoproduktionen, die im Wesentlichen durch einen Handlungsstrang bestimmt sind und daher mit nur geringen Schnittleistungen auskommen – etwa Fernsehshows (Bohnsack, 2009; Hampl, 2010; Sobotka, 2009) und Amateurvideos

---

<sup>26</sup> Durch Feininterpretation ausgewählter Fotogramme, der Montage sowie des Sprechtextes lässt sich dieser Befund validieren (vgl. Bohnsack, 2009; Hampl, 2010).

<sup>27</sup> Das meint Videotranskripte mit einem Bildtakt von mehr als einem Bild pro Sekunde.

(Baltruschat, 2010) – lässt sich mithilfe von Videotranskripten auf einfachem Wege die Sequenzstruktur des Fallmaterials bestimmen. Nach der Einteilung von Bohnsack (Bohnsack, 2009, S. 196) kann prinzipiell zwischen „Hauptsequenzen (HS)“, „Untersequenzen (US)“ und „eingelagerten Sequenzen (ES)“ unterschieden werden. Als Hauptsequenzen lassen sich jene Teile eines Videos oder Films verstehen, die über etwaige Schnitte bzw. Einstellungswechsel hinweg auf immanenter Sinnebene<sup>28</sup> als zusammengehörend identifizierbar sind. Im vorliegenden Fallbeispiel dokumentiert sich im gesamten Ausschnitt der Eingangspassage (95–144 sec.) eine einzige Hauptsequenz: *Der Moderator Stefan Raab auf dem Balkon*.

Als Untersequenzen bezeichnet Bohnsack (2009, S. 196) solche Sequenzen, die als „spezifische Modifikationen der Hauptsequenz identifiziert werden können“. Meist handelt es sich dabei um Variationen der Kameraposition oder des Bildausschnitts, wobei die Identität der abgebildeten Bildproduzent/inn/en bestehen bleibt. Im vorliegenden Fallbeispiel lässt sich im Transkriptabschnitt 140–143 sec. eine Untersequenz erkennen: *Der Moderator Stefan Raab in Totale seitlich am Balkon aufgenommen*.

Als eingelagerte Sequenzen versteht Bohnsack (2009, S. 196) solche, die „fremde Szenerien“ zeigen, „nach denen [...] wieder zur Hauptsequenz zurückgekehrt wird“. Im vorliegenden Fallbeispiel dokumentiert sich im Transkriptabschnitt 105–108 sec. eine eingelagerte Sequenz: *Das Fenerbahçe-Stadion auf der asiatischen Seite des Bosphorus*.

Zusammenfassend zeigt sich bereits in dieser einfachen, im Videotranskript direkt erkennbaren Sequenzierung der hierarchische Grundaufbau des Fallmaterials<sup>29</sup>. Der Moderator der Fernsehsendung dominiert das Geschehen in einer einzigen Hauptsequenz (HS 95–144). Die Untersequenz (US 140–143) zeigt ihn noch einmal, nur aus anderer Perspektive. Nur in einer einzigen Sequenz (ES 105–108) ist der Moderator nicht zu sehen. Bei ihr handelt es sich allerdings um eine eingelagerte Sequenz, auf die unvermeidbar wieder die Hauptsequenz folgt<sup>30</sup>.

### 3.3.2 Einstellungen und Fotogramme

Bohnsack (2009, S. 201) empfiehlt zur detaillierteren Interpretation der Einstellungen die Auswahl repräsentativer Fotogramme. Ausgewählt werden sollen dabei jene

---

<sup>28</sup> Im Falle von Bildern und Videos umfasst dies die vorikonografische und ikonografische Bildebene.

<sup>29</sup> Detailliertere Rekonstruktionen des Fallmaterials finden sich in Bohnsack (2009) und Hampl (2010).

<sup>30</sup> Der Schnitt fällt bei der Betrachtung des Videos durch den mangelnden Farbkontrast des Einstellungswechsels kaum auf. Dieses Phänomen wird anhand der Musikvideos „Araba“ (Kap. 7 und 9) noch genauer erläutert.

Fotogramme, welche „die umfangreichste Hauptsequenz in ihrer häufigsten Einstellungsvariante (HS IA) und in ihrer zweithäufigsten Einstellungsvariante (HS IB) [...] zu repräsentieren vermögen“ (Bohnsack, 2009, S. 201). Die Auswahl kann mithilfe von Videotranskripten geleistet werden. Im vorliegenden Fall wäre als häufigste Einstellungsvariante der Moderator in Frontalaufnahme anzusehen, als zweithäufigste der Moderator seitlich am Balkon (US 140–143).



Abbildung 7: v.l.n.r. Fotogramme 112 und 140

Wie an dieser Stelle deutlich wird, ist mit der Auswahl bestimmter Fotogramme zugleich eine bestimmte Richtung der Interpretation vorweggenommen. Im folgenden Abschnitt möchte ich nun auf weitere Möglichkeiten eingehen, wie auf der Basis von Videotranskripten das Fallmaterial interpretiert werden kann.

### 3.4 Videotranskripte als Grundlage der Interpretation

#### 3.4.1 Fokussierte Gesten bzw. Gebärden

Mithilfe des Videotranskripts lassen sich im Video zwei interessante Gesten bzw. Gebärden <sup>31</sup> ausmachen, die als Fokussierungen im Sinne dramaturgischer Verdichtungen bzw. Brüche (Bohnsack, 2009) anzusehen sind. Die erste Geste kann auf ikonografischer Ebene als Zeigen mit der Hand identifiziert werden. Sie ist durch das Heben und Strecken der Hand relativ rasch und einfach herzustellen. Im Videotranskript

---

<sup>31</sup> Um der inhaltlichen Differenzierung zwischen vorikonografischer und ikonografischer Interpretation hinsichtlich der Beschreibung der Leistungen der abgebildeten Bildproduzent/inn/en Rechnung zu tragen, werden in dieser Arbeit die – im Alltag oft synonym gebrauchten – Begriffe „Gebärde“ und „Geste“ analytisch unterschieden. Als Gebärden werden Körperhaltungen oder -bewegungen auf vorikonografischer Ebene aufgefasst. Als Gesten sind bereits mit Common-Sense aufgeladene Gebärden auf ikonografischer Ebene zu verstehen. In diesem Verhältnis können Gesten (funktional) aus einer oder mehreren Gebärden bestehen. Im Gegensatz zum Deutschen gibt es im Englischen und Französischen (wie auch in vielen anderen Sprachen) für beide Begriffe übrigens nur ein Wort: „geste“.

ist diese Geste zirka sieben Sekunden lang zu sehen (Fotogramme 98–104). Da sich die einzelnen Fotogramme dieser Bildfolge kaum unterscheiden, reicht zur detaillierteren Interpretation dieser Geste im Wesentlichen die Auswahl eines einzelnen Fotogramms (Abbildung 8). Dieses kann nun in weiterer Folge einer klassischen Bildinterpretation zugeführt werden.



**Abbildung 8: Fotogramm 100**

Die zweite Geste lässt sich auf ikonografischer Ebene als „Jacke aufmachen und Fingerzeigen“ identifizieren. Prinzipiell nimmt sie im Video in etwa denselben zeitlichen Raum ein, nämlich zirka sechs Sekunden (Fotogramme 131–136), nur ist sie hinsichtlich ihres formalen Aufbaus komplexer als die erste. Bei genauerer Betrachtung lässt sich die zweite auffällige Geste des Videos nämlich nur durch die Einhaltung einer zeitlichen Ordnung von Teilgesten bzw. Gebärden herstellen (zuerst den Mantel öffnen, dann auf die Innenseite des Mantels zeigen). Zur detaillierten Rekonstruktion des performativen Ablaufs der Geste reicht die Auswahl eines einzelnen Fotogramms daher nicht aus. Es bedarf einer sequenziellen Abfolge mehrerer Bilder, um die Herstellung der Geste nachzuvollziehen (vgl. Abbildung 9, aber auch Abbildung 2). Aus dem Videotranskript kann dieser fokussierte Sequenzausschnitt zur Interpretation ausgewählt werden, um im Vergleich der minimalen Variationen von Bild zu Bild ein tieferes Verständnis der Herstellungsbedingungen der Geste (bzw. der damit in Zusammenhang stehenden Gebärden) zu gewinnen.

Abbildung 9: Sequenz „Jacke aufmachen“



### 3.4.2 Videotranskription und Montage

Die Interpretation der Montage bzw. Bildmischung ist insbesondere zur Rekonstruktion jener Filme und Videos erforderlich, die Produkte professioneller bzw. kommerzieller Film- bzw. Videoproduktionen darstellen sowie prinzipiell auf die Massenkommunikation gerichtet sind. Derartige Fallmaterial durchläuft von der Planung bis zur Postproduktion vielzählige Stationen, über die sich der kollektive Habitus der abbildenden Bildproduzent/inn/en im Material verdichtet. Der Ort, an dem sich dieses hohe Maß an Aufmerksamkeit seitens der abbildenden Bildproduzent/inn/en vorrangig dokumentiert, ist nach Auffassung der Filmwissenschaft (etwa: Mikos, 2003; Pudowkin, 1999; Schumm, 2004; Wiedemann, 2005) die Montage. Diese wird als wesentlichstes Instrument zur Schaffung der filmischen Gesamtkomposition angesehen.

Im Videotranskript findet die Montage in der Bilddimension auf zweierlei Arten Ausdruck: zum einen in den Schnitten zwischen den Einstellungen, zum anderen in den Farb- bzw. Helligkeitskontrasten, die durch diese Schnitte hervorgerufen werden. In Abbildung 6 erfolgen etwa zum einen Schnitte beim Sequenzwechsel zur ES am Übergang von sec. 104 zu sec. 105, zum anderen beim Sequenzwechsel zur US am Übergang von sec. 139 zu sec. 140. Der Einstellungswechsel des Fotogramms von sec. 104 zu sec. 105 ist eher unauffällig, da der Übergang mit keinem Farb- bzw. Helligkeitswechsel einhergeht. Demgegenüber fällt der Schnitt beim Sequenzwechsel 139–140 sec. deutlicher aus, da hier sowohl Farbtöne als auch Helligkeitswerte der

aufeinanderfolgenden Einstellungen eine Veränderung erfahren<sup>32</sup>: Die Untersequenz (140–143 sec.) ist dunkler und „violetter“ als der Rest der Hauptsequenz.

Des Weiteren dokumentiert sich die Montage (und damit die Eingriffe der abbildenden Bildproduzent/inn/en) im Videotranskript auch an den Stellen, wo Bild- und Tondimension technisch entkoppelt werden. Im vorliegenden Beispiel ist etwa beim Einstellungswechsel von Hauptsequenz (HS) zur eingelagerter (ES) oder Untersequenz (US) zwar ein Wechsel in der Bilddimension zu beobachten, aber keiner in der Tondimension. Die Stimme des Moderators ist durchgehend zu hören, was ein weiterer formaler Beleg für seine zentrale Bedeutung im Video ist.

Bei Videos, die zu Forschungszwecken erstellt wurden, spielt die Montage in der Regel keine wesentliche Rolle. Die Leistungen der abbildenden Bildproduzent/inn/en – der Forscher/innen – beschränken sich meist auf die Wahl des Kamerastandorts und -ausschnitts. Diese Beiträge von Seiten der Forscher/innen sollte im Sinne eines methodisch kontrollierten Fremdverstehens reflektiert werden. Da prinzipiell jedoch die Leistungen der abgebildeten Bildproduzent/inn/en das vorrangige Interesse der Forscher/innen darstellen, greifen Letztere in der Regel nicht weiter in die Gestaltung des Materials ein<sup>33</sup>. Wenngleich das Erfordernis der Rekonstruktion der Montage mithilfe von Videotranskripten wegfällt, können diese im Falle von Videos, welche zu Forschungszwecken aufgenommen wurden, jedoch eine wesentliche Unterstützung beim Aufspüren, Dokumentieren und Rekonstruieren fokussierter Gesten bzw. Gebärden der abgebildeten Bildproduzent/inn/en sein.

Einen beachtenswerten Spezialfall hinsichtlich der Rekonstruktion der Montage stellen Amateur- bzw. Privatvideos dar (etwa: Baltruschat, 2010). In Hinblick auf die Verwendung des Instrumentariums des Schnitts unterschieden sich Amateurvideos nicht grundsätzlich von professionellen Video- oder Filmproduktionen, weshalb die Rekonstruktion der Montage auch für diese Materialgattung wesentlich erscheint.

Der maßgebliche Unterschied zwischen professionellen und amateurhaften Videoproduktionen besteht laut Frost (2009) in der Farbkomposition. Während diese bei Amateurvideos eher zufällig entsteht oder bestenfalls intuitiv über das Arrangement der Einstellungen erfolgt, werden die Farben in professionellen Filmproduktionen oft

---

<sup>32</sup> Noch weit drastischere Schnitteffekte sind durch die Montage von Einstellungen zu erreichen, die zueinander farblich oder helligkeitsmäßig in einem komplementären Verhältnis stehen (Kap. 7.3.3).

<sup>33</sup> Eine diesbezügliche Ausnahme stellt der ethnografische Zugang von Mohn (2011, S. 95) dar, die „fokussiertes Schneiden“ durch die Forscher/innen als erkenntnisgenerierende Option hervorhebt.

bereits im Drehbuch festgelegt<sup>34</sup>. Dies geht laut Frost (2009) sogar so weit, dass Farben und Farbkontraste für bestimmte Filmgenres typisch geworden sind. Dadurch ist die Farbwahl indexikal und kann in dieser Weise auch als visuelles Zitat eingesetzt werden<sup>35</sup>. Da sich Farben, Farbkontraste und Helligkeitswerte als visuelle Muster in Videotranskripten dokumentieren, können diese im Zuge der komparativen Analyse auch miteinander verglichen werden. In Kapitel 5 habe ich durch systematischen Vergleich des Videotranskripts der Sendung „Istanbul Total“ (TC 00–99 sec.) mit Transkripten anderer Fernsehsendungen demonstrieren, dass die feststellbar geringe Schnittanzahl in Kombination mit der zentralen Moderatorstellung ein typisches Merkmal der Gattung Late-Night-Shows ist. In Anknüpfung an diese Vorgehensweise könnte auf der Basis des vorliegenden Videotranskripts nun ein weiterer Schritt die Untersuchung der Indexikalität der Farben des Fallmaterials sein. Laut Frost (2009) sind kräftige Farben jedenfalls als typisches Merkmal von Hollywood-Komödien anzusehen. Die empirische Überprüfung dieser Aussage könnte durch einen systematischen Vergleich des Fallmaterials (Abbildung 6) mit Videotranskripten von ebensolchen Komödien geleistet werden.

### **3.5 Anonymisierung**

Im Rahmen der Auseinandersetzung mit Film- und Videomaterial aus den Massenmedien stellt sich die Frage der Anonymisierung nur in geringem Maße. Bei abgebildeten Moderator/inn/en oder anderen Personen des öffentlichen Lebens (Politiker/innen, Schauspieler/innen etc.) wird im Allgemeinen davon ausgegangen, dass diese mit der massenmedialen Verbreitung ihres medialen Auftritts einverstanden sind. In Zweifelsfällen wären zuerst die abbildenden Bildproduzent/inn/en gefordert, eine unvorteilhafte oder für die Betroffenen problematische Szene aus dem Material zu entfernen, bevor es gesendet wird. Fernsehsender und Filmproduktionsstudios sind als abbildende Bildproduzent/inn/en sind üblicherweise penibel darauf bedacht, sich rechtlich gegenüber den abgebildeten Bildproduzent/inn/en abzusichern. Dies geschieht auch gegenüber dem Publikum, das sich in der Regel nicht aus Prominenten, sondern Privatpersonen zusammensetzt. Wie ich selbst bereits mehrmals selbst erleben

---

<sup>34</sup> Die Auswahl und systematische Kontrolle von Farben, Farbkontrasten sowie Helligkeitsniveaus stellen laut Frost (2009) und Arnheim (1979) unverzichtbare Mittel der Sinnvermittlung von professionellen Video- bzw. Filmproduktionen dar, weil diese in wesentlichem Maße zur formalen Fokussierung der inhaltlichen Botschaft beitragen.

<sup>35</sup> Wer heute etwa einen Film in Schwarz-Weiß dreht, zitiert damit auch die Ära der Schwarz-Weiß-Filme (bspw. „The man who wasn't there“, 2001, oder „The artist“, 2011).

konnte, wird man als Zuseher – ohne vorhergehende Unterzeichnung einer Einverständniserklärung, in der sämtliche mit der audiovisuellen Aufzeichnung im Zusammenhang stehende persönliche Bildrechte an den Fernsehsender abgetreten werden – gar nicht erst ins Aufnahmestudio vorgelassen.

Anders verhält es sich mit Videomaterial, welches zu Forschungszwecken erstellt wurde. Dieses sollte sowohl auf der Bild- als auch auf der Textebene modifiziert werden, um im Sinne des Datenschutzes Außenstehenden eine Zuordnung videografisch erfasster Personen und Orte zu verunmöglichen. Hinsichtlich der Datenpräsentation in Publikationen berührt diese Forderung auch die Erstellung von Videotranskripten. Während die Anonymisierung auf der Ebene des Sprechtextes in der Regel bereits durch Veränderungen der Eigennamen zu erreichen ist, erfordert die Anonymisierung in der Bilddimension fortgeschrittene Kenntnisse von Bild- bzw. Videobearbeitungsprogrammen. Das Repertoire an möglichen Eingriffen ins Bild reicht vom schlichten Verdecken der Augen der Abgebildeten durch Farbbalken (etwa: Bohnsack, 2009, S. 98) über das Weichzeichnen der Gesichtszüge (etwa: Dinkelaker & Herrle, 2009) bis zum groben Verpixeln der Gesichter (Abbildung 10). Von Fall zu Fall wird zu entscheiden sein, welche der genannten Maßnahmen zur Anonymisierung ausreichend ist.

Einen Graubereich hinsichtlich des Erfordernisses zur Anonymisierung stellen Amateur- oder Privatvideos dar, die zur Auswertung im Rahmen von Videointerpretationen herangezogen werden. Selbst wenn die an der Produktion beteiligten Personen vor und hinter der Kamera ihr Einverständnis zur Verwendung des Materials gegeben haben, ist damit noch nicht geklärt, ob sie auch mit der Interpretation des Materials einverstanden sind. Videotranskripte sind in besonderem Maße von diesem Umstand berührt, da sie zur Abbildung, Auswahl und Interpretation des Datenmaterials herangezogen werden können.

Der große Nachteil der Anonymisierung der Daten in der Bilddimension ist, dass mit dem Gesicht zwangsläufig auch die interpretationsrelevante Mimik der abgebildeten Bildproduzent/inn/en verdeckt wird. Bei optimaler Anonymisierung des Bildmaterials (etwa durch grobes Verpixeln) sind oft nicht einmal mehr Gebärden gut erkennbar. Dieser Umstand kann durch das folgende Beispiel aus der Unterrichtsforschung illustriert werden, das mir freundlicherweise von Maria Juen (2012) zur Verfügung gestellt wurde. Zwar sind die einzelnen Schüler/innen im Bild nach der Anonymisierung nach wie vor sichtbar. Wie die nachfolgende Abbildung zeigt, geht dadurch aber gerade

der interaktive Gehalt der Sequenz – der durch die Simultaneität von Mimik und Gestik bestimmt ist – größtenteils verloren.



Abbildung 10: Die ersten Minuten des Unterrichts (Juen, 2012)

### 3.6 Zusammenfassung

Bohnsacks (2009, S. 171) Verständnis nach sind Videotranskripte nach dem System MoViQ – ebenso wie Texttranskripte – nicht als Explikationen, sondern lediglich Übertragungen bzw. Wiedergaben von Datenmaterial anzusehen. Im Anschluss daran wurde anhand der Darstellung der Gütekriterien für Transkriptionssysteme (Przyborski & Wohlrab-Sahr, 2009, S. 164) eine Evaluierung des Videotranskriptionssystems MoViQ vorgenommen. Einschränkend wurden lediglich die hohen technischen Anforderungen der Methode der Videotranskription an die Forscher/innen festgestellt. Erst durch die Entwicklung spezifischer Softwarelösungen im Rahmen der Methodenentwicklung werden Videotranskriptionen praktikabel durchführbar. Das Programm MoviScript, das im Rahmen dieser Arbeit entwickelt wurde, stellt dazu einen ersten Ansatz dar.

Wie des Weiteren anhand des Fallbeispiels einer Fernsehsendung deutlich gemacht werden konnte, eignen sich MoViQ-Videotranskripte nicht nur zu Abbildungs-, sondern auch zu Sampling- und Interpretationszwecken. In dieser Eigenschaft ermöglichen sie – in vergleichbarer Weise wie Texttranskripte – zum einen die interindividuelle Kommunikation bei der gemeinsamen Bearbeitung von Fallmaterial im Kontext von Forschungswerkstätten. Zum anderen erhöhen sie in bedeutsamer Weise die Nachvollziehbarkeit der Interpretationsschritte, indem sie eine simultane sprachlich-bildliche Referentialität zu audiovisuellem Fallmaterial herstellen.

Am Beispiel der Interpretation der Montage konnte der wesentliche Beitrag von Videotranskripten zur Methodenentwicklung deutlich gemacht werden: Durch die simultane Abbildungsweise von Bildserien und dazu synchronem Ton bzw. Sprechtext

ermöglichen Videotranskripte die Einnahme zusätzlicher Betrachterperspektiven gegenüber dem Fallmaterial und damit neue Zugänge zu dessen Rekonstruktion. Beispielweise lassen sich die im „Bild“ des Videotranskripts fixierten Montagemuster im Rahmen der komparativen Analyse mit anderem Fallmaterial vergleichen, was eine Gattungsbestimmung ermöglicht. So lässt sich z. B. die ursprünglich rein metatheoretisch fundierte Differenzierung von Videos (Bohnsack, 2009, S. 117), die zu Erhebungszwecken erstellt wurden, und solchen, die als private oder öffentliche Alltagsdokumente angesehen werden können, auf der Basis von Videotranskripten formal empirisch rekonstruieren.

Im Zusammenhang mit den drei von Bohnsack (2009) genannten Typen von Fallmaterial (Erhebungsinstrument, private bzw. öffentliche Alltagsdokumente) wurde abschließend noch auf die komplexe Problematik der Anonymisierung von Videotranskripten aufmerksam gemacht. Die Verantwortung der Forscher/innen für die Anonymisierung des Fallmaterials besteht nicht nur dann, wenn dieses von ihnen selbst als Erhebungsinstrument produziert wurde. Auch in der Auseinandersetzung mit privaten Alltagsvideos bzw. Amateurvideos sollte über Fragen der Anonymisierung nachgedacht werden. Selbst wenn ein Einverständnis der Bildproduzent/inn/en zur Verwendung des Fallmaterials im Rahmen wissenschaftlicher Arbeit vorliegt, ist damit noch nicht automatisch ein Einverständnis zu dessen Interpretation verbunden. Eine zweite Schwierigkeit der Anonymisierung hat formale Gründe und betrifft die Bilddimension. Im Gegensatz zur Anonymisierung in der Textdimension kann durch Anonymisierung der Bilddimension die Aussagekraft des Fallmaterials stark beeinträchtigt werden. Diesbezüglich ist die Forschungspraxis in ähnlicher Weise auf die Entwicklung neuer technischer Ansätze angewiesen, wie sie geeigneter Softwareunterstützung zur Erstellung von Videotranskripten an sich bedarf.

#### **4 Auswahl des Fallmaterials**

Die Auswahl des Fallmaterials wurde grundsätzlich nach den Prinzipien des *Theoretical Samplings* (Przyborski & Wohlrab-Sahr, 2009, S. 177) durchgeführt. Folglich wurden die einzelnen Videos dieser Untersuchung noch nicht zu Beginn festgelegt, sondern auf der Basis des methodologischen Erkenntnisinteresses dieser Arbeit Zug um Zug nach den Kriterien des maximalen und minimalen Kontrasts ausgesucht.

Den Ausgangspunkt der Untersuchung stellte die Eingangspassage der Fernsehshow „Istanbul Total“ dar. Diese wird in Kapitel fünf im Sinne eines minimalen

Kontrasts mit zwei anderen Fernsehshows („TV Total“ und Harald-Schmidt-Show) sowie im Sinne eines maximalen Kontrasts mit dem ZDF-Reisevideo „Reiselust Istanbul“ verglichen. Des Weiteren werden einzelne Fotogramme der Eingangspassage dazu verwendet, mittels bildgebundener Suche im Internet planimetrisch ähnliche Bilder zur komparativen Analyse aufzuspüren.

In weiterer Folge wurde in Kapitel sechs ein Ausschnitt aus der türkischen Beyaz-Show zur Interpretation ausgewählt. Dieser Ausschnitt wurde am zweiten Sendetag in der Sendung „Istanbul Total“ gezeigt. Aus inhaltlicher Sicht stellt die Beyaz-Show einen minimalen Kontrast zu „Istanbul Total“ dar. Von methodologischem Interesse ist der Umstand, dass es sich bei der Beyaz-Show um eine „Show in der Show“ handelt. Dadurch werden Aspekte der Postproduktion sichtbar und interpretierbar. Besondere Aufmerksamkeit erfährt dabei die Rekonstruktion der Relation zwischen Bild- und Tondimension des Videos.

Die Auswahl des Musikvideos „Araba“ (1996), das in Kapitel acht untersucht wird, erfolgte aus zweierlei Gründen. Aus methodologischer Perspektive stellen Musikvideos einen maximalen Kontrast zu Fernsehshows dar, weil sie sich insbesondere hinsichtlich der Montage stark unterscheiden. Des Weiteren eignete sich das Musikvideo „Araba“ (1996) jedoch aufgrund auch besonders gut zur Demonstration von Ittens (2000, 2010) Farbkontrasttheorie. Aus inhaltlicher Sicht war die Interpretation des Musikvideos von Interesse, da eine zweite Bildversion für den internationalen Markt gedreht wurde.

Im Rahmen eines minimalen Kontrasts wurde abschließend für Kapitel neun dieses zweite Musikvideo, „Araba“ (2004), zur komparativen Analyse mit dem ersten ausgewählt. Von methodologischem Interesse waren in diesem Zusammenhang insbesondere der Vergleich der Montagen der beiden Videos (Araba 1996 und Araba 2004) sowie die Interpretation der an der Montage beteiligten Farbkontraste.



**Abbildung 11: Videotranskript der Titelsequenz von „Istanbul Total“ (2004)**

Anmerkung: Der Hinweis „Best of“ wurde in der DVD-Version der Sendung nachträglich hinzugefügt.

## 5 „Istanbul Total“: Eingangspassage

In diesem Kapitel möchte ich zum einen in das ausgewählte Fallmaterial und zum anderen in die Forschungslogik der Film- und Videointerpretation nach der dokumentarischen Methode einführen. Vorausschickend ist anzumerken, dass ich dazu vorläufig noch nicht vollständig in die einzelnen Arbeitsschritte nach Bohnsack (2009, S. 176) eintauche, sondern mich zuerst den medialen Eigenschaften des Videomaterials sowie den damit verbundenen empirischen Ansatzmöglichkeiten für die Interpretation widme. Dieses Kapitel kann damit gewissermaßen als eine Vorrede zur Film- bzw. Videointerpretation verstanden werden, die den Zweck hat für die empirischen Möglichkeiten des damit verbundenen Forschungsprogramms zu begeistern. Mein Beitrag richtet sich nicht zuletzt an all jene, denen aufgrund von zu geringer persönlicher Forschungserfahrung im Umgang mit Film- bzw. Videomaterial ein unmittelbarer Einstieg in die Materie bisher schwergefallen ist. In diesem Zusammenhang kann das hier vorgestellte Repertoire an Interpretationsinstrumenten für Studierende und „Neuankömmlinge“ in der dokumentarischen Methode auch als eine Art Kompass oder „Grundbaukasten“ für eigene erste Schritte in das Feld begriffen werden. Zugleich würde ich mich freuen, wenn dieses Kapitel auch als Inspirationsquelle für erfahrene Forscher/innen dienen könnte. In diesem Zusammenhang können die hier vorgestellten Versuche als Anregung zur eigenständigen Weiterentwicklung der methodischen Instrumente im Rahmen der dokumentarischen Methode aufgefasst werden<sup>36</sup>. Der folgende Text stellt eine aktualisierte Fassung meiner bereits an anderer Stelle entwickelten Überlegungen (Baltruschat & Hampl, 2013; Hampl, 2010) dar, die ich im Kontext der vorliegenden Gesamtarbeit zusammengefügt und nochmals überarbeitet habe.

Methodologisch beziehe ich mich dabei auf die begrifflichen Konzepte und analytischen Einteilungen, die Ralf Bohnsack (2009) – unter Berufung auf Erwin Panofsky und Max Imdahl – für die *qualitative Bild- und Videointerpretation nach der dokumentarischen Methode* erarbeitet hat<sup>37</sup>. Der Vorteil der dokumentarischen Methode besteht meines Erachtens in der anschaulichen Möglichkeit, Videointerpretationen *im*

---

<sup>36</sup> Im vorliegenden Fall wird anhand von konkretem Fallmaterial beispielsweise u. a. die bereits im Kapitel „Zur Montage“ vorgestellte Grundidee der Übertragung des Imdahl'schen Prinzips der Kompositionsvariation (Imdahl, 1994) auf die Film- und Videointerpretation exploriert.

<sup>37</sup> Zur vorbereitenden Lektüre gehörten auch Arbeiten aus den Bereichen der objektiven Hermeneutik und hermeneutischen Wissenssoziologie (etwa: Dinkelaker & Herrle, 2009; Herrle, 2007; Knoblauch, 2004; Raab, 2008; Schnettler, 2007; Schnettler & Raab, 2008; Schnettler, Tuma & Schreiber, 2010).

*Medium des Fallmaterials* durchzuführen. Videos können mittels Internetsuche rasch und direkt mit anderen Videos und Bildern verglichen werden, ohne jeweils den Umweg über die Textebene nehmen zu müssen.

Der spezifischen Eigenlogik (bewegter) Bilder möchte ich durch eine konsequente Orientierung der Arbeitsschritte an der *planimetrischen Komposition* (Kap. 1.2.2) Rechnung tragen, die „die entscheidende Grundlage für das ‚sehende Sehen‘“ (Bohnsack, 2007, S. 167) *sensu* Imdahl (1994) darstellt. Außerdem wird zu einem möglichst frühen Einsatz der *komparativen Analyse* im Rahmen der Videointerpretation ermutigt, indem diese Vorgehensweise anhand empirischer *Montage-, Einstellungs- und Gesten- bzw. Gebärdenvariationen* demonstriert wird. Auf die Interpretation des Sprechtextes sowie außersprachlicher Tonereignisse (Musik, Geräusche etc.) wird in diesem Kapitel verzichtet. Das diesbezügliche Prozedere wird jedoch an vergleichbarem Fallmaterial (Beyaz-Show) im nächsten Kapitel ausführlich erläutert.

## **5.1 Auswahl und Interpretation von Videosequenzen**

Die Frage der Auswahl geeigneter Einstellungen und Videosequenzen stellt für diejenigen, welche in der Methode der Videointerpretation noch unerfahren sind, die erste größere Hürde dar. Zwar bietet die einschlägige Fachliteratur grundsätzliche Auskünfte zur Auswahl von Vergleichsmaterial, etwa über die Orientierung an „Eingangssequenzen“ und die Formulierung der Prinzipien der „Fokussierung“, „Repräsentanz“, „kompositorischen Auffälligkeiten“ (Bohnsack, 2009, S. 174) oder der „maximalen Kontrastierung“ (Raab, 2008, S. 169). Unbeantwortet bleibt dabei jedoch, wie diese Richtlinien ohne eigene empirische Vorerfahrungen am Material angewendet werden sollen.

Meiner Erfahrung nach tendieren Studierende oft dazu, die Kluft zwischen Theorie und Praxis durch ein Mehr an Lektüre überwinden zu wollen. Daran ist die Hoffnung geknüpft, sich so „auf sicherem Wege“ dem Material zu nähern. Dieses in der Sozialforschung an sich bekannte Problem verschärft sich im Bereich der Bild- und Videointerpretation aufgrund des (noch immer verbreiteten) mangelnden Vertrauens in die Leistungsfähigkeit der Instrumente. Methodologisch als zentral erachtete Interpretationsschritte, wie etwa die Rekonstruktion der planimetrischen Komposition, wirken aus Studierendensicht auf den ersten Blick oft beliebig. Die Angst vor interpretativen Willkürakten im Visuellen führt zu einer Klammerung an Texte bzw. an die sprachlichen Äußerungen im Material: Mit beträchtlichem Mehraufwand werden

zusätzliche Common-Sense-Informationen recherchiert und gesammelt, obgleich daraus keine tiefergehenden empirischen Einsichten über das Material zu erwarten sind.<sup>38</sup>

Um solchen Vermeidungstendenzen gezielt entgegenzusteuern, halte ich es für didaktisch sinnvoll, von vornherein eine direkte visuelle Auseinandersetzung mit dem Videomaterial zu forcieren. Im Folgenden möchte ich darstellen, wie dieser Anspruch durch die Suche nach Vergleichshorizonten im Medium des (bewegten) Bildes praktisch umgesetzt werden kann. Stellt man das eigene Fallmaterial anderen Bildern und Videos gegenüber, so werden Gemeinsamkeiten und Eigenheiten oft unmittelbar deutlich und interpretierbar. Die Auswahl *ist* somit bereits Interpretation – die Interpretation ein Vorgang, der vom empirischen Vergleich lebt.

## **5.2 Exemplarische Videointerpretationen**

Um das Sensorium für die Eigengesetzlichkeiten von Bildern und Videos zu schärfen, ist intensives Experimentieren mit Fall- und Vergleichsmaterial erforderlich. Ich möchte eine Herangehensweise an Videomaterial vorstellen, die „von außen nach innen“ über die Variation von Montage, Einstellungen und Gesten bzw. Gebärden verläuft.

Die Montage umfasst laut Bohnsack (2009, S.158) die Leistungen der *abbildenden Bildproduzent/inn/en* (der Personen hinter der Kamera). Gesten und Gebärden umfassen die Leistungen der *abgebildeten Bildproduzent/inn/en* (Personen vor der Kamera). Einstellungen sind als gemeinsame Produkte von abbildenden und abgebildeten Bildproduzent/inn/en anzusehen. Durch *Triangulation*, d. h. die Betrachtung des Materials aus diesen unterschiedlichen Perspektiven, kann ein differenziertes Verständnis über die Sinnstrukturen des Videomaterials in seiner Gesamtheit gewonnen werden.

### **5.2.1 Montagevariation und Gattungsanalyse**

Im Anschluss an meine Überlegungen zur Montage in Kapitel 2 möchte ich in der Folge ein Vorgehen demonstrieren, bei dem die Interpretation der Montage *im Medium des Bildes* erfolgt. Es wird dabei prinzipiell eine Form der komparativen Analyse von Videos favorisiert, die sich an der planimetrischen Komposition (bewegter) Bilder orientiert. Die erste Auswahl von Vergleichsmaterial erfolgt aus didaktischen Gründen dennoch

---

<sup>38</sup> Beliebte Beispiele für Recherchen, mit denen die Konfrontation mit dem Videomaterial vermieden werden kann: Lesen von Filmrezensionen, Abhören der DVD-Audiokommentare der Regisseur/innen, Führen von Interviews mit den Drehbuchautor/inn/en etc.

über die ikonografische Ebene, d. h. die Ebene des Common Sense, da diese für Einsteiger/innen am zugänglichsten ist.

Videointerpretation wird von Angela Keppler (2006, S. 139; zitiert nach Bohnsack, 2009, S. 240) im Wesentlichen als *Gattungsanalyse* begriffen: Im Forschungsprozess wird nicht nur ein tieferes Verständnis über das Fallmaterial allein entwickelt, sondern es lassen sich dabei auch Gattungskriterien empirisch definieren (Keppler, 2006, S. 139). In Umkehrung des Vorgangs kann eine auf ikonografischer Ebene vorgenommene Einordnung des Fallmaterials in eine Gattung – im Sinne einer Bestimmung des Gattungsthemas (Bohnsack, 2003, S. 101) – als Suchhilfe eingesetzt werden. So entspricht im vorliegenden Fall die Fernsehsendung „Istanbul Total“ ikonografisch der Gattung *moderierte Late-Night-Unterhaltungssendungen*<sup>39</sup>. *Gattungsimmanent*, d. h. innerhalb derselben Gattung, wurden zwei weitere Sendungen zum Vergleich ausgewählt: Die Harald-Schmidt-Show (Abbildung 13) und die Sendung „TV Total“ (Abbildung 17), von der „Istanbul Total“ (Abbildung 12) ein „Ableger“ ist. Da das Vorwissen über die Gattungszuordnung im genannten Beispiel deren empirischer Bestimmung vorausgeht, ist es Aufgabe der komparativen Analyse, die a priori festgelegte Gattungszuordnung zu überprüfen. Im nächsten Abschnitt wird das genaue Prozedere vorgeführt.

Es kann vorkommen, dass Gattungszuordnungen, die auf der Common-Sense-Ebene plausibel und klar erschienen, auf der Ebene der formalen Bild- und Textstruktur des Videomaterials nicht überzeugen bzw. sich aus dem Material heraus Anknüpfungspunkte zu anderen Gattungen ergeben. Damit bietet das gattungsimmanente Vorgehen Ansatzpunkte für *gattungsübergreifende* komparative Analysen – für Letztere wurde das Reisevideo „Reiselust Istanbul“ (Abbildung 15) als Beispiel herangezogen.

In der Folge sind die ersten 1:40 Minuten der Sendung „Istanbul Total“ als Videotranskript dargestellt. Auf der linken Seite ist jeweils das fortlaufende Videotranskript der Eingangspassage zu sehen, welche das Ausgangsmaterial der vorliegenden Videointerpretation bildet. Auf der rechten Seite sind zur komparativen Analyse Vergleichssequenzen abgebildet. Diese Form der gegenüberliegenden Darstellung im Buch ermöglicht es, das Verhältnis von Ausgangssequenz und Vergleichsmaterial besonders zu verdeutlichen.

---

<sup>39</sup> Diese Gattung stammt ursprünglich aus den USA. Die Sendungen werden in der Regel spätabends ausgestrahlt und von einem Moderator in Anzug und Krawatte präsentiert.



**Abbildung 12: Istanbul Total (1/3)**

Passage (oder Sequenz): Eingangspassage  
 Film (oder Video): Istanbul Total  
 Dateiname: Istanbul Total.wmv  
 Datum der Transkription: 03.05.2009  
 Time Code: 00 - 382  
 Dauer: 382 sec.  
 Transkription: Stefan Hampf  
 Korrektur: -

TC	0 sec.	1 sec.	2 sec.	3 sec.	4 sec.
Am	Ich befinde mich hier				
Geräusche	(				
TC	5 sec.	6 sec.	7 sec.	8 sec.	9 sec.
Am	n kann hier hat man einen fantastischen Blick über die ganze Stadt				Hier ist der
Geräusche	(Motorengeräusche)				
TC	10 sec.	11 sec.	12 sec.	13 sec.	14 sec.
Am	Bosporus hier geht's ins goldene Horn ja?		Die Schiffe fahren rein s		
Geräusche					
TC	15 sec.	16 sec.	17 sec.	18 sec.	19 sec.
Am	ie fahren raus ein ständiges Heckmeck ja?		und äh da drüben kann man fantastisch		
Geräusche					
TC	20 sec.	21 sec.	22 sec.	23 sec.	24 sec.
Am	äh	ähm	die Moscheen sehen da vorne ist die Hagia Sofia eine der ält		
Geräusche					
TC	25 sec.	26 sec.	27 sec.	28 sec.	29 sec.
Am	ersten Kirchen	äh hier in Istanbul	ähm gla'b über tausend Jahre a		
Geräusche					

**Abbildung 13: Harald-Schmidt-Show**

Passage (oder Sequenz): Eingangspassage  
 Film (oder Video): Harald Schmidt Show  
 Dateiname: Harald Schmidt.avi  
 Datum der Transkription: 15.06.2009  
 Time Code: 00 - 69  
 Dauer: 69 sec.  
 Transkription: Stefan Hampl

TC	0 sec.	1 sec.	2 sec.	3 sec.	4 sec.
Am	Willkommen bitte um Verständnis für Sie zu Hause dass ich den Text hier oben ablese				aber im neuen Stud
Pmf	((verebbender Applaus))				@@
TC	5 sec.	6 sec.	7 sec.	8 sec.	9 sec.
Am	io läuft der Text hier oben rechts auf einem Laufband			Herzlich willkomme	
Pmf				@@@	
TC	10 sec.	11 sec.	12 sec.	13 sec.	14 sec.
Am	n hier im wie es jetzt neu heißt off=ziell im Ersten im Palast der Republik				@((ha ha ha))
Pmf					@@
TC	15 sec.	16 sec.	17 sec.	18 sec.	19 sec.
Am	@	Neu aufgebaut und mich ham viele Leute heut angesimst und a			
Pmf	@				
TC	20 sec.	21 sec.	22 sec.	23 sec.	24 sec.
Am	nge-MMSt und äh	Faxe geschrieben und auch persönlich vorbeigekommen und gefragt was is eigentlich oben im alten Studio ich weiß es ni			
TC	25 sec.	26 sec.	27 sec.	28 sec.	29 sec.
Am	cht aber ich habe vorhin gesehen es kam der Möbelwagen	ich glaube in unserem alten Studio genau hier in unserm Studio vier-			
Bm					ja

Abbildung 14: „Istanbul Total“ (2/3)

TC	30 sec.	31 sec.	32 sec.	33 sec.	34 sec.
Am	It und äh daneben		ist der Sultanspalast dort vorne auf dieser Anhöhe		ja? und wenn ma
Geräusche					
TC	35 sec.	36 sec.	37 sec.	38 sec.	39 sec.
Am	n mal auf die andere Seite rüberschaut			dann kann man den Bosphorus entla	
Geräusche					
TC	40 sec.	41 sec.	42 sec.	43 sec.	44 sec.
Am	ng kucken und sieht dort hinten die Bosphorusbrücke			die	Asien und Europa mi
Geräusche					
TC	45 sec.	46 sec.	47 sec.	48 sec.	49 sec.
Am	einander verbindet ja? die einzige Stadt der Welt auf zwei Kontinenten				
Geräusche					
TC	50 sec.	51 sec.	52 sec.	53 sec.	54 sec.
Am	und	da vorne kommt gerade ein Schlepper an			ja?
Geräusche					
TC	55 sec.	56 sec.	57 sec.	58 sec.	59 sec.
Am	sieht man hier unten das sieht man das sind immer die Boote die Boote mit den kann		ähm		
Geräusche	man das mal zeigen? die Boote mit diesen Motorengeräusch))				
TC	60 sec.	61 sec.	62 sec.	63 sec.	64 sec.
Am	mit diesen Reifen die schl=		die schleppen immer die großen äh Schiffe deswegn		

**Abbildung 15: Reiselust Istanbul**

Passage (oder Sequenz): Beyoğlu  
 Film (oder Video): ZDF Reiselust Istanbul  
 Dateiname: Reiselust.mpg  
 Datum der Transkription: 25.06.2009  
 Time Code: 00 - 55  
 Dauer: 55 sec.  
 Transkription: Stefan Hampl  
 Korrektur: -

TC	0 sec.	1 sec.	2 sec.	3 sec.	4 sec.
Am	Beyoğlu 24 saat yaşar... ((spricht auf Türkisch))				
Bm	Beyoğlu lebt vier				
TC	5 sec.	6 sec.	7 sec.	8 sec.	9 sec.
Am	spricht weiter auf Türkisch; Lautstärke leiser geregelt))				
Bm	undzwanzig Stunden am Tag.		Hier gibt es alles		soziale kulturelle und
TC	10 sec.	11 sec.	12 sec.	13 sec.	14 sec.
Am	Freizeitaktivitäten		eigentlich bräuchten die Leute ihr Viertel gar nicht zu verl		
Bm					
TC	15 sec.	16 sec.	17 sec.	18 sec.	19 sec.
Am	assen		sie haben ja alles		Kino Theater Bars Restau
Bm					((Läuten))
TC	20 sec.	21 sec.	22 sec.	23 sec.	24 sec.
Am	rants		seit über einhundert Jahren ist Beyo		
Bm					
TC	25 sec.	26 sec.	27 sec.	28 sec.	29 sec.
Am	glu das Unterhaltungszentrum Istanbuls und der Türkei				
Bm					

Abbildung 16: Istanbul Total (3/3)

TC	65 sec.	66 sec.	67 sec.	68 sec.	69 sec.
Am					
Geräusche	gibt's auch hier den äh Spruch Nepper-Schlepper-Bauernfänger @ja@ ja und hier war viele				
Pmf	((Mot				
TC	70 sec.	71 sec.	72 sec.	73 sec.	74 sec.
Am					
Geräusche	icht was los ich weiß nicht wie das bei Ihnen war ein fantastisches Fußballwochenende oder? ein großartige				
TC	75 sec.	76 sec.	77 sec.	78 sec.	79 sec.
Am					
TC	80 sec.	81 sec.	82 sec.	83 sec.	84 sec.
Am					
Pmf	r Auswärtssieg und damit vorzeitig Meister herzlichen Glückwunsch Fenerbahce Istanbul				
TC	85 sec.	86 sec.	87 sec.	88 sec.	89 sec.
Am					
Pmf	übrigens				
TC	90 sec.	91 sec.	92 sec.	93 sec.	94 sec.
Am					
Pmf	Gejohle, Klats				
TC	95 sec.	96 sec.	97 sec.	98 sec.	99 sec.
Am					
Pmf	die Mannschaft von Christoph Daum ist vorzeitig Meister geworden hier in der Türkei Fenerbahce liegt übrigens drüben auf der asiatischen Seite				

**Abbildung 17: TV Total**

Passage (oder Sequenz): Eingangspassage  
 Film (oder Video): TV Total  
 Dateiname: TV Total.avi  
 Datum: 24.03.2010  
 Time Code: 00 - 513  
 Dauer: 513 sec.  
 Transkription: Stefan Hampl  
 Korrektur: -

TC	0 sec.	1 sec.	2 sec.	3 sec.	4 sec.
Am	Erst einmal möchte ich Ihnen eine Schlagzeile zeigen vom heutigen Tage		Folgendes	Sesamstraße	Gibt's jetzt auch
Pmf				@@	@
TC	5 sec.	6 sec.	7 sec.	8 sec.	9 sec.
Am	in Afghanistan		kein Scherz	es gibt jetzt auch	
Pmf				@@	
TC	10 sec.	11 sec.	12 sec.	13 sec.	14 sec.
Am	die Sesamstraße in Afghanistan ja?	die erste Straße in Afghanistan übrigens ja?		und die äh	
Pmf	@				@
TC	15 sec.	16 sec.	17 sec.	18 sec.	19 sec.
Am	und die habn ja@haha@		und	ja	da sehn die Kinder
Pmf					Bildungsfernsehe
TC	20 sec.	21 sec.	22 sec.	23 sec.	24 sec.
Am	ja da sehn die Kinder dann auch Graf Zahl		ja?	wie=er sagt	eins
Pmf					
TC	25 sec.	26 sec.	27 sec.	28 sec.	29 sec.
Am	zwei	drei	vier	fünf	nachladen ja?
Pmf					@@@

### 5.2.1.1 Gattungsimmanenter Vergleich

Der gattungsimmanente Vergleich bietet sich ideal als erster Einstieg in die Praxis der Videointerpretation an. Dabei wird die Auswahl von Vergleichsmaterial anhand von inhaltlichen Kriterien auf ikonografischer Ebene vorgenommen. Die Fernsehsendung „Istanbul Total“, die hier als Ausgangsmaterial der Videointerpretation dient, lässt sich ikonografisch als Fernsehshow, oder noch spezifischer als Late-Night-Show, klassifizieren. Dazu passend wurden die Harald-Schmidt-Show sowie die Sendung „TV Total“ als weitere Vertreter dieses Sendungstyps ausgewählt.

Nachdem die Gattungszugehörigkeit als ikonografisches Auswahlkriterium für Vergleichsmaterial gedient hat, ist in weiterer Folge die empirische Überprüfung dieser Zuordnung erforderlich. Aus forschungsökonomischen und didaktischen Überlegungen empfiehlt es sich, die methodische Kontrolle auf der Ebene der planimetrischen Komposition und der Montage zu beginnen, da so das Fallmaterial als visuelles *Ganzes* (Deleuze, 1996) besser verstanden werden kann.

Bei einer ersten Betrachtung der Videotranskripte (Abbildung 12) wird hinsichtlich der Montage der Eingangspassage der Harald-Schmidt-Show (Abbildung 13) deutlich, dass diese mit einer einzigen Einstellung auskommt. Darin ist der Moderator im Bild zu sehen. Zur leichteren Verständigung wird diese Einstellung aufgrund ihrer konstitutionellen Bedeutung für die Sequenz als *Hauptsequenz* (Bohnsack, 2009, S. 196) bzw. *Haupteinstellung* bezeichnet.

Empirisch charakterisieren sich Sequenzen und Einstellungen auf formaler Ebene zuallererst durch planimetrische Gestaltungselemente, die im Sinne einer *Übergegensätzlichkeit* (Bohnsack, 2009, S. 36; Imdahl, 1994, S. 312; Ruck & Slunecko, 2008) das visuelle Geschehen bestimmen. Laut den genannten Autoren bezeichnen Übergegensätzlichkeiten besondere Spannungsverhältnisse, die typisch für visuelle Medien sind und letztlich den Schlüssel zu deren empirischem Verständnis darstellen: Charakteristisch ist die *gleichzeitige* Erzeugung und Lösung von Spannungen, durch die sich die Kernaussage des (bewegten) Bildes bestimmt.

In Abbildung 13 resultiert die planimetrische Spannung der Haupteinstellung der Harald-Schmidt-Show (unter Einbeziehung vorikonografischen Wissens) aus der horizontalen, dunklen Schichtung des Hintergrundes und der dazu vertikal aufgerichteten, hell akzentuierten Person im Vordergrund. Die Person im Bild (ikonografisch: der Moderator) durchbricht einerseits die horizontale Grundstruktur dunkler Farbtöne (Blau-Schwarz-Rot) durch ihre aufrechte Gestalt, ihre hellen

Körpermerkmale (Kopf, Hemd, Hände) sowie ihre Arm- und Rumpfbewegungen (gegenüber dem unbewegten Hintergrund). Andererseits nimmt die Figur des Moderators (im Sinne einer Lösung) auch zentrale Umgebungsaspekte in sich auf: dunkle Farbtöne durch das schwarze Jackett, die Schichtlinien des Hintergrunds durch die horizontalen Kontrastlinien von Kinn- und Handansatz sowie die Brille, die Unbewegtheit des Hintergrundes durch minimale Rumpfbewegungen und Symmetrie. Planimetrisch wird die Spannung in Bild und Sequenz damit von der Gestalt des Moderators sowohl induziert als auch reduziert – und somit in Form einer Übergegensätzlichkeit fokussiert.

Stellt man die Eingangspassagen von „Istanbul Total“ und der Harald-Schmidt-Show einander gegenüber, so dokumentiert sich auf der Ebene der häufigsten Einstellung (Abbildung 12: Sekunden 0, 16–23, 49–57, 65–67, 69–81, 88–99) eine *homologe*, d. h. vergleichbare, planimetrische Fokussierung, in der der Moderator bildbestimmend ist. Homologien bestehen auch auf der Ebene der *Perspektive*: Diese ist in beiden Fällen flach (ohne Fluchtlinien, d. h. räumliche Tiefenachsen). Es handelt sich um eine *Parallelperspektive*, die sich aus einer zu Moderator und Umgebung rechtwinkligen Kameraposition ergibt. Aufgrund kleiner perspektivischer Verzerrungen an den Bildrändern (Harald-Schmidt-Show: gelber Vorhang links im Bild, „Istanbul Total“: gekrümmtes Balkongitter unten) lässt sich in beiden Fällen eine Aufnahmeposition im oberen Bilddrittel (etwa auf Kopfhöhe) sowie ein leicht abwärtsgeneigter Blickwinkel rekonstruieren. Schließlich bestehen auch auf der Ebene der szenischen Choreografie deutliche Übereinstimmungen: Beide Moderatoren sind körperlich den Bildbetrachter/innen weitgehend zugewandt und deuten mit ihren Händen in verschiedene Richtungen. Wie in Abbildung 17 zu sehen, dokumentiert sich die – auf allen drei Ebenen der formalen Komposition des Bildes – festgehaltene Homologie der Haupteinstellungen von „Istanbul Total“ und der Harald-Schmidt-Show auch in der Haupteinstellung der Vergleichssendung „TV Total“.

Nachdem die Gattungszugehörigkeit zu Beginn der Untersuchung als rein ikonografische Suchhilfe gedient hatte, erlaubte die darauf aufbauende komparative Analyse eine sendungsübergreifende Charakterisierung von Late-Night-Shows. Auf empirischem Wege konnte so die zentrale Stellung der Moderator-Hauptsequenz in Eingangssequenzen der Gattung Late-Night-Show herausgearbeitet und strukturell charakterisiert werden. Auf der Ebene der Montage treten nun zwei grundlegende Gestaltungsaspekte hervor, die für die weitere Betrachtung von Interesse sind: erstens,

dass sich die Sendung „Istanbul Total“ gegenüber Vergleichssequenzen derselben Gattung eingangs durch einen weit häufigeren Einstellungswechsel auszeichnet, sowie zweitens, dass die Frequenz des Einstellungswechsels nach den ersten sechzig Sekunden deutlich abnimmt und sich formal den Vergleichsshows annähert (Abbildung 16). Im Vorgriff auf die abschließende Gesamtinterpretation könnte der zweite Aspekt für die Wirksamkeit der zuvor als zentral identifizierten Bildstellung des Moderators in der Gattung der Late-Night-Shows gewertet werden, zu der die Sendung – nach anfänglicher Abweichung – wieder hintendiert. Die Frage, worin diese Abweichung besteht und was sich in ihr dokumentiert, kann durch den gattungsübergreifenden Vergleich geklärt werden.

### **5.2.1.2 Gattungsübergreifender Vergleich**

Der gattungsübergreifende Vergleich von Videosequenzen setzt direkt an den planimetrischen Strukturen von Einstellungen *und* Einstellungswechseln an. Die Erweiterung des Konzepts der planimetrischen Komposition auf bewegte Bilder wird dabei durch die Lektüre von Max Imdahl (1994, 1996a) nahegelegt. Er versteht nämlich die planimetrische Komposition als die primordiale Totalstruktur des Bildes, die sich durch eine spezifische Konstellation von Formen, Farben, Richtungen, Kontrasten, Spannungen, Lösungen und letztlich Übergegensätzlichkeiten ergibt. Wird anstelle des Standbilds das Bewegtbild betrachtet, so lassen sich im Sinne Imdahls sowohl einzelne Einstellungswechsel als auch Muster von Einstellungswechseln interpretieren. Im Rahmen der Gattungs- und Montageanalyse interessieren aus planimetrischer Sicht etwa Fragen des Kontrasts, der Frequenz bzw. des Rhythmus von Einstellungswechseln. Im konkreten forschungspraktischen Vorgehen wird dabei nach Vergleichsmaterial gesucht, das dem Ausgangsmaterial auf möglichst vielen Ebenen planimetrisch ähnlich ist, aber nicht notwendig derselben Gattung entspricht. Auf diese Weise können Gattungsgrenzen überschritten und gegebenenfalls neue Gattungstypen festgelegt werden.

Zu Beginn der gattungsübergreifenden Interpretation können wir uns auf ein Ergebnis des vorangegangenen gattungsimmanenten Vergleichs beziehen: Die Hauptsequenz von „Istanbul Total“ wird – für Late-Night-Shows untypisch – auffällig von anderen Sequenzen unterbrochen. Dies betrifft sowohl die Frequenz als auch den planimetrischen Bildinhalt der in den Einstellungswechseln aufeinandertreffenden Einstellungen.

Sequenzen, die eine Hauptsequenz unterbrechen, können laut Bohnsack (2009, S. 162) in *Untersequenzen* und *eingelagerte Sequenzen* differenziert werden, wobei Untersequenzen im Gegensatz zu eingelagerten Sequenzen als Einstellungsvariationen innerhalb der Hauptsequenz aufgefasst werden. Die Interpretation der Einstellungen erfolgt im nachfolgenden Kapitel im Rahmen einer *komparativen Einstellungsvariation*. Für die Darstellung der Interpretation einzelner *Fotogramme*, also Standbilder, sei an dieser Stelle auf Bohnsack (2009, S. 202) verwiesen. Die Eingangssequenz von „Istanbul Total“ besteht aus einer Untersequenz, in der der Moderator in seitlicher Totale am Balkon stehend zu sehen ist (Abbildung 12 ff.: 1–11, 29–33, 39–43 sec.). Weiters lassen sich zwei Arten von eingelagerten Einstellungen unterscheiden: In der einen ist Publikum zu sehen – was, wie sich empirisch rekonstruieren ließe, für Late-Night-Shows *gattungstypisch* ist; in den anderen sind (entfernte) Bauwerke und Fahrzeuge erkennbar – was für Late-Night-Shows *gattungsuntypisch* ist. Auf der Ebene des Verhältnisses von Haupt- bzw. Untersequenzen zu eingelagerten Sequenzen sind somit die ersten sechzig Sekunden von „Istanbul Total“ einem Reisevideo weit ähnlicher als einer Late-Night-Show.

Im Rahmen der komparativen Analyse der Sendungen „Istanbul Total“ und „Reiselust Istanbul“ (Abbildung 15) können Homologien und Kontraste empirisch herausgearbeitet werden. In den Haupt- und Untersequenzen beider Sendungen ist jeweils eine Person im Bild zu sehen, die sich vor einem entfernten (Gebäude-) Hintergrund befindet. Die eingelagerten Sequenzen unterscheiden sich inhaltlich: Im Gegensatz zu den entfernten, diffusen Objekten in „Istanbul Total“ (z. B. Abbildung 14: 35–38, 44–48, 58–64 sec.) zeigt „Reiselust Istanbul“ Menschenmengen (Abbildung 15: 7–11 sec.) und öffentliche Transportmittel (Abbildung 15: 12–18 sec.), die klar umrissen sind und den Betrachter/inne/n recht nahe kommen. Ikonologisch dokumentiert sich in der Sendung „Istanbul Total“ eine Distanzierung gegenüber der Umgebung. Diese Distanzierungsleistung kann anhand eines Vergleichs der Haupteinstellungen (z. B. Abbildung 14: 49–57 sec. und Abbildung 15: 0–6 sec.) exemplarisch noch weiter ausgeführt werden: Sowohl in „Istanbul Total“ als auch in „Reiselust Istanbul“ befinden sich die Sprecher auf Balkonen in erhöhter Position. Während der Balkon bei „Reiselust Istanbul“ *in* der Straße der Stadt hängt, ist er bei „Istanbul Total“ *über* der Stadt und über dem Meer angebracht. Der Moderator von „Istanbul Total“ ist sowohl planimetrisch als auch perspektivisch (im Rahmen der Untersequenzen) fokussiert. Demgegenüber tritt der Sprecher in „Reiselust Istanbul“ planimetrisch zwar durch Farbe und Form hervor.

Jedoch ist er seitens der abbildenden Bildproduzent/inn/en deutlich am linken Bildrand angeordnet, sodass rechts das Blickfeld frei bleibt und perspektivisch in die Stadt erweitert wird. Im Gegensatz zu Late-Night-Shows unterliegen Reisevideos folglich keiner „Monostrukturierung durch den Showmaster“ bzw. „Hyperzentrierung auf dessen Person“ (Bohnsack, 2009, S. 240). Den Menschen und ihren Umgebungen wird in Reisevideos mehr Raum gegeben, um ihre eigenen Relevanzstrukturen zu entfalten. Dies demonstrieren auch andere im Handel erhältliche Videobeispiele: z. B. „Städtereisen – Istanbul“, „Istanbul On Tour“ oder „Crossing the Bridge: The Sound of Istanbul“, ein Dokumentarfilm von Fatih Akin.

In knapper Form, aber durch die Interpretationen sowie die vorliegenden Videotranskripte (Abbildung 12 ff.) gestützt, kann hier der ikonologische bzw. ikonische Schluss (Bohnsack, 2007, S. 158 ff.) gezogen werden, dass die Sendung „Istanbul Total“ zu Beginn deutliche Anleihen beim Genre Reisevideos nimmt. Erklärbar wird dieses Vorgehen durch den ungewöhnlichen Sendestandort in Istanbul, der es erfordert, den Zusehenden die neue (Studio-)Umgebung als Teil der Show näher zu erklären. Gleichzeitig unterliegen Late-Night-Shows wie „Istanbul Total“ gattungs- und produktionsbedingt einer starken Monostrukturierung durch den Moderator. Ein gattungsimmanenter und ein gattungsübergreifender Vergleich fördern die Kompromissbildung zutage, die sich im Verlauf der Sendung dokumentiert: Nachdem in der ersten Sendeminute gattungsuntypisch einige Gebäude und andere Umgebungsobjekte hervorgehoben werden, kehrt die Sendung sukzessive zu jener prominenten Fokussierung des Moderators im Bild zurück, die für Late-Night-Shows üblich ist. Schon von vornherein sind der Aufnahmestandort der Kameras sowie die Position des Moderators festgelegt: Das Fernsehstudio bzw. dessen Balkon ermöglichen einen friktionsfreien Wechsel vom Modus des Reisevideos zurück zum Modus der Late-Night-Show. Geradezu kompensatorisch verstärkt sich in diesem Zuge gegenüber Reisevideos das Prinzip der Distanzierung<sup>40</sup>.

Nachdem für das Medium Video eine Form der Kompositionsvariation vorgeführt werden konnte, die sich im Sinne einer „Relationierung der Relationen“ (Bohnsack, 2009, S. 165) auf den Vergleich der formalen Strukturen der Einstellungen *und* des Einstellungswechsels bezog, wird nun mit der *Einstellungsvariation* ein Verfahren

---

<sup>40</sup> Derselbe Wirkmechanismus dokumentiert sich neben „Istanbul Total“ (2004) bis heute auch in anderen Song-Contest-Specials der Sendung „TV-Total“. Das jüngste Beispiel, „Unser Star für Oslo“ (2010), wirkt wie eine norwegische Variante von „Istanbul Total“.

vorgeführt, dass zur Charakterisierung der planimetrischen „Grundstimmung“ von Videos dienlich ist. Es geht dabei um das Herausarbeiten jener Bildstrukturen, die das Videomaterial sequenzübergreifend in seiner Totalstruktur prägen.

## **5.2.2 Einstellungsvariation**

Zur Entwicklung eines tieferen Verständnisses der Gesamtkomposition von Bildern schlägt Imdahl (1994) deren experimentelle Veränderung durch die Forscher/innen vor. Das von Bohnsack (2001a, S. 77) als „Kompositionsvariation“ bezeichnete Verfahren lässt sich prinzipiell auch auf die Film- und Videointerpretation übertragen. Im Gegensatz zur Bildinterpretation dokumentieren sich Kompositionsleistungen eines Films oder Videos nicht nur auf der Ebene des Einzelbildes bzw. Fotogramms, sondern auch auf der von Einstellung und Montage. Im Sinne eines differenzierten Verständnisses der Kompositionsvariation soll deshalb das Verhältnis zwischen Fotogramm und Einstellung näher spezifiziert werden.

### **5.2.2.1 Einstellung und Fotogramm**

Unter dem Begriff der Einstellung versteht Bohnsack (2009, S. 160) die „Art der Gestaltung des Bildausschnitts, wie er durch die Kamera geleistet wird“. Von daher ist die Einstellung in besonderem Maße durch die Leistungen der abbildenden Bildproduzent/inn/en – d. h. der Personen hinter der Kamera – definiert. Dieser Umstand hat Konsequenzen für die Interpretation und Kompositionsvariation von Fotogrammen, die zur Rekonstruktion der Einstellung ausgewählt werden. Das durch das Medium Film bzw. Video aufgespannte Verhältnis von Einstellungen und Fotogrammen markiert eine ontologische Differenz zwischen Bild- und Videointerpretation. Während sich in Fotografien oder anderen (Stand-)Bildern die Leistungen von abbildenden *und* abgebildeten Bildproduzent/inn/en in prinzipiell gleichberechtigter Weise dokumentieren, gibt die Einstellung in asymmetrischer Weise den Orientierungsrahmen der abbildenden Bildproduzent/inn/en für die Fotogramme vor, innerhalb dessen sich die Leistungen der abgebildeten Bildproduzent/inn/en – etwa im Bereich der Mimik und Gestik<sup>41</sup> – artikulieren können.

Unter Berücksichtigung der angeführten Einschränkung ist es zusammenfassend möglich, die Kompositionsvariation von Einstellungen mittels repräsentativer Fotogramme durchzuführen. Gegenstand der Variation sind dabei die Leistungen der

---

<sup>41</sup> Zur Interpretation siehe Kap. 5.2.3 bzgl. „Gesten- bzw. Gebärdenvariation“.

abbildenden Bildproduzent/inn/en, die sich auf inhaltlicher und formaler Bildebene dokumentieren. In einem ersten Schritt kann die Einstellungsvariation fallintern in der Gegenüberstellung ausgewählter Einstellungen des Film- bzw. Videomaterials bestehen. Die für diese Einstellungen als repräsentativ ausgewählten Fotogramme wiederum können als Ausgangspunkte für die Suche nach Vergleichsbildern außerhalb des Fallmaterials dienen. Das genaue Prozedere wird in der Folge dargestellt. Zuvor sei noch ein Exkurs über das Internet als technisches Hilfsmittel bei der Bildersuche eingefügt.

#### **5.2.2.2 Exkurs: Text- und bildgebundene Suche im Internet**

Hinsichtlich der technischen Möglichkeit, mithilfe von Internetsuchmaschinen zu Vergleichsbildern für die komparative Analyse zu kommen, lassen sich prinzipiell zwei Arten von Bildersuche im Internet unterscheiden: Zum einen gibt es die klassische textgebundene Suche nach Bildern. Mit ihr können Bilder recherchiert werden, die zuvor beschlagwortet wurden. Gibt man als Nutzer/in etwa das Wort „Baum“ ins Suchfeld ein, werden verschiedene Bilder von Bäumen angezeigt. Für die Bild- und Videointerpretation besteht der Nutzen der textgebundenen Bildersuche in der Möglichkeit mittels (vor-)ikonografisch rekonstruierter Kriterien nach anderen Bildern zu suchen, die als empirische Vergleichshorizonte genutzt werden können. Die Grenzen der textgebundenen Bildersuche liegen in ihrer Orientierung am wiedererkennenden Sehen. Es kann nur nach sprachlichen Begriffen gesucht werden, die zuvor aus dem eigenen Material oder gedankenexperimentell entwickelt wurden. Der Umweg über die Textebene schränkt bei dieser Suchmethode die Möglichkeit zur direkten komparativen Entschlüsselung der Formalstruktur des Ausgangsbildes ein. In Abbildung 18 sind neben dem Ausgangsbild in „Parallelprojektion“ (Müller, 2012, S. 149) auf der linken Seite zwei Bilder aus dem Internet zu sehen, die nach Eingabe der Kriterien „Mann, Meer, Balkon“ angezeigt wurden. Dabei werden die Spielräume der textgebundenen Suche sichtbar. Im mittleren Bild etwa ist das Meer gar nicht *im* Bild zu sehen, wenngleich es vorstellbar wäre, dass der Mann vom Balkon aus *auf* das Meer blickt.



**Abbildung 18: Beispiel des Ergebnisses einer textgebundenen Bildersuche.** Parallelprojektion des Ausgangsbilds auf der linken Seite mit zwei weiteren Bildern, die mithilfe der vorikonografischen Suchkriterien „Mann, Meer und Balkon“ im Internet gefunden wurden. Quelle der Vergleichsbilder: [images.google.com](https://images.google.com)

Einen neueren Ansatz der Bildersuche im Internet stellt die sogenannte Ähnliche-Bildersuche bzw. similar image search (vgl. Abbildung 19) dar. Dieses Verfahren könnte auch als *bildgebundene Suche*<sup>42</sup> bezeichnet werden, da sich dabei direkt Bilder *durch* Bilder suchen lassen. Um die Suche nach Vergleichsbildern zu starten, ist es nicht erforderlich, einen sprachlichen Suchbegriff einzugeben – vielmehr dient das Ausgangsbild selbst als Suchkriterium. Dieses wird von der Suchmaschine zuerst mittels technischer Algorithmen auf formale Muster hin (Farben, Formen, Bildaufteilung etc.) untersucht. Auf dieser Basis werden den Nutzer/inne/n dann Vergleichsbilder vorgeschlagen, die formal ähnlich aufgebaut sind. Der Vorteil dieser Suchmethode liegt darin, dass sie ohne Umweg über die Sprache auskommt und sich rein im Medium des Bildes bewegt. Als Nachteil kann die daraus resultierende hohe (vor-)ikonografische Heterogenität der Ergebnisse angesehen werden. Da sich diese Suchmethode aber *im* Medium des Bildes selbst vollzieht, steht sie insgesamt dem sehenden Sehen näher als die textgebundene Suche und ist für die explorative Generierung von Bildkontrasten ergiebiger. In Abbildung 19 sind neben dem Ausgangsbild links zwei weitere Bilder zu sehen, die hinsichtlich ihrer planimetrischen Gesamtkomposition (Formen, Farben, Bildaufteilung; Symmetrie etc.) dem Ausgangsbild nahe kommen. In der Parallelprojektion werden unmittelbar mehrere bestimmende formale Elemente des Ausgangsbildes deutlich: Die auffällige Symmetrie, die Rundheit des Turbans, der Kuppel oder des Kopfes im oberen

<sup>42</sup> Hinsichtlich der Favorisierung des direkten Bildvergleichs ist der methodologische Ansatz der bildgebundenen Suche prinzipiell mit dem Konzept einer „figurativen Hermeneutik“ (Müller, 2012) verwandt. Die Unterschiede der beiden Verfahren liegen in ihren Erkenntniszielen, Analyseinstellungen sowie den damit verbundenen Arten der Bildzusammenstellung. Indem die „figurative Hermeneutik“ vornehmlich auf die Analyse „körpersprachlicher“ Merkmale, „zeichen-symbolische[r] ... Referenzen“ sowie die Gattungsanalyse (Müller, 2012, S. 151) abzielt (was der Analyseinstellung des wiedererkennenden Sehens entspricht), erfolgt die Zusammenstellung kontrastiver Vergleichsbilder auf Basis (vor-)ikonografischer Kriterien. Demgegenüber erfolgen die Bildzusammenstellungen im Rahmen der hier vorgestellten bildgebundenen Suche auf Basis planimetrischer Kriterien (was der Analyseinstellung des sehenden Sehens entspricht). Das vorrangige Erkenntnisziel der bildgebundenen Suche besteht in der kontrastiven Vertiefung der formalen Aspekte der Ikonik des Bildes. In Komplementarität zu Müller (2012) könnte dieser methodologische Ansatz damit ggf. auch als „planimetrische Hermeneutik“ aufgefasst werden.

Bilddrittel, der Komplementärkontrast von Blau-Violett und Gelb-Orange, das starke rechtwinklige Verhältnis von Horizontalität und Vertikalität etc.



**Abbildung 19: Beispiel einer bildgebundenen Suche.** Parallelprojektion des Ausgangsbilds auf der linken Seite mit zwei weiteren Bildern, die aufgrund ihrer ähnlichen planimetrischen Komposition im Internet gefunden wurden. Quelle der Vergleichsbilder: [images.google.com](http://images.google.com)

Nach der Darstellung der forschungspraktischen Möglichkeiten der Bildersuche im Internet seien abschließend aber nicht die *caveats* dieses Zugangs verschwiegen. Zum einen zeigt die Internetbildersuche nicht *alle* möglichen Bilder, sondern nur diejenigen, die irgendjemand irgendwann eingegeben hat<sup>43</sup>. In diesem Zusammenhang ist davon auszugehen, dass ein/e Suchanbieter/in jede Suchanfrage und jedes Bild, das zu ihm/ihr hochgeladen wird, auch speichert. Daher ist der Einsatz der Internetbildersuche für vertrauliches oder urheberrechtlich geschütztes Bildmaterial nicht zu empfehlen. Letztlich steht hinter dem kostenlosen Suchangebot der privaten Internetsuchanbieter/innen auch ein ökonomisches Interesse. Es wird jedoch keine transparente Information darüber zur Verfügung gestellt, inwieweit dies die Ergebnisse von Internetsuchen beeinflusst.

Zusammenfassend kann somit sowohl die text- als auch die bildgebundene Suche von Vergleichsbildern die komparative Analyse forschungspraktisch unterstützen und ökonomisieren. Für Forscher/innen ist es jedoch stets erforderlich, die Grenzen und Gefahren des Einsatzes dieser technischen Möglichkeiten ausreichend zu reflektieren und im Hinblick auf die Auswertung des eigenen Materials zu berücksichtigen.

### 5.2.2.3 Fallbeispiel

Für die folgende Einstellungsvariation wurden insgesamt sieben Fotogramme der zirka dreieinhalbminütigen Eingangspassage der TV-Sendung „Istanbul Total“<sup>44</sup> ausgewählt. Das Auswahlkriterium war die minimale Variation der planimetrischen Grundstruktur des Videomaterials: Planimetrisch gesehen stellen der hellblaue Farbton sowie die

<sup>43</sup> Nach Angaben von [www.bildersuche.org](http://www.bildersuche.org) (Stand: Juli 2011) sind in der Google-Bildersuche bereits über zehn Milliarden Bilder verfügbar, wovon täglich über eine Milliarde abgerufen wird.

<sup>44</sup> Link zum Video: <http://tvtotal.prosieben.de/tvtotal/videos/player/index.html?contentId=14144>

ausgeprägte Horizontalität und Vertikalität der meisten Einstellungen eine wesentliche Fokussierung dieser TV-Sendung dar. Ausgangspunkt der Einstellungsvariation sind nun Einzelbilder, in denen diese Fokussierung in kleinen Veränderungen (im Sinne eines minimalen Kontrasts) zum Ausdruck kommt. Es handelte sich dabei um die Fotogramme der Sekunden 0, 5, 12/44, 24/36 und 79, die in der obersten Zeile von Abbildung 20 angeordnet sind. Direkt darunter sind die zugehörigen Sprechtexte des abgebildeten Bildproduzenten (Am) angeführt<sup>45</sup>. Unterhalb davon sind jeweils planimetrisch ähnliche Bilder zu finden, die nicht dem Fallmaterial entstammen, sondern im Internet mittels bildgebundener Suche recherchiert wurden.<sup>46</sup> Ergänzend zum bekannten Einzeichnen planimetrischer Linien in die Fotogramme hat sich in der Arbeit mit Studierenden das Herausarbeiten planimetrischer Strukturen durch planimetrisch homologe Vergleichsbilder bewährt. Die Diskussion der planimetrisch-ikonischen Grundstruktur des Fallmaterials lässt sich auf diese Weise anschaulich empirisch führen, ohne es dabei zeichnerisch zu verändern<sup>47</sup>.

TC	0 sec.	5 sec.	12 sec./44 sec.	24 sec./36sec.	79 sec.
					
Am	„Ich befinde mich“	„wie man sehen kann“	„Asien und Europa“	„ältesten Kirchen“/ „die andere Seite“	„Fenerbahçe Istanbul“

**Abbildung 20: Einstellungsvariation von „Istanbul Total“**

Die Darstellung der Einstellungsvariation orientiert sich am Layout von Videotranskripten nach dem System MoViQ (Hampl 2008, 2010; Przyborski & Wohlrab-Sahr 2008), was an der Sprecherspur (Am) erkennbar ist. Der Unterschied zu Videotranskripten besteht in den unregelmäßigen Zeitabständen zwischen den Fotogrammen. Bei 12 sec./44 sec. und 24 sec/36 sec. steht ein Fotogramm jeweils für zwei homologe Einstellungen (vgl. Hampl 2010).

<sup>45</sup> Auch anhand der Interpretation der Textfragmente ließe sich übrigens das Pars-pro-Toto-Prinzip demonstrieren.

<sup>46</sup> Im Internet gibt es einige kostenlose Anbieter für die bildgebundene Suche (z. B.: <http://www.google.com/imghp>, <http://www.tineye.com/> etc.). Auf diesen Seiten können Fotogramme hochgeladen und dann ähnliche Bilder abgerufen werden. Vom Hochladen urheberrechtlich geschützten oder vertraulichen Materials sollte Abstand genommen werden, da jedes hochgeladene Bild online gespeichert wird.

<sup>47</sup> In Anlehnung an Dagobert Freys Feldtheorie hat Imdahl (1996a) zur Verdeutlichung der für die Ikonik konstitutiven Planimetrie im Bild das Einzeichnen planimetrischer Linien in das Bild vorgeschlagen. Wenngleich Imdahls ursprüngliche Idee die Verdeutlichung dominanter Bildstrukturen gewesen sein mag, so handelt es sich dabei doch um einen in der Literatur nur vage spezifizierten, konstruktiven Zeichenakt der Forscher/innen, der das angestrebte „sehende Sehen“ einschränken kann.

## Fotogramm 0: „Ich befinde mich“



Abbildung 21: v.l.n.r. Fotogramm 0 sec., Tagesschau, Urlaubsfoto. Quelle: images.google.com

Die Planimetrie des allerersten Videobildes (Fotogramm 0) wird – wie die gesamte restliche Videosequenz – farblich maßgeblich durch eine hellblaue Fläche geprägt, die im konkreten Fall horizontal in verschiedenen Blauabstufungen geschichtet ist (Himmel, Uferlinie, Meer, Balkongitter). Formal hebt sich davon farblich sowie durch ihre Lage und Ausrichtung eine braun-weiß-goldene Gestalt ab, die senkrecht rechts der Bildmitte aufragt (ikonografisch: der Moderator<sup>48</sup> in seiner Verkleidung). Im Sinne einer Übergegensätzlichkeit (Imdahl, 1994) wird das Bild – trotz dieser prinzipiell gegensätzlichen Strebungen – durch horizontale Farbkontraste (Turban bzw. auf Türkisch „Sarık“ und Bauchbinde) in der Gestalt des Moderators integriert. Die Gestalt des Moderators ist somit planimetrisch fokussiert, indem sie einerseits aus dem Ensemble farblich und lagemäßig herausfällt, andererseits durch die spezifischen Akzente ihrer (Ver-)Kleidung jedoch das Bild zusammenhält.

Durch die Gegenüberstellung der Vergleichsbilder in Abbildung 21 tritt die formale Struktur des Ausgangsbildes (Fotogramm 0) profilierter zutage: Sie besteht im farblichen Kontrast von Blau- und komplementären Gelbtönen, der rechts der Mitte (jeweils klar abgrenzbar) eine schmale, vertikale Figur aus dem Bild hervortreten lässt. Wenn Personen im Bild leicht außerhalb der Bildmitte angeordnet sind, eröffnet dies den Blick auf den Hintergrund, welcher dadurch thematisiert und im Modus der Gleichzeitigkeit ins Verhältnis zur Person im Vordergrund gestellt wird. Wie der Bildvergleich zeigt, ist dieser Effekt sowohl in professionellen Bildern (hier aus der Nachrichtensendung Tagesschau<sup>49</sup>) als auch in amateurhaften Urlaubsfotos realisiert.

---

<sup>48</sup> Erklärung: Es handelt sich um den bekannten Entertainer Stefan Raab, der seit über zehn Jahren die Sendung „TV Total“ sowie auch deren Ableger (z. B. „Istanbul Total“) moderiert.

<sup>49</sup> Im Falle des Bildes E 0 der ARD-Tagesschau handelt es sich ikonografisch um einen stilisierten Hintergrund, also ein Bild im Bild (vgl. Bohnsack 2003, S. 244).

## Fotogramm 5: „Wie man sehen kann“



Abbildung 22: v.l.n.r. Fotogramm 5 sec., Menschen vor/auf Wohnhaus, Hotelanlage. Quelle: images.google.com

Planimetrisch ist auch Fotogramm 5 durch seine hellblaue Farbe bestimmt, wobei der Farbton insgesamt etwas violetter ist als im Fotogramm 0. Im direkten Vergleich ist Fotogramm 5 weniger klar strukturiert und stattdessen eher durch kleingliedrige Elemente bestimmt. Wie schon zuvor ist das 90°-Verhältnis von Vertikale und Horizontale ein bestimmender Bildaspekt, der sich in den scharfen Kanten der hellblau-weißen rechten oberen Bildhälfte dokumentiert. Damit besteht ein Gegensatz zwischen der diffusen Kleingliedrigkeit und der rechtwinkligen Klarheit des Bildes.

Die Person des Moderators integriert diese beiden Aspekte im Sinne einer Übergegensätzlichkeit. Der Moderator fügt sich einerseits durch seine eigene Kleinheit in die detailreiche Umgebung ein, ist gleichzeitig aber durch einen weißen Farbakzent (Hemd und Turban) markiert. Des Weiteren befindet er sich direkt unterhalb der Stelle, wo horizontale und vertikale Linien sich treffen. Damit ist die Person des Moderators zwar erneut planimetrisch fokussiert, geht jedoch als Einzelperson in der Umgebung unter. Dieser Effekt tritt auch in den planimetrisch homologen Vergleichsbildern von Abbildung 22 zutage: Die Grenzen zwischen Personen und Gebäuden verschwimmen. Hierin zeigt sich ein deutlicher Kontrast der Einstellung von Fotogramm 5 zu jener von Fotogramm 0, welche durch eine klare Farb- und Formgebung bestimmt ist.

## Fotogramm 12 bzw. 44: „Asien und Europa“



Abbildung 23: v.l.n.r. Fotogramm 12 bzw. 44 sec., Hügelketten im Dunst, Windsurfer am Meer. Quelle: images.google.com

Die beiden Fotogramme 12 und 44 aus dem Videomaterial sind zueinander planimetrisch homolog und werden deshalb gemeinsam betrachtet. Im Wesentlichen ist das Bild erneut stark von hellblauen Farbtönen geprägt, die sich in horizontaler Schichtung übereinander befinden (Himmel, Ufer, Schiff, Meer). Die Farben kontrastieren kaum miteinander. Im rechten Bilddrittel ist eine breite, mittelblaue Linie (Brückenpfeiler der Bosphorusbrücke) erkennbar. Diese Linie stellt, zusammen mit dem horizontalen Ufer, eine Homologie zu den vorhergehenden Bildern dar. Insgesamt gesehen dominiert hier jedoch die horizontale Schichtung gegenüber der Vertikalität, die in den Fotogrammen 0 und 5 stärker ausgeprägt war. Planimetrisch fokussiert ist damit das Verhältnis „oben – unten“ in verschiedenen Blautönen. Da kein Moderator zu sehen ist, wird das Bild durch diese Fokussierung vollständig bestimmt und zusammengehalten. Die in Abbildung 23 neben Fotogramm 12 angeordneten Vergleichsbilder illustrieren den planimetrisch-ikonischen Befund: Himmel, Berge und Meer liegen schichtweise und in atmosphärisches Blau getüncht übereinander, worin sich eine große Distanz zu den betrachteten Objekten dokumentiert.

Fotogramme 24 bzw. 36: „ältesten Kirchen“ bzw. „die andere Seite“



**Abbildung 24: v.l.n.r. Fotogramm 24 sec., Fotogramm 36 sec., Stephansdom in Wien. Quelle: images.google.com**

Im vorliegenden Fall werden zwei Fotogramme betrachtet, die zueinander hinsichtlich ihres hellblauen Farbtons homolog sind und die beide eine dunkle vertikale (Turm-) Spitze in der rechten Bildhälfte aufweisen. Es bestehen jedoch auch Kontraste zwischen den Bildern. Wie die Vergleichsbilder zeigen, unterscheidet sich Fotogramm 24 planimetrisch von Fotogramm 36 hinsichtlich seiner stärkeren Symmetrie und „Rundheit“ (Kuppeldach). Diese Spannung stellt damit die Fokussierung der beiden Bilder dar. Wie das Vergleichsbild in Abbildung 24 verdeutlicht, dominiert jedoch in beiden Fotogrammen letztlich die Turmspitze gegenüber einem niedrigeren Gebäudekorpus, was insgesamt an christliche Kirchenbauten erinnert.

## Fotogramm 79: „Fenerbahçe Istanbul“



Abbildung 25: v.l.n.r. Fotogramm 79 sec., Taj Mahal in Indien, Jesusdarstellung von Giotto. Quelle: images.google.com

Das Fotogramm 79 stellt eine Homologie zu Fotogramm 0 dar. Der Unterschied zwischen beiden besteht jedoch darin, dass die Gestalt des Moderators diesmal in das Bildzentrum gerückt ist, wodurch das Bild insgesamt symmetrischer ausfällt. Mit beidseitig waagrecht auf Hüfthöhe ausgestreckten Armen integriert der Moderator vollständig die bereits bei Fotogramm 0 herausgearbeiteten Gegensätze zwischen Horizontale und Vertikale sowie zwischen Hellblau und Weiß.

Vor dem Hintergrund der vorhergehenden Bilder – und den dabei herausgearbeiteten gemeinsamen Bildstrukturen (kontrastreich – kontrastarm, horizontal – vertikal, oben – unten, rund – spitz, rechtslastig – symmetrisch) – zeichnet sich das letzte Fotogramm insbesondere durch die Integration der bisher herausgearbeiteten Bildaspekte einerseits sowie die zusätzliche Betonung von Kontrast und Symmetrie andererseits aus.

Damit handelt es sich in Fotogramm 79 um eine *Metafokussierung*, die als Ausgangspunkt für die komparative Analyse auf der Ebene der Einstellungen („simultane Performanz“) sowie auch auf der Ebene der Montage („sequenzielle Performanz“) herangezogen werden kann. Der Stellenwert der Symmetrie für die Bildgestaltung lässt sich anhand der planimetrisch homologen Vergleichsbilder in Abbildung 25 demonstrieren. In anderen Bildern mit dieser spezifischen Symmetrie dokumentiert sich stets die Metaphorik des Majestätischen und Außergewöhnlichen. Wie die Beispiele des palastartigen indischen Königsgrabes Taj Mahal sowie einer frühmittelalterlichen Jesusabbildung<sup>50</sup> zeigen, trägt zu diesem Eindruck übrigens auch die spezielle Kopfbedeckung des Moderators bei. Diese weist formale Ähnlichkeiten sowohl zum Kuppeldach des Taj Mahal als auch zum Heiligenschein der Jesusdarstellung auf.

<sup>50</sup> Giotto (1304): Das Jüngste Gericht (Ausschnitt)

In Hinblick auf die Gesamtinterpretation des Videomaterials von „Istanbul Total“ (vgl. Bohnsack 2009; Hampl 2010) zeigt der ikonische Vergleich, dass der Dokumentsinn der Fotogramme pars pro toto die wesentlichen Aspekte des Videomaterials enthält. Auch die Rekonstruktion der Montage (Bohnsack 2009; Hampl 2010), die laut Stanley Kubrick das Alleinstellungsmerkmal von Filmen darstellt (Bohnsack 2009), liefert den empirischen Befund, dass ein majestätischer Moderator das Bildgeschehen von „Istanbul Total“ dirigiert und sich damit zugleich von der diffusen türkischen Umgebung distanziert, in der er sich befindet.

Wie die Einstellungsvariation verdeutlicht, ist der Dokumentsinn des Videos pars pro toto in allen Fotogrammen des Videos angelegt, jedoch in unterschiedlicher Fokussierung ausgeprägt. Nachdem bei diesem Interpretationsschritt vor allem die Leistungen der abbildenden Bildproduzent/inn/en im Vordergrund standen, soll in weiterer Folge nun eine Technik zur komparativen Rekonstruktion der Leistungen der abgebildeten Bildproduzent/inn/en vorgestellt werden. Wie bereits erwähnt, sind Filme und Videos in höherem Maße als Bilder durch die Leistungen der abbildenden Bildproduzent/inn/en bestimmt.

Innerhalb des durch die abbildenden Bildproduzent/inn/en vorgegebenen Rahmens besteht für die abgebildeten Bildproduzent/inn/en in der Bilddimension lediglich auf der Ebene ihrer Körperhaltung (Gebärden) und Gesten ein gewisser Gestaltungsspielraum. Mittels des Instruments der Gesten- bzw. Gebärdenvariation soll exemplarisch die Interpretation einer besonders auffälligen Gebärde demonstriert werden.

### **5.2.3 Exkurs: Gesten- bzw. Gebärdenvariation**

Aus methodologischer Sicht fällt die Gesten- bzw. Gebärdenvariation ein wenig aus der Gesamtlage dieser Arbeit heraus, da sich der analytische Fokus dabei sowohl vom sehenden zum wiedererkennenden Sehen als auch von den abbildenden zu die abgebildeten Bildproduzent/inn/en verlagert. Hinter diesem Exkurs stehen aber pragmatische Überlegungen. Zum einen lässt sich an dieser Stelle gut demonstrieren, wie sich das Grundprinzip der Kompositionsvariation auch auf andere empirische Instrumente der Bild- bzw. Videointerpretation übertragen lässt. Zum anderen stellt die Gesten- bzw. Gebärdenvariation ein Verfahren dar, das neben Erweiterungsmöglichkeiten für die dokumentarische Methode auch Anknüpfungspunkte an andere

methodologische Ansätze, etwa den der „figurativen Hermeneutik“ (Müller, 2012) im Rahmen der visuellen Wissenssoziologie (Breckner, 2010; Raab, 2008)“, bietet.

Die Rekonstruktion von Gesten und Gebärden<sup>51</sup> ist als Bestandteil der Videointerpretation sowohl für die Untersuchung von Videomaterial aus Massenmedien als auch für die Betrachtung von Videomaterial aus eigenen Erhebungen relevant. Die Verwendung von Videotranskripten nach dem System MoViQ stellt in diesem Zusammenhang eine große Arbeitserleichterung dar, sollen Bewegungsabläufe in ihrer Genese Schritt für Schritt festgehalten und interpretiert werden<sup>52</sup>. Wie bereits erwähnt, sind Auffälligkeiten und Diskontinuitäten in der Regel Kennzeichen von Fokussierungen im Material (Bohnsack, 2009, S. 215). Die in der Folge betrachtete Geste erstreckt sich über das Intervall von Sekunde 131 bis Sekunde 136 (Abbildung 26) und wird mit planimetrisch homologen Fotos aus dem Internet (Abbildung 27) kontrastiert, um den Dokumentsinn der Geste auf den Punkt zu bringen.

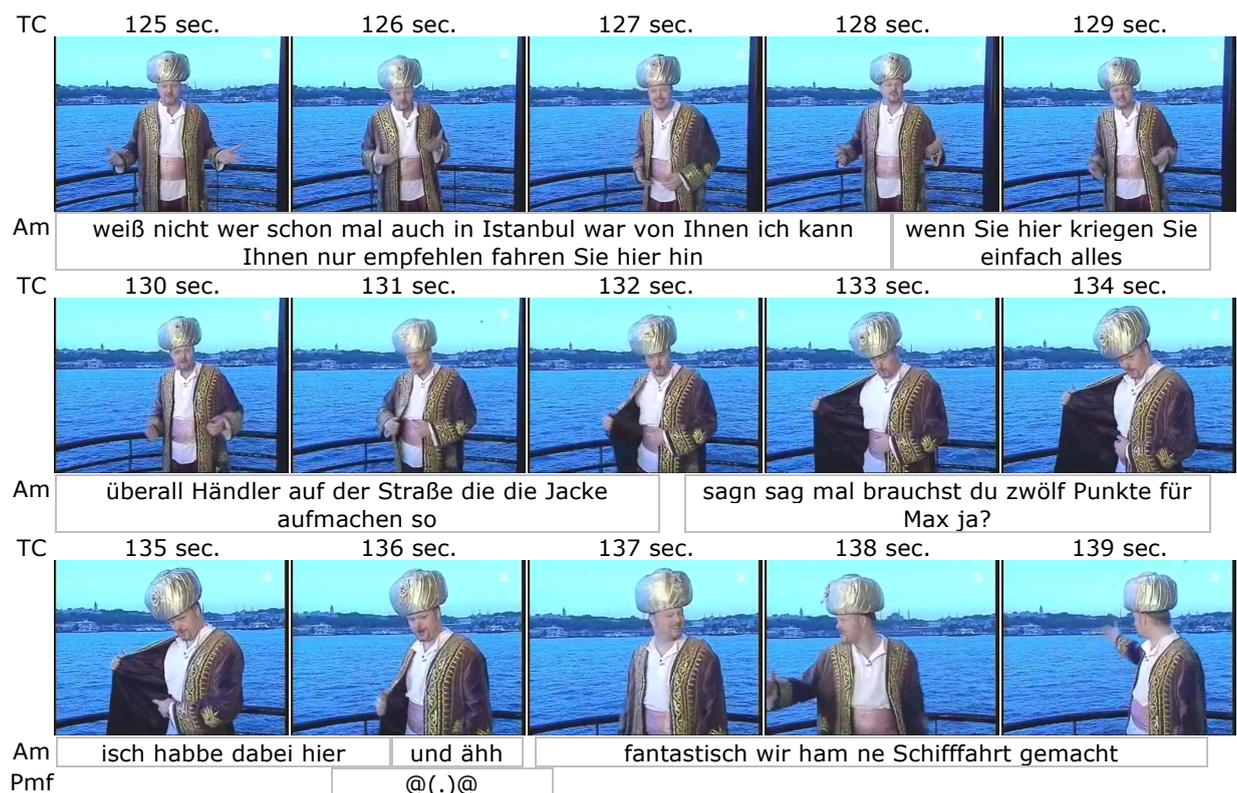


Abbildung 26: Sequenz „Jacke aufmachen“

<sup>51</sup> Zur Differenzierung der Begriffe siehe Kap. 3.4.1, Fußnote 31.

<sup>52</sup> Medienhistorisch betrachtet besteht die Möglichkeit des filmischen Festhaltens von Bewegungsabläufen seit Ende des 19. Jahrhunderts. Unter dem Titel „Animal Locomotion“ publizierte der Fotograf Eadweard Muybridge 1877 u. a. seine berühmten Serienfotografien (siehe etwa Abbildung 2).



Abbildung 27: Vergleichsbilder „Jacke aufmachen“

Vorikonografisch betrachtet öffnet der Mann in Abbildung 26 von Sekunde 131 bis Sekunde 134 seinen Mantel. In Sekunde 135 streckt er seinen Zeigefinger aus und richtet ihn auf die Innenseite des Mantels. Auf planimetrischer Ebene verändert sich durch den Vorgang des Mantelöffnens die symmetrische Grundkonstitution der Einstellung: Im Bild entsteht eine dunkle Fläche, die sich im Kontext eines starken Schwarz-weiß-Kontrasts wie ein Keil zwischen den abgebildeten Bildproduzenten und dessen Umgebung schiebt. Planimetrisch gesehen erhöht sich durch den Kontrastanstieg die Fokussierung der Person im Bildvordergrund; nicht nur wird die Person selbst kontrastreicher und schärfer – auch gegenüber dem Bildhintergrund gewinnt sie zusätzlich an Profil. Die planimetrische Einstellungsvariation, bei der das Gesamtbild im Fokus steht, lässt sich so in eine Gesten- bzw. Gebärdenvariation überleiten. Zusätzlich kann es, wie im vorliegenden Fall geschehen, hilfreich sein, vorikonografische Kriterien bei der Auswahl der Vergleichsbilder mit einzubeziehen. Für die komparative Analyse wurden somit Vergleichsbilder gesucht, die der Sequenz der Geste (Abbildung 26, 131–134 sec.) sowohl planimetrisch als auch vorikonografisch entsprachen.

Die Vorzüge der empirischen Gesten- bzw. Gebärdenvariation werden aus Abbildung 27 ersichtlich. In der Gegenüberstellung der Bilder treten uns schon bei flüchtiger Betrachtung – im Sinne eines Pars pro Toto – jene Bedeutungshorizonte entgegen, die für die Konstitution der Ausgangsgeste (Abbildung 26) relevant sind: Keines der Vergleichsbilder durchdringt die Ausgangsgeste vollständig hinsichtlich ihrer Sinnstruktur. Jedoch ermöglicht es der Gesamteindruck aller Bilder, mit verschiedenen „Brillen“ auf die Ausgangsgeste zu blicken.

Vorikonografisch lassen sich unmittelbar folgende Bestimmungsmerkmale der Geste identifizieren: Erstens ist in der Sequenz eine der Kamera zugewandte Person mit Mantel abgebildet. Diese öffnet während einer Bewegung von sieben Sekunden (131–137) mit der rechten Hand ihren Mantel und schließt ihn dann wieder. Im Zuge dessen

dreht die Person ihren Körper den Bildbetrachter/inne/n gegenüber aus einer Frontalposition in eine Dreiviertelpose. Dabei neigt sie den Kopf leicht abwärts und dreht ihn mit dem Körper mit. Die Augen sind in Position 131 kurz nach unten gewandt, bleiben die restliche Zeit der Ausführung der Geste aber direkt auf die Bildbetrachter/innen gerichtet. Der geöffnete Zustand des Mantels bleibt drei Sekunden (133–135) lang aufrecht. Die Finger der rechten Hand sind dabei auf Brusthöhe der Person zirka eine Unterarmlänge vom Körper entfernt. Gleichzeitig vollzieht sich eine Bewegung des linken Arms der abgebildeten Person Richtung Mantelinnenseite, die durch das Ausstrecken des linken Zeigefingers abgeschlossen wird. Auf ikonografischer Ebene lässt sich diese auffällige Gebärde als *Geste des Zeigens* identifizieren.

Die Betrachtung der Vergleichsbilder setzt beim planimetrisch fokussiertesten Element, dem geöffneten Mantel, ein (Abbildung 27). Insgesamt sind fünf verschiedene Variationen dieser Grundstruktur dargestellt. Bei erster genauerer Inspektion der Bilder fällt auf, dass Arm-, Hand- und Fingerhaltungen der Ausgangssequenz nicht in allen Vergleichsbildern gegeben sind. Sie fehlen etwa in den Bildern *Cora* und *Daniel*. Von den verbleibenden drei Bildern scheinen auch *Birgit* und *Emil* auf den ersten Blick nicht ganz zur Ausgangssequenz zu passen, da in ihnen beide Arme angehoben und beide Mantelseiten geöffnet sind. Übrig bleibt vorläufig somit nur *Anna*. Sie hat eine Mantelseite mit ihrem Arm geöffnet und blickt frontal in die Kamera. Was dieser Aufnahme fehlt, ist die spezifische Haltung der anderen Hand, die auf ikonografischer Ebene als Zeigen identifiziert werden konnte. Im direkten Vergleich zwischen der Ausgangssequenz und Anna wird nun deutlich, dass die Idee des Zeigens bereits im reinen Öffnen des Mantels enthalten ist. Wie die ikonografische Interpretation der Vergleichsbilder ergibt, handelt es sich dabei jedoch um ein Zeigen *von* gegenüber eines Zeigens *auf*. Anna und Birgit zeigen ihren Mantel, genauer gesagt dessen Innenseite. Sie zeigen jedoch nicht *auf* diese Innenseite. Das Zeigen auf die Innenseite des Mantels, wie sie sich in der spezifischen Arm- und Fingerhaltung in der Ausgangssequenz dokumentiert, stellt somit eine besondere Fokussierung dar.

Irritierend ist an dieser Stelle, dass auf der Innenseite des Mantels in Abbildung 26 nichts weiter zu sehen ist als eben der Mantel selbst. Nachdem der Mantel selbst nicht gemeint sein kann – da es dazu keines zusätzlichen Zeigens bedürfte – bestehen nun zwei prinzipielle Interpretationsmöglichkeiten: Entweder es wird tatsächlich auf *nichts* gezeigt oder es wird auf *etwas* gezeigt, das sich gerade nicht an der bedeuteten Stelle befindet. Phänomenologisch betrachtet ist es ein maßgebliches Charakteristikum

des Nichts, dass sich darauf nicht zeigen lässt. Ein Nichts, auf welches gezeigt wird, ist somit indexikal für ein Etwas aufzufassen, das (derzeit) nicht anwesend ist. Im Vergleich zum Bild *Emil* (Abbildung 27) erfährt dieser Gedanke seine konkrete Visualisierung. Emil trägt etwas im Mantel, das sich nicht genau identifizieren lässt. Formal festzuhalten ist jedoch, dass diese Dinge bei geschlossenem Zustand des Mantels nicht zu sehen sind. Der Träger des Mantels entscheidet folglich selbst, ob er dessen Inhalt anderen preisgibt. Andernfalls bleibt der Inhalt des Mantels Außenstehenden *verborgen*.

Zur vollständigen Rekonstruktion der Geste fehlt nun noch die abschließende Kontextualisierung dieses Mechanismus. Vorikonografisch ist im Bild *Emil* die Strichzeichnung eines Mannes mit Hut, Krawatte, Hemd, Hosenträgern, Hosen und Schuhen zu sehen. Ikonografisch handelt es sich beim Hut um ein Modell der 1930er/40er Jahre – ein Zeitraum, in dem auch die im Bild gezeigten Nadelstreifenhosen und Schuhe mit Gamaschen in der westlichen Welt verbreitet waren. Der Mantel ist unspezifischer, doch lässt er sich im Kontext der übrigen Kleidungsstücke als Trenchcoat identifizieren. Zusammengefasst ist der Schnitt der dargestellten Kleidung im Grunde typisch für die Mode der 1930er/40er Jahre, der Zeit der Prohibition und Weltwirtschaftskrise sowie der Hochblüte des organisierten Verbrechens. Insbesondere Gamaschen und Nadelstreifen waren typische Accessoires von Mafiagangstern<sup>53</sup>. Die betrachtete Geste des Zeigens auf etwas Verborgenes erhält in diesem Zusammenhang die Kontextualisierung des kriminellen Aktes, des Schwarzhandels.

Der Moderator der Sendung „Istanbul Total“ trägt kein *Mafiakostüm*, aber auch er ist *verkleidet*. Wie die ikonografische Betrachtung ergibt, trägt er das Kostüm eines osmanischen Sultans. In diesem Zusammenhang wird vom abgebildeten Bildproduzenten, dem Moderator Stefan Raab, die Geste des Schwarzhändlers in orientalischer Verkleidung dargestellt. Noch überspitzter könnte man formulieren: Osmanische Sultane (und damit das Türkische an sich) werden mit mafiösen Schwarzmarkthändlern gleichgesetzt. Da auf der Mantelinnenseite des als Sultan verkleideten Moderators nichts zu sehen ist, erhält die Geste in der Bilddimension den Charakter des Gespielten, nicht Ernsthaften. Im Modus des Komischen gelingt die Vermittlung stereotyper Inhalte über das Türkische bei gleichzeitiger Vermeidung eines ernsten Beigeschmacks. So ermöglichen es Komik, Verkleidung und Rolle dem

---

<sup>53</sup> Ein weiterführender Hinweis kann auf den Film „Manche mögen's heiß“ (USA, 1959) gemacht werden, der im Chicago des Jahres 1929 spielt. Die kriminellen Bösewichte des Films sind ebenso in Nadelstreifen und Gamaschen gekleidet. Der Anführer der Bande heißt im Film sogar „Gamaschen-Colombo“ oder einfach kurz „Gamasche“.

Moderator, zu Inhalten der Sendung „Istanbul Total“ auf Distanz zu gehen – insbesondere dann, wenn er diese selbst lanciert hat. Letztlich tritt in der Übergegensätzlichkeit der vorgestellten Geste „Jacke aufmachen“ auch eine allgemeine Immunisierungsstrategie von Late-Night-Shows hervor: Von vornherein haben diese durch die Verkleidung eine Art *DistanzierungsfILTER* eingebaut, der Handlungen und Äußerungen des Moderators vor Kritik oder Anfechtungen zu schützen vermag (vgl. Hampl 2010).

#### **5.2.4 Reflektierende Gesamtinterpretation und Zusammenfassung**

Die reflektierende Gesamtinterpretation des Videomaterials von „Istanbul Total“ setzt sich aus den Teilergebnissen der komparativen Montage-, Einstellungs- und Gesten- bzw. Gebärdenvariationen zusammen. Auf der Basis des gattungsimmanenten und gattungsübergreifenden empirischen Vergleichs der Eingangspassage der Fernsehsendung „Istanbul Total“ mit anderen Sendungen sowie mit Vergleichsbildern aus dem Internet konnten spezifische Sendungs- und Gattungscharakteristika in der Bilddimension herausgearbeitet werden. Dabei haben wir uns dem Material „von außen nach innen“ (d. h. von der Interpretation der Montage zu jener der Einstellungen und Gesten und Gebärden) genähert. Im Rahmen der Gattungsanalyse ließ sich die 1:40-minütige Sequenz von „Istanbul Total“ (Abbildung 12 ff.) als Eingangssequenz der deutschen Late-Night-Show „TV Total“ identifizieren, wobei sich auf der Ebene der Montage eine für Eingangspassagen von Late-Night-Shows untypische Schnittfolge zeigte: Üblicherweise dominieren nämlich in den ersten Minuten von Late-Night-Shows Kameraeinstellungen, in denen der Moderator fokussiert ist, oder solche in denen kurz das Publikum eingeblendet wird. Wie mittels des Vergleichs mit der Sendung „Reiselust Istanbul“ (Abbildung 15) gezeigt werden konnte, entspricht die Eingangssequenz von „Istanbul Total“ zu Beginn formal gesehen eher einem Reisevideo als einer Late-Night-Show. Erst im späteren Verlauf nähert sich die Sendung wieder dem gattungsspezifischen Montagegestus von Late-Night-Shows an. So wird im Falle von „Istanbul Total“ das anfängliche Spannungsverhältnis zwischen der Fokussierung des Moderators und der Fokussierung der Stadt Istanbul letztlich zugunsten der Fokussierung des Moderators gelöst. Der vorgegebene Anspruch von „Istanbul Total“ ist dem Produktionsformat der Sendung untergeordnet. Die Stadt Istanbul ist nichts weiter als eine *Kulisse*, vor der der Moderator steht. Dieser Hintergrund ist mit dem großen Fernsehbildschirm in „TV Total“ (Abbildung 17) oder der gemalten Abendskyline der

Harald-Schmidt-Show (Abbildung 13) vergleichbar. Wie die Einstellungsvariation von „Istanbul Total“ (Abbildung 20) ergibt, ist die Kameraperspektive der Sendung flach, die Farben der türkischen Kulisse sind hellblau-blass-kühl verwaschen. Der Moderator steht planimetrisch in deutlichem Farbkontrast zu seiner Umgebung und ist auch durch seine klar umrissene Formgebung von ihr abgegrenzt. Insgesamt dokumentiert sich in diesen Aspekten eine deutliche Opposition zur Stadt Istanbul. Demgegenüber befindet sich der Sprecher des Reisevideos „Reiselust Istanbul“ (Abbildung 15) direkt *in* der Stadt; formal gesehen aus dem Zentrum genommen und in der linken Bildhälfte platziert, sodass die Stadt im perspektivischen Zentrum der Einstellungen aufscheinen kann. Istanbul kommt im Falle des Reisevideos auch auf der Textebene stärker zur Sprache, indem der einheimische Erzähler der Sequenz auf Türkisch zu sprechen beginnt. Zusammenfassend dokumentiert sich in der Eingangspassage von „Istanbul Total“ eine eindeutige Distanzierung der Moderatorperson von seiner Umgebung. Stefan Raab ist in Istanbul und zugleich doch nicht dort. Es handelt sich sozusagen um einen Besuch aus sicherer Distanz, wie er dem deutschen Late-Night-Show-Publikum zumutbar ist.

Auf der Ebene der planimetrischen Komposition konnte die zentrale symmetrische Bildposition des Moderators in der Sendung „Istanbul Total“ (Abbildung 20, 0 sec. und 79 sec.) als typisches Element von Late-Night-Shows identifiziert werden, das seine kulturhistorische Verankerung in christlich-hegemonialen Bildtraditionen findet. Zur Geltendmachung und Behauptung ihrer zentralen Bildstellung sind Moderatoren generell auf eine mit den abbildenden Bildproduzent/inn/en akkordierte Einstellungs- und Montageleistung angewiesen. Dieses Zusammenspiel kann als allgemeines Merkmal massenproduzierter Fernsehsendungen betrachtet werden. Die Vorausplanung und Berechenbarkeit – auch von „spontanem“ Moderatorverhalten – ist eine unvermeidliche Konsequenz, die sich aus den sich regelmäßig wiederholenden Produktionsabläufen sowie dem Ökonomisierungsdruck dieser Sendungen ergibt. Bohnsack (2009, S. 240) bezeichnet die Kollusion von abbildenden und abgebildeten Bildproduzent/inn/en als „Monostrukturierung durch den Showmaster“.

Im Rahmen der Gesten- bzw. Gebärdenvariation konnte schließlich anhand einer besonders auffälligen Gebärde (Abbildung 26), des „Jackeaufmachens“ sowie der (Ver-)Kleidung des Moderators, die Übergegensätzlichkeit zwischen hegemonialer Moderatorstellung und Distanzierung gegenüber der Umgebung genauer herausgearbeitet werden. Im Gewand eines Sultans (d.h. eines ehemaligen osmanischen Herrschers) nimmt der Moderator die Pose eines Schwarzmarkthändlers ein. Dass er auf

die leere Mantelinnenseite zeigt, macht klar, dass wir es hier mit einer Als-ob-Handlung zu tun haben müssen – einer Handlung, die der Moderator außerhalb seiner Verkleidung nicht ausführen würde. Damit entpuppt sich die gespielte Handlung selbst als Verkleidung, wodurch ihr denotativer Kern zum Thema wird: Wie bei einem Theaterstück verweist das in der Verkleidung Gespielte stets zurück auf die in der Verkleidung eingenommene Rolle sowie die umgebende Kulisse. Nicht der Moderator entlarvt sich somit durch die Geste als Schwarzhändler. Vielmehr wird all das, worauf die Verkleidung verweist, mit Schwarzhandel in Verbindung gebracht: die Einwohner und einstigen Herrscher der Stadt Istanbul und der Türkei. Im Sinne einer Übergegensätzlichkeit relativiert die komische Verkleidung einerseits die stereotype Deutung der Handlung durch humoristische Brechung. Andererseits ermöglicht der Modus des Komischen aber erst das unproblematische Einschleusen stereotyper semantischer Viren. Als reine Überbringer von Botschaften getarnt, können sich Sendung und Moderator jeglicher inhaltlicher Kritik entziehen. In Hinblick auf die Gattungsanalyse von Late-Night-Shows kann der Modus der Verkleidung und Distanzierung somit als ein praktikables und effektives Instrument angesehen werden, womit sich dieser Sendungstyp gegen eindeutige Zuschreibungen und Vorwürfe immunisieren kann, was ihn weniger angreifbar macht.





**Abbildung 28 : Stefan Raab präsentiert die Beyaz Show in TV Total, 1. Dez. 2005**

Quelle: <http://tvtotal-bilder.homepage24.de/Showgalerie%20vom%201-12-2005>

## **6 „Istanbul Total“: Die Beyaz-Show**

Bei der Interpretation der Eingangspassage des ersten Sendetags von „Istanbul Total“ ging es mir im vorhergehenden Kapitel insbesondere um die Sensibilisierung für praktische Interpretationsansätze in der Bilddimension. In weiterer Folge möchte ich mich nun stärker an jenen Arbeitsschritten orientieren, die Bohnsack (2009) als allgemeine Ansätze für die Film- und Videointerpretation vorgestellt hat. Eine besondere Rolle spielen dabei zum einen die Interpretation auffälliger Verhaltensmuster seitens der abgebildeten Bildproduzent/inn/en sowie zum anderen die Interpretation des Sprechtextes in Relation zur Bilddimension.

### **6.1 Auswahl der Passage**

Nachdem der Showmoderator Stefan Raab am ersten Sendetag von „Istanbul Total“ vom Balkon aus sein Publikum mit der türkischen Studioumgebung am Bosphorus vertraut gemacht hat, wagt er sich am zweiten Sendetag in das fremde Produktionsstudio einer türkischen Fernsehshow. Der nun folgende Vergleich der Beyaz-Show mit der Sendung „Istanbul Total“ liefert zusätzliche Aufschlüsse hinsichtlich der zuvor in Kapitel 5 anhand der Eingangspassage (erster Sendetag) herausgearbeiteten „Distanzierung von der türkischen Umgebung“, der Mechanismen der „Stereotypisierung“ und „Nostrifizierung“ sowie der „Einheit von abgebildeten und abbildenden Bildproduzent/inn/en“. Letztere ist im Falle der Beyaz-Show am zweiten Sendetag von „Istanbul Total“ in doppelter Weise vollzogen, da es sich bei der türkischen Beyaz-Show praktisch um eine „Show in der Show“ handelt. Die Beyaz-Show wurde vor ihrer Ausstrahlung im Rahmen von „Istanbul Total“ im deutschen Fernsehen nochmals einer Postproduktion unterzogen, wodurch diese Sequenz eine starke Fokussierung erfahren hat. In der knapp zweiminütigen Passage der Beyaz-Show sind anfangs der deutschstämmige Show-Moderator Stefan Raab und später auch sein musikalischer Protegé Max als abgebildete Bildproduzenten in einer fremden, türkischstämmigen Produktionsumgebung beobachtbar. Von hohem Interesse für die Interpretation ist dabei der Umgang Stefan Raabs mit Unsicherheit und Unbeholfenheit. Diese treten sowohl auf der Bild- als auch auf der Textebene in deutlichem Maße zutage. Im Gegensatz zu „Istanbul Total“ sind Stefan Raab als Gast einer fremden Show die Möglichkeiten entzogen, die abbildenden Bildproduzent/inn/en (Kameramänner/-frauen, Regisseur/innen etc.) auf direktem Wege live zu dirigieren. Jedoch stellt das Instrument der Postproduktion ein Mittel dar,

um die bereits aufgenommene Beyaz-Show im Nachhinein zu modifizieren und im Kontext der eigenen Show „Istanbul Total“ zu rekontextualisieren.

## 6.2 Videotranskript

Abbildung 29: Beyaz-Show

Passage (oder Sequenz):	„Auftritt in der Beyaz-Show“
Film (oder Video):	<a href="http://tvtotal.prosieben.de/tvtotal/videos/player/?contentId=14157">http://tvtotal.prosieben.de/tvtotal/videos/player/?contentId=14157</a>
Datum:	11. Mai 2004
Time Code (min.):	03:53–05:43
Dauer:	1:50 min.
Transkription:	Stefan Hampf
Korrektur:	-

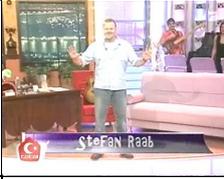
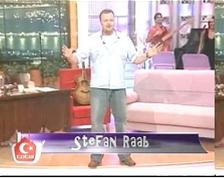
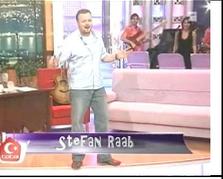
TC:	3:53	3:54	3:55	3:56	3:57
					
Musik:	Signation der Beyaz-Show (Melodie von „Anton aus Tirol“)				
Geräusch:	Brüllen			Klatschen	

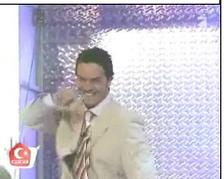
TC:	3:58	3:59	4:00	4:01	4:02
					
Musik:	Signation der Beyaz-Show (Melodie von „Anton aus Tirol“ mit türkischem Text)				
Geräusch:	Klatschen und Pfiffe				

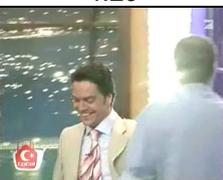
TC:	4:03	4:04	4:05	4:06	4:07
					
Musik:	Signation der Beyaz-Show (Melodie von „Anton aus Tirol“ mit türkischem Text)				
Geräusch:	Klatschen und Pfiffe				

TC:	4:08	4:09	4:10	4:11	4:12
					
Am:		Locker, locker.	Vielen Dank, danke, schön	danke schön.	Hallo.
Geräusch:	Brüllen, Klatschen und Pfiffe				

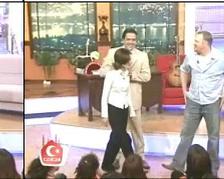
TC:	4:13	4:14	4:15	4:16	4:17
					
Am:	Hallo, mein Name ist Stefan Raab, ich komme aus Deutschland.				Und jetzt, äh,
Geräusch:				Gelächter	

TC:	4:18	4:19	4:20	4:21	4:22
					
Am:		ja,	und jetzt kommt euer Gastgeber.		Hier ist Bey=
Musik:					
Geräusch:	Gelächter				

TC:	4:23	4:24	4:25	4:26	4:27
					
Am:	=yaz				
Musik:	Signation der Beyaz-Show (Melodie von „Anton aus Tirol“ mit türkischem Text)				
Geräusch:	Gebrüll	Gebrüll und Pfiffe			

TC:	4:28	4:29	4:30	4:31	4:32
					
Am:					
Musik:	Signation der Beyaz-Show (Melodie von „Anton aus Tirol“ mit türkischem Text)				
Geräusch:	Gebrüll und Pfiffe			Klatschen	

TC:	4:33	4:34	4:35	4:36	4:37
					
Am:					
Musik:	Signation der Beyaz-Show (Melodie von „Anton aus Tirol“ mit türkischem Text)				
Geräusch:	Klatschen				

TC:	4:38	4:39	4:40	4:41	4:42
					
Am:					
Bm:					Ehhh
Musik:	Signation der Beyaz-Show				
Geräusch:	Klatschen				

TC:	4:43	4:44	4:45	4:46	4:47
					
Bm:	[Spricht türkisch]				
Cm:				Wir könnten natürlich jetzt das übersetzen,	
Musik:					
Geräusch:	Klatschen				

TC:	4:48	4:49	4:50	4:51	4:52
					
Am:					
Cm:	lassn wa aba.	Der Chef hat ja auch nichts verstanden	Zum Glück gibt es aber eines, was alle verstehen		
Musik:					
Geräusch:					

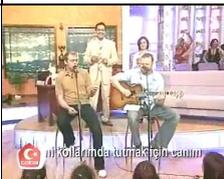
TC:	4:53	4:54	4:55	4:56	4:57
					
Am:					
Cm:		Musik			
Musik:	Ohhho	J u s t c a n ' t w a i t u n t i l t o n i = i g h t , b a b y			
Geräusch:	Rhythmisches Klatschen				

TC:	4:58	4:59	5:00	5:01	5:02
					
Am:					
Cm:				Aba Stefan Raab wäre nicht	
Musik:	T i l l I h a v e y o u b y m y [ l e i s e r ] s i d e , b a = a b y				
Geräusch:	Rhythmisches Klatschen				

TC:	5:03	5:04	5:05	5:06	5:07
					
Am:					
Cm:	Stefan Raab, wenn er sich nicht noch was Besonderes ausgedacht hätte.		Und so haben Stefan und Max		
Musik:	[leise] J u s t c a n ' t w a i t u n t i l t o n i = i g h t , b a b y				
Geräusch:					

TC:	5:08	5:09	5:10	5:11	5:12
					
Am:				[lauter] selbes Lied auf Türkisch gesungen	
Cm:	natürlich noch eine Überraschung mitgebracht				
Musik:				[lauter] selbes Lied auf Türkisch gesungen	
Geräusch:				Rhythmisches Klatschen	

TC:	5:13	5:14	5:15	5:16	5:17
					
Am:	Lied auf Türkisch gesungen				
Musik:	Lied auf Türkisch gesungen				
Geräusch:	Rhythmisches Klatschen				

TC:	5:18	5:19	5:20	5:21	5:22
					
Am:	Lied auf Türkisch gesungen				
Musik:	Lied auf Türkisch gesungen				
Geräusch:	Gelächter		Rhythmisches Klatschen		

TC:	5:23	5:24	5:25	5:26	5:27
					
Am:	Lied auf Türkisch gesungen				
Cm:					One more time
Musik:	Lied auf Türkisch gesungen				Instrumental
Geräusch:	Rhythmisches Klatschen				

TC:	5:28	5:29	5:30	5:31	5:32
					
Am:				Hep beraber	
Musik:	Instrumental	Düdüdü	düduuuu		
Geräusch:	Rhythmisches Klatschen				Gelächter

TC:	5:33	5:34	5:35	5:36	5:37
					
Am:				Hep beraber	
Cm:				Damit haben wir die Türken wohl	
Musik:					
Geräusch:	Gelächter	Klatschen und Pfiffe			

TC:	5:38	5:39	5:40	5:41	5:42
					
Am:					
Cm:	endgültig auf unserer Seite.		Danke, Beyaz!		
Musik:				Schlussakkorde	
Geräusch:	Klatschen			Klatschen, Brüllen und Pfiffe	

### 6.3 Interpretation in der Bilddimension

Wird die Gesamtinterpretation eines Videos oder Films sowohl hinsichtlich der Bild- als auch der Textdimension angestrebt, so empfiehlt Bohnsack (2009, S. 195), aus methodologischen Gründen bei der Interpretation der Bilder zu beginnen. Dadurch ist der/die Forschende von vornherein dazu aufgefordert, sich auf die eigensinnige „Sprache“ der Bilder „einen Reim zu machen“, bevor er/sie diese mit den Interpretationen der Sprechtexte vergleicht. Die so im Vergleich von Bild- und Textdimension zutage tretenden Homologien und Kontraste wären bei einem umgekehrten Prozedere weit weniger ausgeprägt. Dies hat laut Bohnsack (2009) mit dem Erfordernis der grundlegenden Unterscheidung von „sehendem“ und „wiedererkennendem Sehen“ (Imdahl, 1996a) im Rahmen der Bild- und Videointerpretation zu tun. Bereits vorliegende Ergebnisse der Textinterpretation eines Videos bzw. Films vertiefen eher das „wiedererkennende Sehen“, indem sie zum Auffinden der rekonstruierten Konzepte im Bild einladen. Dabei tritt laut Bohnsack

(2009) erschwerend hinzu, dass das Medium Text tendenziell eher eine sequenzielle und das Medium Bild eher eine simultane Struktur aufweist. Die Besonderheit der Bildstruktur ist es, Gegensätze im Sinne sogenannter „Übergegensätzlichkeiten“ (Imdahl, 1996a) zugleich aufwerfen und auflösen zu können. Diese Fähigkeit ist letztlich als grundlegendes Alleinstellungsmerkmal von Bildern anzusehen, selbst wenn Film- bzw. Videobilder – in Abgrenzung zu Standbildern – ebenso einer sequenziellen Struktur folgen wie der Text.

### **6.3.1 Formulierende Interpretation: Sequenzen, Einstellungswechsel und Montage**

Anhand des Fallmaterials von „Istanbul Total“ hat Bohnsack (2009, S. 198) ein Schema entwickelt, nach dem sich die Sequenzen von Fernsehshows und auch einige andere Film- bzw. Videogattungen in ihrer hierarchischen Ordnung verstehen lassen. Prinzipiell lassen sich drei Gruppen von Sequenzen unterscheiden: Hauptsequenzen (HS), Untersequenzen (US) und eingelagerte Sequenzen (ES).

Die gesamte Dauer der Fernsehshow „Istanbul Total“ beträgt pro Sendetag zirka 35 Minuten. Am zweiten Sendetag wurde eine Passage der Beyaz-Show ausgewählt, die insgesamt 1 Minute und 51 Sekunden dauert. Die Szene zeigt die Studiobühne der Beyaz-Show, bestehend aus Showtreppe, Podium mit Schreibtisch, rotem Sessel und beiger Couch. Links im Hintergrund ist durch Fensterrahmen eine Stadt bei Nacht zu sehen. Rechts im Hintergrund sitzt eine fünfköpfige Gruppe von Musikern.

Innerhalb der vorliegenden Passage lassen sich 2 Hauptsequenzen (HS) identifizieren:

- Hauptsequenz I: 3:53–4:52 (60 Sekunden)  
„Stefan Raab, Beyaz und Assistentin kommen auf die Bühne“
- Hauptsequenz II: 4:53–5:43 (51 Sekunden)  
„Stefan Raab und Max singen auf der Bühne“

**Hauptsequenz I** (HS 1) ist die umfangreichste Sequenz. Sie wird in unterschiedlichen Einstellungsgrößen und mit unterschiedlichen Personen in folgenden Untersequenzen (US) realisiert. Der Einstieg erfolgt über einen Establishing Shot<sup>54</sup> (Halbtotale):

- US 1: Stefan Raab alleine auf der Bühne (Halbtotale): 3:53–3:54; 3:59–4:25 (**28“**)
- US 2: Stefan Raab (amerikanische Einstellung): 3:55–3:58
- US 3: Beyaz (amerikanischer Einstellung): 4:26–4:30

---

<sup>54</sup> „meist in der Totale oder Halbtotale durchgedrehte Einstellung, die beim Schnitt als Grundlage für eine Szene genommen werden kann“ (Monaco 2003, S. 103)

- US 4: Stefan Raab und Beyaz auf der Bühne (Zoom Totale zu Halbtotale): 4:31–4:35 (5“)
- US 5: Stefan Raab, Beyaz und Assistentin auf der Bühne (Zoom Halbtotale zu Totale): 4:36–4:50 (15“)

In die Hauptsequenz I sind folgende Sequenzen eingelagert (ES):

- ES 1: Publikum (Totale): 4:06–4:07, 4:51–4:52 (in Summe: 4“)

**Hauptsequenz II** ist beinahe genauso lang wie HS I. Sie wird in unterschiedlichen Einstellungsgrößen und mit unterschiedlichen Personen in folgenden Untersequenzen (US) realisiert. Der Einstieg erfolgt über einen Establishing Shot (Halbtotale). Naheinstellungen sind jeweils Ausschnitte der Halbtotale:

- US 6: Stefan Raab, Beyaz, Frau und Max auf der Bühne (Halbtotale): 4:53–5:02, 5:13–5:24, 5:29–5:33, 5:40–5:43 (36“)
- US 7: Stefan Raab und Max schräg im Profil (Amerikanische): 5:03–5:07
- US 8: Stefan Raabs Gesicht (Close-up): 5:10–5:12

In die Hauptsequenz II sind folgende Sequenzen eingelagert (ES):

- ES 1: Publikum (Halbtotale): 5:08–5:09, 5:25–5:28, 5:33–5:39 (in Summe: 12“)

**Hauptsequenz I:** 3:53–4:52 (60“)

„Stefan Raab, Beyaz und Assistentin kommen auf die Bühne“

Übergang	Sequenz	In	Out	Einstellung	Standort/Bewegung	Thema
	US 1	3:53	3:54	Halbtotale	Fahrt nach rechts; Schwenk nach links	Stefan Raab auf der Bühne
Schnitt						
	US 2	3:55	3:58	Amerikanische	folgt Stefan Raab nach links	Stefan Raab
Schnitt						
	US 1	3:59	4:05	Halbtotale	schneller Zoom aus der Totalen in die Halbtotale und kurz zurück; Fahrt nach links, Schwenk nach rechts	Stefan Raab auf der Bühne
Schnitt						
	ES 1	4:06	4:07	Halbtotale	Fahrt nach links	Publikum
Schnitt						
	US 1	4:07	4:25	Halbtotale	Fahrt nach unten; nach links; Zoom in; Zoom out	Stefan Raab auf der Bühne
Schnitt						
	US 3	4:26	4:30	Amerikanische	folgt Beyaz nach links	Beyaz
Schnitt						

	US 4	4:31	4:35	Totale	Zoom in; Fahrt nach links und Schwenk nach rechts;	Stefan Raab und Beyaz auf der Bühne
	US 5	4:36	4:50	Halbtotale	Zoom out auf Totale; Fahrt nach rechts und Schwenk nach links; Zoom in auf Halbtotale	Stefan Raab, Beyaz und Assistentin auf der Bühne
Schnitt						
	ES 1	4:51	4:52	Halbtotale	Handkamera	Publikum
Schnitt						

**Hauptsequenz II: 4:53–5:43 (51“)**  
„Stefan Raab und Max singen auf der Bühne“

Übergang	Sequenz	In	Out	Einstellung	Standort/Bewegung	Thema
	US 6	4:53	5:02	Halbtotale	Fahrt nach rechts und Schwenk nach links; leichter Zoom in; leichter Zoom out	Stefan Raab, Beyaz, Assistentin und Max auf der Bühne
Überblendung 1“						
	US 7	5:03	5:07	Amerikanische	Drehen aus rechter Schräglage in die Horizontale	Stefan Raab und Max schräg im Profil
Schnitt						
	ES 1	5:08	5:09	Halbtotale	leichter Schwenk nach links unten	Publikum
Schnitt						
	US 8	5:10	5:12	Close-up	unbewegt	Stefan Raabs Gesicht
Überblendung 1“						
	US 6	5:13	5:24	Halbtotale	Fahrt nach rechts; Zoom out; leichte Wackler am Ende	Stefan Raab, Beyaz, Assistentin und Max auf der Bühne
Schnitt						
	ES 1	5:25	5:28	Halbtotale	Handkamera	Publikum
Schnitt						
	US 6	5:29	5:32	Halbtotale	Fahrt nach links; Zoom out	Stefan Raab, Beyaz, Assistentin und Max auf der Bühne
Schnitt						
	ES 1	5:33	5:39	Halbtotale	Handkamera	Publikum
Schnitt						
	US 6	5:40	5:43	Halbtotale	Zoom in; Fahrt nach links und Schwenk nach rechts	Stefan Raab, Beyaz, Assistentin und Max auf der Bühne

### 6.3.2 Auswahl der Fotogramme

Laut Bohnsack (2009, S. 201) gilt als Auswahlkriterium geeigneter Fotogramme für die detaillierte Einzelbildinterpretation, „dass diese Fotogramme zumindest die umfangreichsten Hauptsequenzen (HS) in deren umfangreichsten Einstellungsvarianten, also in den Untersequenzen (US), repräsentieren sollten“. Auf der Basis dieser Überlegungen wurden folgende drei Fotogramme für die weiterführende Interpretation ausgewählt:

#### **Fotogramm 1: US 1 „Stefan Raab auf der Bühne“ (4:08)**

Ist der umfangreichsten Untersequenz (US 1) entnommen und eignet sich gut für die komparative Analyse mit der Studiosequenz der Eingangspassage von „Istanbul Total“ (Abbildung 30).



Abbildung 30: Das „Außenstudio“ von „Istanbul Total“

#### **Fotogramm 2: US 5 „Beyaz, Assistentin und Raab“ (zweitumfangreichste Sequenz) (4:45)**

#### **Fotogramm 3: HS 2, US 6 „Songcontest Song“ (absolut umfangreichste Sequenz) (5:21)**

### 6.3.3 Formulierende und reflektierende Interpretation der Fotogramme

#### 6.3.3.1 Fotogramm 1: Stefan Raab auf der Bühne

##### Formulierende Interpretation



Abbildung 31: Stefan Raab in der Beyaz-Show (4:08)

##### *Vorikonografische Ebene*

Nach den Titeleinblendungen einer Abendshow kommt ein Mann über die Treppe im Bildhintergrund auf die Bühne. Er bleibt an der Vorderkante eines Podests im Bildmittelgrund stehen. Der Körper des Mannes befindet sich in der Bildmitte, ist zur Gänze erkennbar und den Zuschauer/inne/n frontal zugewandt. Seine beiden Arme liegen am Körper an, die Unterarme sind symmetrisch im 45°-Winkel nach unten abgespreizt. Die beiden Handflächen sind zum Boden flach ausgestreckt einander zugewandt. Für die Bildbetrachter/innen sind die leicht nach oben erhobenen Handrücken erkennbar. Der Kopf des Mannes ist leicht nach links gekippt. Sein Gesichtsausdruck ist aufgrund der niedrigen Bildauflösung uneindeutig.

Das Bühnenpodest hat eine helle, glatte Holzoberfläche und wird vorne von einer sanften Krümmung abgeschlossen. Es ist zirka 25 cm hoch und an der Vorderseite weiß lackiert oder von innen beleuchtet. Auf dem Podest befinden sich links ein kleiner, brauner Schreibtisch, neben dem rechts ein roter Sessel steht. Auf der gegenüberliegenden rechten Seite des Podests befindet sich eine beige Couch. Alle drei Möbel sind den Betrachter/inne/n schräg zugewandt, sodass jeweils jene Seite weiter hinten ist, die sich näher an der Bildmitte befindet. Die Vorderseite des Schreibtischs hat an der Unterkante drei abwärtsgekrümmte Halbbögen. Auf der Schreibtischplatte befinden sich verschiedene kleinere Gegenstände. Auf dem roten Ledersessel daneben

ist eine Gitarre in aufrechter Position erkennbar. Sowohl Sessel als auch Gitarre werden vom Mann auf der Bühne teilweise verdeckt.

Unterhalb des Podests sind im Bildvordergrund zirka ein Dutzend Köpfe von hinten zu sehen. Jeweils in Kopfhöhe sind zusätzlich einige erhobene Hände erkennbar. Der Bildhintergrund setzt sich mit einer ca. 35 cm hohen Stufe schräg unterhalb der Bildhälfte vom Rest des Bildes ab. Die Stufe verläuft links und rechts des Podests im Mittelgrund und ist weiß. Unterhalb ihrer Mitte ist ein schmaler, blauer Längsstreifen sichtbar. Auf der Oberseite der Stufe ist ein dunkelbrauner Fußboden zu erkennen, der bis zur Rückwand verläuft. Auf der linken Seite ist die Rückwand gelb-orange bemalt. Auf ihr sind zwei quadratische, dunkelbraune Holzrahmen zu sehen, in denen grau-weiße Punkte und Linien erkennbar sind. An der Unterseite der Rahmen sind jeweils zwei grüne Blumenkistchen befestigt. Zwischen den Rahmen, an deren Oberkante, ist ein kleiner, brauner Punkt zu sehen; genau in der Mitte steht ein Objekt, das an eine stilisierte Katzenfigur mit langem Hals erinnert. Auf der rechten Seite ist die Rückwand in Violettönen gehalten. Sie ist vertikal abwechselnd durch weiße und violette Streifen, Quadrate und Sterne strukturiert. Vor der Rückwand sind vier rot gekleidete Personen erkennbar. Die linken drei sind den Betrachter/inne/n zugewandt. Vor ihnen sind auf Hüfthöhe graue Tafeln angebracht. Die Person an der rechten Bildkante ist im Profil zu erkennen. Sie ist bis zur Hüfte von einer rechteckigen, schwarzen Fläche verdeckt, die an der Oberkante durch eine graue Fläche abgeschlossen ist. Zwischen den beiden Rückwänden ist ein weißer Türrahmen erkennbar, an dessen Unterseite eine weiße Treppe zu sehen ist. Durch den Türrahmen ist schemenhaft eine violett-weiße Fläche erkennbar.

### ***Ikongrafische Ebene***

Der deutsche Showmoderator Stefan Raab befindet sich im Studio der türkischen Beyaz-Show. Das Studio ist nach dem Vorbild typischer Late-Night-Shows eingerichtet. Die Rahmen auf der linken Seite stellen Fensterattrappen dar, die den Blick auf das nächtliche Istanbul vortäuschen. Der weiße Türrahmen ist der Bühneneingang mit Showtreppe. Die vier Personen im rechten Bildhintergrund stellen die Studioband dar. Bei den Köpfen im Bildvordergrund handelt es sich um das sitzende Studiopublikum.

## Reflektierende Interpretation

### *Planimetrische Komposition*



Abbildung 32: Planimetrie

Planimetrisch ist das Bild durch ein Gitter von kleinen und großen Quadraten, Rechtecken und Linien in fünf Bereiche zerteilt: Das linke und rechte obere Bildviertel, dazwischen der weiße Bühneneingang, darunter die freie Bodenfläche mit Podest sowie der Publikumsbereich. Durch seine zentrale Positionierung im Bild verbindet die Person Stefan Raab alle diese Bereiche, mit Ausnahme des Publikumsbereichs. Wie durch die eingezeichnete Linie erkennbar, ist dieser vom Rest des Bildes horizontal klar abgesetzt.

### *Perspektivische Projektion*



Abbildung 33: Perspektivische Projektion

Als perspektivische Anhaltspunkte im Bild dienen die parallelen, horizontalen Linien der Fensterrahmen im Bildhintergrund. Sie verweisen auf eine Frontalperspektive. Da sich die vertikalen Linien nach unten hin verjüngen, ist von einer leicht abwärtsgeneigten

Kamera – und damit von einem zentralen Fluchtpunkt in der oberen Bildhälfte – auszugehen. Er dürfte sich auf der Horizontlinie in etwa zwischen Brust und Bauch von Stefan Raab befinden.

### ***Szenische Choreografie***

Auf der Ebene der szenischen Choreografie sind drei Personenkreise zu erkennen, die in bestimmter Weise aufeinander bezogen sind. Stefan Raab steht in der Bühnenmitte frontal einer im Bildvordergrund sitzenden Menschengruppe sowie den Bildbetrachter/inne/n gegenüber (beides Publikum). Ebenso frontal zugewandt ist die Gruppe von Musikern, die im Hintergrund des rechten, oberen Bildviertels zu sehen ist. Die Musiker scheinen für die Menschen im Vordergrund zu spielen, die mit ihnen händeklatschend in Verbindung stehen. Die Geste Stefan Raabs, der mit dem Publikum im Mittelgrund im Blickkontakt steht, ist beschwichtigend.

### ***Ikonologisch-ikonische Interpretation***

Das Fotogramm eignet sich zur komparativen Analyse mit den Studioaufnahmen der Eingangspassage von „Istanbul Total“ (Abbildung 30). Auf der Ebene der formalen Struktur des Bildes dokumentieren sich zunächst folgende zwei Gemeinsamkeiten zwischen den Fotogrammen: In beiden Fällen ist eine Trennung und Distanzierung erkennbar, die sich zwischen Bühnenpodest und Publikum manifestiert, welches planimetrisch vom Rest des Bildes abgetrennt ist, in szenischer Hinsicht sind sie in der Beyaz-Show sowohl Stefan Raab als auch den Musikern im Hintergrund oppositionell zugewandt. Weiters ist auch im vorliegenden Fotogramm von einer mehrfachen Zentrierung der Person auszugehen. In planimetrischer Hinsicht repräsentiert Stefan Raab eine Übergegensätzlichkeit. Durch seine räumliche Positionierung entsteht eine Synthese der oberen Bildbereiche. Gleichzeitig wird durch seine Position an der Podestkante, durch den deutlichen Farbkontrast zwischen Hemd und Hose sowie die flache Handhaltung die räumliche Trennung zum Publikum verstärkt. Auf der Ebene der szenischen Choreografie wird dieser Befund durch die frontale Gegenposition gestützt, die die abgebildeten Bildproduzenten Raab und Publikum einnehmen. Sowohl im Klatschen des Publikums als auch in der abwiegelnden Gestikulation von Raabs Händen dokumentiert sich eine gegenseitige Bezogenheit im Modus der Distanzierung. Raab wirkt abwartend befremdet. Der türkische Bühnenkontext scheint ihm nicht ganz geheuer.

So dokumentiert sich im Falle Stefan Raabs eine große Unsicherheit sowie ein „Kontrollverlust, der nicht eigentlich komisch ist“ (Bohnsack, 2009, S. 222). Er manifestiert sich in der „beruhigenden Gestik“, in der psychischen Distanz zum Publikum durch Verbleib am Podium, sowie – wie die Zeitlupenanalyse (s.u.) erweist – im mehrfachen Umdrehen zum Bühneneingang. In dieser letzten Geste dokumentiert sich Raabs Bemühung, mittels Blickkontakt den Beistand des Showmasters Beyaz herbeizurufen, der aber nicht erscheint. Dass Raab auf die Unterstützung Beyaz‘ hofft, zeigt sich darin, dass er – nach dreimaligem Umdrehen – verbal seinen Auftritt in der Show ankündigt.

### ***Einschub: Zeitlupenanalyse***

#### *Detailinterpretation einer fokussierten Geste – Das mehrmalige Umdrehen Stefan Raabs*

Stefan Raab dreht sich während seines Auftritts dreimal deutlich nach hinten zum Bühneneingang um. Die Körperdrehungen dauern jeweils etwa zehn Sekunden. Im Folgenden sind die Fotogramme der Bewegungsabläufe dargestellt. Raabs Bühnenauftritt beginnt in Abbildung 34 mit der Geste des sogenannten „Beyaz-Moves“ (s. u.). Danach erfolgt die (unwillkürliche) Gebärde der ersten Blickwendung.



**Abbildung 34: Beyaz-Show (4:02-4:03), 4 Bilder pro Sekunde**



Abbildung 35: Beyaz-Show (4:04–4:05), 4 Bilder pro Sekunde



Abbildung 36: Beyaz-Show (4:09–4:10), 4 Bilder pro Sekunde



Im Vergleich zu den Erkenntnissen zur Eingangspassage werden daneben folgende Kontraste deutlich: Aus planimetrischer Sicht ist das vorliegende Fotogramme von oben bis unten klar und detailliert strukturiert. Es gibt keine diffusen Bereiche wie beispielsweise in der mit Plüschpolstern und orientalischen Teppichen ausgestatteten Sitzzecke des „Istanbul Total“-Studios (Bohnsack, 2009, S. 210). Die klare Struktur wird auch durch die Zentralperspektive unterstützt, die dem Bild einen symmetrischen Aufbau verleiht. Im Gegensatz zur Eingangspassage von „Istanbul Total“ (Kap. 5) findet sich weiters keine formale Distanzierung zwischen Innen- und Außenwelt (Bohnsack, 2009, S. 210) – auf ein Außen wird bestenfalls durch die beiden stilisierten Fenster verwiesen. Dieses Außen ist aber nicht eindeutig identifizierbar, sondern lediglich erahnbar. In der Gattung der Late-Night-Shows handelt es sich in der Regel um eine

Nachtaufnahme der Stadt<sup>55</sup>, aus der die Sendung stammt. Das Außen ist hier folglich nur Schablone.

In Bohnsacks (2009, S. 208) Bildinterpretation der Sitzecke im „Istanbul-Total“-Studio dominiert ein opakes und diffuses Bild der Türkei. In komparativer Analyse mit dem Studio der Beyaz-Show ergibt sich eine deutliche Validierung dieser Annahme. Das Bild wirkt hier ausgewogener und ruhiger. Es ist klar strukturiert und aufgeräumt. Sämtliche Dinge sind an ihrem Platz und funktional zugeordnet: der Schreibtisch des Moderators, die Couch für die Gäste, die Ecke für die Musiker etc. Auffallend ist außerdem, dass das Podium, auf dem sich das Handlungsgeschehen abspielt, dem Publikum direkt zugewandt ist. Dies ist ein Kontrast zum „Istanbul-Total“-Studio, wo die „orientalische Sitzecke“ ja seitlich angeordnet und sogar noch durch eine Bar vom Publikum getrennt ist. Im „Istanbul-Total“-Studio sprechen der Moderator und seine Gäste in der Sitzecke nur direkt in die Kamera, im Studio der Beyaz-Show zur Kamera und zum Publikum. Es gibt hier keine versteckten, dem Publikum abgewandten Räume. In der größeren Offenheit dieser Konstellation dokumentiert sich auch eine geringere Exklusivität der Podiumsplätze als im Studio von „Istanbul Total“.

### 6.3.3.2 Fotogramm 2: Beyaz, Assistentin und Stefan Raab

Beim Vergleich zwischen dem ersten (Abbildung 31) und dem zweiten Fotogramm der Beyaz-Show (Abbildung 37) beschränken sich die Beschreibungen und Interpretationen im Sinne der Forschungsökonomie auf die *Veränderungen* des Bildes.

#### Formulierende Interpretation



Abbildung 37: Beyaz, Assistentin und Stefan Raab (4:45)

<sup>55</sup> Eine Nachaufnahme gibt es auch bei der Harald-Schmidt-Show (vgl. Kap. 5)

### ***Vorikonografische Ebene***

Der Bildausschnitt hat sich im zweiten Fotogramm auf die Halbtotale verengt und nach links verschoben. Nach wie vor sichtbar sind die beiden stilisierten Fensterrahmen im Hintergrund sowie die Treppe dort. Von den Personen rechts im Bild ist nur noch die äußerst linke erkennbar. Es handelt sich um einen Mann mit schwarzen Haaren, dessen Kopf und Oberkörper nach rechts gewandt sind. Die Personen am unteren Bildrand sind fast zur Gänze verschwunden – fünf Kopfrundungen sind noch zu erkennen. Das Podium, der darauf befindliche Schreibtisch, der rote Sessel sowie die Couch werden von drei Personen im Bildvordergrund verdeckt. Die Füße der Personen sind abgeschnitten. Bei der linken Person handelt es sich um einen Mann mit kurzen, schwarzen Haaren. Sein Kopf sowie sein rechtes Bein sind den Bildbetrachter/inne/n frontal zugewandt. Der Mund ist leicht geöffnet. Der Mann trägt einen beige Herrenanzug mit weißem Hemd und einer Krawatte mit breiten, braunen Schrägstreifen. Das Sakko ist geöffnet und fällt aufgrund der Körperhaltung des Mannes auseinander, sodass ein brauner Gürtel gut erkennbar ist. Seine rechte Hand hat der Mann auf Schulterhöhe nach vorne erhoben und geöffnet, der Daumen weist nach oben. Sein rechtes Bein ist in Richtung Bildmitte gedreht, die linke Hand ruht auf der Schulter einer zweiten Person in der Bildmitte. Diese Person ist zirka einen ganzen Kopf kleiner als die erste. Es handelt sich um eine zierliche Frau, deren Körper den Bildbetrachter/inne/n frontal zugewandt ist. Ihre Beine und Schultern sind parallel, die Arme fallen links und rechts gerade herab. Die Frau trägt eine schwarze Hose und eine schlichte, weiße Kragenbluse darüber. Sie hat dunkles, schulterlanges Haar. Ihr Kopf ist beinahe vollständig ins Profil gedreht. Ihr Blick fällt auf die dritte Person rechts im Bild, die aus Fotogramm 1 bekannt ist. Es handelt sich um einen Mann mit kurzen, brünetten Haaren und Schnurrbart. Er trägt über einer dunkelblauen Jeans ein hellblaues Hemd mit hochgekrempeelten Ärmeln. Unter dem Hemd ist beim Kragen ein weißes Unterhemd erkennbar. Der Mann ist den Bildbetrachter/inne/n zu drei Vierteln zugewandt. Sein rechtes Bein und sein Kopf weisen Richtung Bildmitte. Beide Arme sind angewinkelt. Die Finger der Hände stecken in den vorderen Hosentaschen. An der linken Hand trägt der Mann eine Uhr mit braunem Armband.

### ***Ikongrafische Ebene***

Zuzüglich der Informationen aus dem ersten Fotogramm lässt sich im vorliegenden Bild der türkische Showmoderator Beyaz erkennen. Bei der Person in der Mitte dürfte es sich um eine Assistentin handeln, die vermutlich als Übersetzerin für Stefan Raab agiert.

### **Reflektierende Interpretation**

#### ***Planimetrische Komposition***



**Abbildung 38: Planimetrie**

Das Bild ist annähernd symmetrisch. Die schwarzen, parallelen Hosenbeine der Assistentin stellen einen starken visuellen Kontrast dar. Zwischen ihnen entsteht ein Einschnitt, der das Bild vertikal in zwei annähernd gleich große Hälften teilt. Die Linie setzt sich bis ganz nach oben fort – entlang der unten spitz geöffneten Bluse über das mittig am Kragen befestigte Mikrofon und den rechten Fensterrahmen. Die Vertikalität des Bildes wird durch die links und rechts stehenden Personen Beyaz und Stefan Raab verstärkt, gleichzeitig aber auch abgemildert. Entlang ihrer zur Mitte gedrehten Körper sind leicht nach außen gekippte Vertikalen erkennbar. Insgesamt strebt dadurch das Bild symmetrisch wie ein Springbrunnen nach oben und außen. Dieser Befund steht im Kontrast zum ersten Fotogramm der Beyaz-Show, das planimetrisch von einer starken Horizontalität geprägt war. Wie ersichtlich, ist dieser Wechsel auf Leistungen der abgebildeten Bildproduzent/inn/en zurückzuführen.

## ***Perspektivische Projektion***



**Abbildung 39: Perspektive**

Eine eindeutige Rekonstruktion der Perspektive ist aufgrund mangelnder Cues im Bild nicht möglich. Die gekrümmte Vorderkante des Podests im Mittelgrund könnte sowohl für eine Parallel- als auch eine Schrägperspektive sprechen. Bedingt durch die nach rechts hinten verlaufende Podestkante entsteht der Eindruck einer leichten Schrägperspektive. Die annähernd gleich großen Männer im Vordergrund legen aber eher eine Zentralperspektive nahe. Zusammenfassend ist von zumindest einem Fluchtpunkt auszugehen, der sich auf der Mittellinie oder leicht rechts davon befindet. Die Horizontlinie dürfte sich zirka auf Augenhöhe der Männer befinden, da das Gesicht von Beyaz den Bildbetrachter/inne/n frontal und unverzerrt zugewandt ist. Dies bedeutet, dass die Kamera nach unten geneigt sein muss, was sich auch darin dokumentiert, dass die Oberseiten von Podest, Sofa und Schreibtisch sowie der Fußboden gut erkennbar sind.

## ***Szenische Choreografie***

Wie der formulierenden Interpretation zu entnehmen, sind im Bild mehrere Personen erkennbar. Im Bildzentrum befinden sich drei Personen. Die linke von ihnen, der Showmoderator Beyaz, blickt und spricht frontal in Richtung der aufzeichnenden Kamera. Das am Boden sitzende Publikum wird von ihm nur in zweiter Hinsicht adressiert. Zwar ist er ihm in seiner Körperhaltung zugewandt, sein Blick geht jedoch über das Publikum hinweg. Die Haltung von Beyaz' rechter Hand deutet auf eine argumentative oder erläuternde Geste hin. Mit seiner linken Hand hält er die Schulter der Assistentin fest, deren Blick nach rechts zu Stefan Raab gewandt ist. Dass es sich nicht nur um ein Abstützen der Hand, sondern tatsächlich um ein Festhalten an einer

bestimmten Position handelt, dokumentiert sich sowohl am erkennbaren linken Daumen von Beyaz als auch deutlich in der Zeitlupenanalyse 4:41–4:50. Die Assistentin ist somit von Beyaz zwischen ihm und Raab in Position gebracht. Ihre Aufgabe an dieser Stelle scheint jedoch nicht klar definiert zu sein. Dies wird daran deutlich, dass Stefan Raab über die kleine Assistentin hinweg zu Beyaz blickt und lächelt. Raabs Hände stecken beide bis zur Hälfte in den vorderen Hosentaschen seiner Jeans. Offensichtlich versucht er sie gerade hineinzustecken oder herauszuziehen, was einer gewissen Anstrengung bedarf, da Jeanstaschen eng sind. Die Finger der linken Hand sind jedenfalls gut erkennbar. Wie die Zeitlupenanalyse 4:42–4:28 zeigt, handelt es sich dabei um eine „Verlegenheitsgeste“.

Der Musiker in der rechten oberen Bildhälfte steht mit keinem anderen Bildakteur in Verbindung. Sowohl sein Blick als auch sein Körper sind vom zentralen Bildgeschehen abgewandt. Seine Tätigkeit bezieht sich somit offensichtlich auf Vorgänge, die außerhalb des Bildes stattfinden. Das Publikum am unteren Bildrand ist streng nach vorne hin orientiert. Aufgrund der Sitzposition am Boden ist davon auszugehen, dass es zu den Akteur/inn/en auf der Bühne aufblicken muss, um der Show zu folgen.

### ***Ikonologisch-ikonische Interpretation***

Im Vergleich zu Fotogramm 1 fällt planimetrisch unmittelbar die Aufhebung der formalen Distanzierung durch die abgebildeten Bildproduzent/inn/en ins Auge. Dadurch, dass die Personen vom Podium heruntergestiegen sind, gibt es keine physischen Barrieren zum Publikum mehr. In dieser Position begreift sich Beyaz klar in seiner Rolle als Showmoderator, wie an seiner aufrecht nach vorne gerichteten Körperhaltung, dem entschlossenen Blick und der Gestik der rechten Hand deutlich wird. Mit seiner linken Hand versichert er sich, dass sich auch die Übersetzerin an ihrem Platz befindet und von dort nicht fortbewegt. Diese scheint sich ihrer Position weniger sicher zu sein als Beyaz, lässt ihn gewähren und verharnt steif in der ihr zugewiesenen Lage. Lediglich ihren Kopf bewegt sie in Richtung Stefan Raab, der von ihr jedoch keine Notiz nimmt. Insgesamt entsteht dadurch der Eindruck, dass die Übersetzerin weder fixer Bestandteil der Show zu sein scheint noch im Vorhinein klar vereinbart wurde, in welcher Weise sie daran teilnehmen sollte. Offensichtlich ist jedenfalls, dass die konkrete Position und Rolle der Übersetzerin für sie selbst und für Stefan Raab im vorliegenden Bild nicht vollständig geklärt zu sein scheinen. Dies steht im Gegensatz zu ihrer klaren Positionierung durch Beyaz sowie ihrer markanten Dreifachfokussierung

im Bild: Die Übersetzerin ist in planimetrischer Hinsicht fokussiert und steht außerdem am nächsten zur Bildmitte. Im Gegensatz zu Stefan Raab hat Beyaz damit nur eine gewisse Kontrolle über die abgebildeten Bildproduzent/inn/en. Somit besteht eine deutliche Differenz zur Monostrukturierung der Sendung „Istanbul Total“ (Bohnsack, 2009) oder „TV Total“ durch den Moderator Stefan Raab. Dieser dirigiert sowohl die abgebildeten wie die abbildenden Bildproduzent/inn/en.

Verwirrung besteht in der Beyaz-Show nur für den uneingeweihten Kreis der abgebildeten Bildproduzent/inn/en. Im vorliegenden Falle gehört dazu auch Stefan Raab. Seine Orientierungslosigkeit dokumentiert sich insbesondere in seinem Blick zu Beyaz, aber auch in der – wie die Zeitlupenanalyse zeigt, sogar mehrfachen – Bemühung, durch Einstecken der Hände in die Hosentaschen lässige Entspannung zu vermitteln.

### 6.3.3.3 Fotogramm 3: Songcontest Song



Abbildung 40: Deutscher Songcontest-Song in der Beyaz-Show

## Formulierende Interpretation

### *Vorikonografische Ebene*

Im Vergleich zum zweiten Fotogramm ist hier wieder etwas mehr vom Studio zu sehen, gleichzeitig hat sich der Bildausschnitt nach links verschoben, sodass der rechte Fensterrahmen nun genau auf der Mittellinie des Bildes liegt. Am unteren Bildrand nimmt das Publikum wieder zirka ein Viertel des Bildes ein. Der Musiker am rechten Bildrand ist nur noch ansatzweise erkennbar.

Im Bildmittelgrund hat sich ein neues Ensemble formiert. Es sind nun vier Personen den Betrachter/inne/n frontal zugewandt. Beyaz und seine Assistentin stehen auf dem Podium, während vor ihnen Stefan Raab und links neben ihm ein Mann in

hellbraunem T-Shirt und dunkler Hose auf Barhockern Platz genommen haben. Stefan Raab hat auf seinem rechten Oberschenkel eine Gitarre aufgestützt. Seine linke Hand hält das Griffbrett. Beide Füße ruhen auf dem Fußraster des Barhockers. Der Mann neben ihm hält auf Brusthöhe einen kleinen Zettel in der linken Hand und hat seine Augen darauf gerichtet. Sein rechter Unterarm ruht auf dem rechten Oberschenkel, die Finger befinden sich in einer schnipsenden Position. Das linke Bein ist angewinkelt, das rechte hängt am Barhocker hinunter. Vor beiden Männern befindet sich auf Mundhöhe ein Mikrofonständer mit Mikrofon. Durch das Ensemble der vier Personen sind der Schreibtisch sowie der rote Sessel fast völlig verdeckt. Im unteren Bildbereich ist die Schrift „(e)kleyemem akşama kadar canım“ auszumachen. Dies ist die türkische Übersetzung des Songtextes „Waiting until tonight, baby“.

### ***Ikonomografische Ebene***

Der neu hinzugekommene Mann heißt Max Mutzke und ist der deutsche Kandidat zum Eurovisions-Songcontest 2004. Er ist der Gewinner der speziellen „TV-Total“-Show „SSDSGPS“ (Stefan sucht den Super-Grand-Prix-Star), die im Vorfeld des Songcontests stattfand. Max' Song wurde von Stefan Raab geschrieben und produziert. Aufgrund der Körperhaltung von Max und Raab sowie der verwendeten Gitarre ist anzunehmen, dass beide gerade gemeinsam singen. Bei der eingeblendeten Schrift handelt es sich um türkische Liedtexteinblendungen.

### ***Planimetrische Komposition***



**Abbildung 41: Planimetrie**

Auf planimetrischer Ebene hebt sich die Personengruppe im Bildzentrum durch ihre runde Form von der sonst eher horizontal strukturierten Umgebung ab. Durch diese

Formation fällt das Bild jedoch keineswegs auseinander, da sich im Personenensemble auch Linien wiederfinden, die an die Umgebung anschließen – so etwa die Vertikalen der Mikrofonständer oder die horizontal angeordneten Köpfe von Max und Raab. Auch das Griffbrett der Gitarre ist horizontal ausgerichtet.

### ***Perspektivische Konstruktion***



**Abbildung 42: Perspektive**

Aus perspektivischer Sicht handelt es sich beim vorliegenden Bild erneut um eine Frontalperspektive. Der Fluchtpunkt befindet sich vermutlich auf der linken Hand von Beyaz, ist aber aus dem Bild nicht eindeutig rekonstruierbar. Die Horizontlinie verlief damit etwas oberhalb der Bildmitte auf Augenhöhe von Max und Stefan Raab.

### ***Szenische Choreografie***

Im Wesentlichen sind alle vier Bühnenakteure den Betrachter/inne/n frontal zugewandt. Die einzige Ausnahme bildet Max, dessen Kopf ins Dreiviertelprofil gedreht ist. Er ist auch der Einzige, dessen Blick nicht zu den Bildbetrachter/inne/n, sondern auf einen Zettel in seiner Hand<sup>56</sup> gerichtet ist. Vor dem Bühnenensemble ist das ihm zugewandte Publikum zu sehen. Die Zuschauer/innen sind allerdings nicht identifizierbar, sondern können nur anhand der Rückseite ihrer Köpfe wahrgenommen werden.

### ***Ikonologisch-ikonische Interpretation***

Im vorliegenden Fotogramm ist der Showmoderator Beyaz mehrfach fokussiert. Er steht genau in der Bildmitte an oberster Position sowie im planimetrischen Zentrum. Seine Hand befindet sich genau im Fluchtpunkt der perspektivischen Projektion. Farblich hebt

---

<sup>56</sup> vermutlich ein sogenannter „Spickzettel“ mit dem türkischen Text des Liedes.

sich Beyaz noch zusätzlich durch seinen hellen Anzug von der Umgebung ab. In seinem Lächeln und seiner Gestik dokumentiert sich entspannte Zufriedenheit mit dem gegenwärtigen Bühnengeschehen.

Im Gegensatz zum zweiten Fotogramm sind diesmal die Rollen klar verteilt und die genauen Positionen der Agierenden auf der Bühne vorgegeben. Stefan Raab und sein Songcontest-Kandidat Max Mutzke haben sich auf zwei Barhocker vor dem Publikum gesetzt und eine aktive Rolle als Musiker eingenommen. Diese aus seiner eigenen TV-Show bzw. auch aus anderen Showauftritten (z. B. bei Anke Engelke) vertraute Rolle stiftet deutlich Sicherheit, wie an der entspannten Haltung Stefan Raabs sichtbar wird.

Die Unsicherheiten und „Befremdungen“, die sich in Form von Distanzierungen sowohl formal im Bild wiedergefunden als auch in Raabs Körperhaltungen (Gebärden) und Gesten dokumentiert haben, sind verschwunden. Wie den türkischen Liedtexteinblendungen entnommen werden kann, singt Stefan Raab jetzt sogar auf Türkisch. Letzteres wird daran deutlich, dass der Sänger Max Mutzke von einem Blatt Papier abliest, was bei seinem vertrauten Liedtext nicht nötig wäre. Die türkische Texteinblendung entspricht wohl sinngemäß den Worten des deutschen Songcontest-Liedes.

#### **6.3.4 Reflektierende Interpretation von Einstellungswechseln und Montage**

Im Kontrast zum Bildmaterial der Stefan-Raab-Shows „Istanbul Total“ und „TV Total“, wo Szenen- bzw. Sequenzwechsel in der Regel durch Zeige- oder Blickgesten des Moderators Stefan Raab dirigiert werden (vgl. auch Bohnsack 2008), fällt in der Beyaz-Show ein davon abweichendes Prinzip des Einstellungswechsels und der Montage auf. Dieses abweichende Prinzip ergibt sich aus den Leistungen zweier abbildender Bildproduzenten – dem Produktionsteam der Beyaz-Show einerseits und dem Produktionsteam von „Istanbul Total“ andererseits.

Deutlich wird dieser Umstand im Vergleich zweier Einstellungswechsel, bei dem jeweils kurz das Publikum in der eingelagerten Sequenz ES 1 zu sehen ist. Zum Zeitpunkt 4:06–4:07 wird US 1 durch ES 1 unterbrochen: Stefan Raab ist gerade auf die Bühne gekommen, winkt dem Publikum zu, das Publikum wird gezeigt. Anschließend wird die Sequenz US 1 fortgesetzt. Diese Sequenz entspricht einem für die Gattung der Fernsehshows typischen Montageschritt, bei dem durch einen Einstellungswechsel das Publikum als Gegenüber des Bühnenperformers gezeigt wird. Empirisch lässt sich dies auch an der Eingangssequenz von „Istanbul Total“ verdeutlichen. Bei 1:08, 1:22–1:27,

2:27 etc. gibt es jeweils eingelagerte Sequenzen, auf denen das Publikum zu sehen ist. Besonders bei Applaus und Gelächter wird dieses Mittel häufig eingesetzt, um den Echtheitsgehalt dieser Ereignisse zu unterstreichen und darüber hinaus zu dokumentieren, dass der Raum der Show mehr als nur die Bühne umfasst. Im Sinne Hickethiers (1996) dient diese Vorgangsweise schließlich zur Herstellung des „filmischen Raums“. Ein wichtiges Charakteristikum dieses Montagetyps ist die Gleichzeitigkeit, die nur durch Parallelmontage hergestellt werden kann. Somit gilt als sehr wahrscheinlich, dass der vorliegende Einstellungswechsel vom Produktionsteam der Beyaz-Show vorgenommen wurde.

Anders verhält es sich mit dem Einstellungswechsel 4:50/4:51. Auch hier wird eine laufende Untersequenz (US 5) von der Publikumseinstellung (ES 1) unterbrochen. Allerdings wird die einleitende Sequenz (US 5) danach nicht fortgesetzt, sondern von einer neuen Sequenz (US 6) abgelöst, auf der ein anderes Bühnenarrangement und andere Personen zu sehen sind. Damit kann es sich um keine Parallelmontage bzw. Bildmischung mehr handeln. Der daraus resultierende Bruch in der Zeitkoordinate ist ein typisches Zeichen einer sogenannten Postproduktion. Der kontinuierliche Zeitablauf wird gerafft, für die abbildenden Bildproduzent/inn/en Irrelevantes herausgeschnitten. Nun wäre es zwar vorstellbar, dass die Beyaz-Show in der Türkei in geraffter Weise ausgestrahlt würde. Wenig wahrscheinlich ist allerdings, dass ausgerechnet der Gastgeber Beyaz mitten in seiner eigenen Show herausgeschnitten würde. Insbesondere aufgrund der herausgearbeiteten Personalunion zwischen abbildenden und abgebildeten Bildproduzent/inn/en erscheint dies wenig plausibel. Beim vorliegenden Einstellungswechsel kann es sich somit nur um eine Leistung der abbildenden Bildproduzent/inn/en der „Istanbul-Total“-Show im Rahmen der Postproduktion handeln. Was die Zusehenden folglich im Rahmen dieser Darbietung der Beyaz-Show zu sehen bekommen, ist letztendlich kein Eindruck der Sendung an sich, sondern dessen, was als Ergebnis der Postproduktion des „Istanbul-Total“-Teams angesehen werden kann.

### **6.3.5 Fokussierungen: Steigerungen, Verdichtungen und Diskontinuitäten**

#### **6.3.5.1 Fokussierungen im Bereich der Performanz der abbildenden Bildproduzent/inn/en**

In der Beyaz-Show gibt es keine fixen Kameras wie in „Istanbul Total“. Die gesamte Show wird entweder mit Handkameras oder mit einer an einem Kran befestigten Kamera

aufgenommen, die direkt über dem Publikum schwebt. Die Krankamera bewegt sich ständig frei im Raum und ermöglicht dadurch den Fernsehzuschauer/innen ein sehr unmittelbares Bühnenerlebnis. Seitens der abgebildeten Bildproduzent/innen kann diese Kamera rasch und spontan zum Einsatz gebracht werden, was auf Kosten perfekt ausgerichteter Einstellungen geht (etwa der Zentralperspektive). Es handelt sich dabei aber um visuelle Einschränkungen, die offensichtlich zugunsten der erhöhten Beweglichkeit in Kauf genommen werden. In „Istanbul Total“ kommen – neben Handkameras – für viele Einstellungen hingegen fixe Studiokameras zum Einsatz. Diese ermöglichen ein höheres Maß der Kontrolle über den Bildausschnitt und den Aufnahmewinkel, als dies bei einer Krankamera der Fall ist, die sich in allen drei Raumdimensionen bewegen kann.

### **6.3.5.2 Fokussierungen im Bereich der Performanz der abgebildeten Bildproduzent/innen**

Als Fokussierung findet sich der sogenannte „Beyaz-Move“, wie Stefan Raab ihn stets bezeichnet. Es handelt sich dabei um eine eigentümlich rhythmische Bewegung der Arme und Beine, die Beyaz üblicherweise zu Beginn seiner Show zur Anwendung bringt. Der „Beyaz-Move“ wird von Stefan Raab nachgeahmt und sowohl beim Interview mit Beyaz<sup>57</sup> als auch bei der Beyaz-Show<sup>58</sup> von ihm mehrmals ausgeführt. Wie Beyaz im Interview erklärt, habe er sich die Bewegungen seinerzeit aus dem Nachtprogramm des deutschen Privatfernsehens abgeschaut. Nun ist sie zu seinem Markenzeichen geworden.

Zu Beginn der Beyaz-Show tritt Stefan Raab an Beyaz' Stelle als Erster mit dem Beyaz-Move auf, was ihm einen sicheren und wirksamen Auftritt verschafft. Die anschließende spontane Kommunikation mit dem Publikum misslingt. Raab kommt weder mit seinen Gesten und seiner Sprache noch mit seinem Humor bei den Leuten an. Bei Beyaz' Auftritt wiederholt Raab zur Begrüßung den Beyaz-Move. Deutlich ist somit die Kommunikationsfunktion dieser Bewegung für Raab – Kommunikation sowohl mit dem Publikum als auch mit Beyaz. Gleichzeitig kopiert Raab dabei den Bewegungsablauf des türkischen Moderators. Gelungene Kommunikation mit dem Türkischen ist somit – nach Vorstellung von „Istanbul Total“ – nur in der Nachahmung bzw. im Nachäffen von standardisiertem Verhalten möglich. Wie „Istanbul Total“ den Zusehenden zu verstehen gibt (dies ist in der Vorgeschichte zur Beyaz-Show massiv spürbar), ist die Nachahmung

---

<sup>57</sup> am 11. Mai 2004, link: <http://tvtotal.prosieben.de/tvtotal/videos/player/?contentId=14156>

<sup>58</sup> <http://tvtotal.prosieben.de/tvtotal/videos/player/?contentId=14157>

legitim, da die Türken ja selber alles von den Deutschen kopieren. Eine interessante Homologie zu dieser Erkenntnis findet sich auch auf der Textebene. Am Ende des Liedes fordert Raab das türkische Publikum mit dem türkischen Ausruf „Hep beraber!“ („Alle zusammen!“) zum Mitsingen auf. Auch dieser Ausruf schlägt die Brücke zum Publikum, das mit Applaus antwortet.

Das eingeübte türkische Handlungsrepertoire bewahrt den Moderator Stefan Raab somit davor, vor der Kamera als unsicher und hilflos bloßgestellt zu werden. Wie die komparative Analyse und das Beispiel der Beyaz-Show zeigen, scheitert Raab regelmäßig in Situationen, wo echte Spontaneität gefragt ist. Er ist eher Experte für das Sprunghafte, die Diskontinuitäten, die das „Raabhafte“ ausmachen. Aufschlussreiche Erkenntnisse über das Scheitern von Raab liefert die Interpretation der Text- bzw. Tönebene.

## 6.4 Interpretation in der Dimension von Text und Ton

### 6.4.1 Text-Transkript

Am=Stefan Raab, Bm=Beyaz, Cm=Off-Kommentator, Dm=Max Mutzke

	4:13	
Am:	Locker, locker, vielen Dank, danke schön, danke schön. Hallo. Hallo. Mein Name ist Stefan Raab, ich komme aus Deutschland.	
		4:17
	<p style="text-align: right;">↳((mehrstimmiges Gelächter))</p> <p style="text-align: right;">↳Und jetzt, äh, ja, und jetzt</p>	
	kommt euer Gastgeber. Hier ist Beyaz!	
		4:46
	<p style="text-align: right;">↳((Gedrüll))</p> <p style="text-align: right;">↳((Signation der Beyaz-Show (Melodie von „Anton aus Tirol“ mit türkischem Text))</p> <p style="text-align: right;">↳((Gedrüll und Pfiffe)) ((Klatschen))</p>	
Bm:	↳Ehh ((spricht türkisch weiter))	
Cm:	↳Wir könnten	
	natürlich jetzt das übersetzen, lassn wa aba. Der Chef hat ja auch nichts verstanden. Zum Glück gibt es aber eines, was alle verstehen.	
Dm:	↳Ohhho ((singend))	
Cm:	↳Musik!	
Am+Dm:	5:01	
	ton::ight, baby, till I have you by my side, ((singend)) ba::by just can't wait until to-	
Cm:	↳Aba Stefan Raab wäre nicht Stefan Raab, wenn er sich nicht noch was Besonderes ausgedacht hätte,	
Am+Dm:	↳-night, baby	
Cm:	↳und so haben	
	Stefan und Max natürlich noch eine Überraschung mitgebracht	
Am+Dm:	↳((selbes Lied auf Türkisch gesungen))	



UT: 4:13–4:15 Stefan Raab aus Deutschland

Im Modus der Unsicherheitsbewältigung folgen mit der Begrüßung („Hallo“) und der Vorstellung („Hallo. Mein Name ist Stefan Raab. Ich komme aus Deutschland“) weitere standardisierte Handlungen, die im Alltagskontext oft beim erstmaligen Kontakt mit Fremden gesetzt werden. Aus der Reaktion des Publikums (Gelächter) ist jedoch erkennbar, dass diese Art der Vorstellung eine komische Note besitzt<sup>59</sup>. Der Bewältigungsversuch war erfolglos, das Thema der Unsicherheit ist nicht vom Tisch. Im weiteren Sprechen von Stefan Raab dokumentiert sich Irritation bei dem gleichzeitigen Versuch, einen neuen Bewältigungsansatz zu finden („Und jetzt, äh, ja, und jetzt kommt [...]“). Es folgt ein erneuter, dreimaliger Ansatz.

UT: 4:17–4:23 Der Gastgeber Beyaz

Erst der letzte Ansatz löst die Situation der Unsicherheit, indem Stefan Raab sein Wort an den türkischen Showmoderator Beyaz abgibt. Dieser bedient sich in türkischer Sprache eines ähnlichen Begrüßungsmodus wie Stefan Raab, verkörpert diesen aber mit weniger Unsicherheit. Im Gegensatz zur Vorstellung Stefan Raabs bleibt Beyaz' Redefluss kontinuierlich, ohne Pausen und Brüche. Das Publikum bleibt ruhig.

### **OT: 4:46–5:00 Verstehen**

UT: 4:46–4:50 Was auch der Chef nicht versteht, muss nicht übersetzt werden

An dieser Stelle setzt ein Kommentar aus dem Off ein: „Wir könnten natürlich jetzt das übersetzten, lassn wa aba. Der Chef hat ja auch nichts verstanden.“ Was vom türkischen Moderator Beyaz gesagt wurde, wird einfach nicht übersetzt. Die Begründung wird lediglich nachgeliefert. Darin dokumentieren sich zwei Bedeutungsgehalte. Mit der Begründung, der Chef habe ja auch nichts verstanden, wird einerseits Beyaz, dem *eigentlichen* Chef der Sendung, die Geltung entzogen. Andererseits fungiert diese Begründung auch als Rechtfertigung für den Entzug des Wortes sowie für die Handlungen, die bereits auf anderen Interpretationsebenen herausgearbeitet wurden: Der On-Ton der Show wird leiser gemacht und eigene, neue Anschlüsse für die Montage werden gesetzt. Nachdem Stefan Raab also aus Gründen der Unsicherheit in der fremden Situation sein Wort an Beyaz abgegeben hatte, wird es nun – stellvertretend für ihn –

---

<sup>59</sup> Eine ähnlich deplatzierte Äußerung macht Raab übrigens auch beim Interview mit der Deutsch-Türkin Nina Öger. Als diese ihm zum Abschluss für seine Bemühungen um Istanbul dankt, rutscht ihm die Floskel heraus: „Es lohnt sich wirklich hierher [nach Istanbul] zu kommen. Das kann ich ihnen nur empfehlen.“

Beyaz wieder entzogen und Raab damit eine erneute Bewältigungschance eröffnet: „Musik“.

UT: 4:51–5:00 Musik verstehen alle

Musik wird „von allen verstanden“ und damit als positiver Horizont zur Überwindung von Sprach- und Kulturgrenzen hervorgehoben. In diesem Rahmen singt Max seinen Songcontest-Song in englischer Sprache: „Can’t wait until tonight, baby“, den, so wird suggeriert, verstehe man in Deutschland *und* der Türkei.

### **OT: 5:01–5:40 Überraschungen**

UT: 5:01–5:10 Stefan Raabs besondere Überraschung

Nach den Worten des Kommentators ist Raab ein besonderer Mensch, der sich „Besonderes einfallen lässt“ (5:01) und damit die Leute überrascht. Das Besondere hat demnach mit Überraschung zu tun, ein Element, das sich in vielen Handlungen Stefan Raabs wiederfindet. Die besondere Überraschung besteht offensichtlich darin, das Lied auf Türkisch zu singen. Dies ist vor allem deshalb bemerkenswert, da zuvor der türkische Text von Beyaz weggeredet wurde, da „der Chef“ (Stefan Raab) diesen nicht verstanden habe. Auch an anderen Stellen von „Istanbul Total“ wird wenig Wert auf türkische Übersetzung gelegt. Beispiele sind etwa das Interview mit der türkischen Bauchtanzlehrerin oder die Äußerungen des Kapitäns des FeuerwehrschiFFs. An vielen anderen Stellen hingegen wird das Türkische explizit verballhornt oder bewusst als Kauderwelsch missverstanden, wie in den Sequenzen „Alidissimo“, „Elton vor der Beyaz-Show“ („Wadde-Hadde-du-de-da“, 562-00-05, 1:50) oder „FeuerwehrschiFF“ („Wir sind hier in einem Dance-Külübü“, 563-00-01). Das intendierte Missverstehen von Sprachen findet sich auch in der Stefan-Raab-Show „TV Total“ relativ häufig, etwa in der Sequenz „Lebensmittelmesse“. Dort wird auf Eigentümlichkeiten der französischen Sprache herumgeritten. So manches Mal beinhalten die Missverständnisse auch eine sexuelle Konnotation, beispielsweise in der Sequenz „Elton und Gülcan 1“ (561-00-03). Gülcan spricht auf Türkisch und Elton antwortet: „Ja das machma dann später im Hotel, ähm.“

UT: 5:31–5:36 Alle zusammen (Hep beraber)

Stefan Raab und Max singen den Text vom Blatt. Völlig frei skandiert Stefan Raab dann zweimal „Hep beraber!“ (auf Deutsch: „Alle zusammen!“). Wie bereits an früherer Stelle dargelegt, dokumentiert sich darin die Nachahmung eines standardisierten Ausrufs mit

der Erfolgsgarantie, das türkische Publikum zu erreichen. Dieses antwortet erwartungsgemäß mit Applaus, was in der abschließenden Rahmung des Off-Kommentators als umfassende Zustimmung zu Raab und seiner Show zu werten ist.

UT: 5:36–5:40 Dank Beyaz sind die Türken auf unserer Seite

Der Off-Text signalisiert eine „geglückte“ Verständigung und dokumentiert den Anspruch von „Istanbul Total“: „Damit haben wir die Türken wohl endgültig auf unserer Seite.“ Das Gemeinsame zwischen Deutschen und Türken wird also bestenfalls durch eine Verbündung gegen Dritte realisiert, wobei die Türken dabei eindeutig instrumentalisiert werden. Letztlich impliziert die getätigte Aussage ja auch, dass man sich selber nicht auf die Seite der Türken schlagen kann. Dies korrespondiert mit der erklärten Zielsetzung von „Istanbul Total“ – 12 Songcontest-Punkte für Max –, die gebetsmühlenartig in verschiedenen Kontexten der Show hervorgehoben wird.

Der abschließende Ausruf aus dem Off, „Danke, Beyaz!“, besitzt im Kontext von „Hep beraber!“ eine freundschaftliche Note. Vor dem Hintergrund dessen, dass Beyaz zuvor das Wort und die Kontrolle über seine eigene Show entzogen wurde, erhält diese Aussage zynische Züge. Der Dank wird Beyaz letztlich dafür ausgesprochen, dass er Stefan Raab und Max seine Bühne für ihr eigenes Programm überlassen hat. Letztlich reichte das Interesse an ihm und seiner Show nur so weit, wie er für die Vermarktung von Raab, Max und „Istanbul Total“ von Nutzen war.

## **6.5 Reflektierende Gesamtinterpretation**

Am zweiten Sendetag von „Istanbul Total“ wird die Beyaz-Show als türkisches Pendant zur Sendung „TV Total“ des Moderators Stefan Raab präsentiert. Dieser 1:1-Vergleich wird dadurch nahegelegt, dass die Beyaz-Show im Wesentlichen nichts anderes als eine „abgekupferte“ Version von „TV Total“ sei. Die Vorgehensweise der Stereotypisierung der Türken als unbeholfene Nachahmer wurde in dieser Arbeit sowie von Bohnsack (2009) bereits mehrfach erwähnt. Die Rahmung der Beyaz-Show fügt sich nahtlos in diesen Kontext ein.

In der Beyaz-Show dokumentieren sich jedoch auch Nachahmungsleistungen des deutschen Moderators Raab, die legitimiert werden müssen. Dies gelingt dadurch, dass diese Nachahmungen entweder ironisierend verkleidet oder – entsprechend der öffentlichen PR von „Istanbul Total“ – als „kulturverbindender Brückenschlag“ dargestellt werden. Die Präsentation von Stefan Raab als „interkulturellen

Brückenbauer“ immunisiert ihn gegen die unmittelbare Zuschreibung egoistischer Motive, wie sie durch eine „Hyperzentrierung“ (Bohnsack, 2009, S. 234) auf seine Person entstehen könnten. Bei genauerer Betrachtung dokumentiert sich dennoch die unilaterale Ausnutzung der Beyaz-Show durch „Istanbul Total“ auf sämtlichen Interpretationsebenen.

### **6.5.1 Der Gebrauchswert von Faksimiles und Pendants**

Im Rahmen der Beyaz-Show kommt dem türkischen Stereotyp des Nachahmens und Fälschens eine produktive Bedeutung zu. Es eignet sich perfekt zur Verdoppelung medialer Ereignisse, was dem „parasitären Aufmerksamkeitsgewinn“ von „Istanbul Total“ (Bohnsack, 2009, S. 238) zuträglich ist. Nachdem die Beyaz-Show als türkische Kopie der Raab-Show entlarvt ist, liegt offensichtlich nichts näher, als Stefan Raab an dieser Show teilnehmen zu lassen und Beyaz zu „Istanbul Total“ einzuladen<sup>60</sup>.

In diesem Wechselspiel kann selbst Raab sich „kleine Gaunereien“ erlauben (sonst Stereotyp der Türken), die zur kulturellen Verständigung beitragen. Der nachgemachte Beyaz-Move etwa sowie die türkischen Worthülsen (z. B. „Hep beraber!“), die Stefan Raab beim Singen einsetzt, generieren jedenfalls die gewünschte Zustimmung des Publikums. Wie gleich zu Beginn der Sendung ersichtlich ist, scheitert Raab mit anderen Kommunikationsbemühungen kläglich. Im Orientierungsrahmen von „Istanbul Total“ stellt die Nachahmung somit die einzige funktionale Möglichkeit zur interkulturellen Kommunikation dar.

### **6.5.2 Instrumentalisierung des Türkischen zur Selbstinszenierung**

Je weiter die Analyse der Sendung „Istanbul Total“ voranschreitet, desto deutlicher wird, dass die „türkische Innenwelt“ (Bohnsack, 2009, S. 237) nicht von Interesse ist. Das Türkische entfaltet seine Relevanz lediglich in Kontexten, die der Selbstinszenierung des Moderators Stefan Raab bzw. der Vermarktung seiner Person sowie der Show insgesamt dienlich sind. Dieser Umstand dokumentiert sich in der Beyaz-Show auf zwei Arten: einerseits im offenkundigen Desinteresse am Türkischen auf der Ton- und Textebene. In Minute 4:46 werden die Äußerungen des türkischen Moderators einfach unübersetzt mit der Begründung ausgeblendet: Was der Chef (d. h. Stefan Raab) nicht versteht, muss auch nicht übersetzt werden. Andererseits bedient sich gerade dieser „Chef“ der

---

<sup>60</sup> Beyaz tritt seinerseits am 11. Mai 2004 in der Sendung „Istanbul Total“ auf. Link: <http://tvtotal.prosieben.de/tvtotal/videos/player/?contentId=14156>

türkischen Sprache, um in seiner Gesangseinlage mit Max Aufmerksamkeits- und Popularitätsgewinne zu erreichen.

### **6.5.3 Die Immunisierung des Moderators von „Istanbul Total“**

Bei der Betrachtung des Videomaterials der Beyaz-Show wird deutlich, dass das in „Istanbul Total“ aufgeworfene Stereotyp der Türken als ungeschickte Akteure brüchig zu werden droht. Stefan Raab verhält sich zu Beginn der Show überaus unbeholfen (vgl. Fotogramm 1), während der türkische Moderator Beyaz selbstsicher die Sendung leitet (vgl. Fotogramme 2 u. 3). Allein auf der Basis der Leistungen der abgebildeten Bildproduzent/inn/en wäre dieses Videomaterial folglich nur bedingt für den Einsatz in „Istanbul Total“ geeignet, da es im Widerspruch zu den dort wirksamen Orientierungen steht; insbesondere der „Inszenierung von Spontaneität“ und der „Hyperzentrierung auf die Person des Moderators“ Stefan Raab (Bohnsack, 2009, S. 234).

Als letzte Möglichkeit zur Aufrechterhaltung der bestehenden Showkonzepte bleibt nur die Postproduktion, d. h. der Rückgriff auf Leistungen der abbildenden Bildproduzent/inn/en. Dazu wird das fertige Videomaterial der Beyaz-Show einer nochmaligen Montage durch das Produktionsteam von „Istanbul Total“ unterzogen. Resultat: Beyaz wird über den Off-Text die Kontrolle über seine Show entzogen und diese im Sinne von „Istanbul Total“ „solide“ zu Ende geführt. „Istanbul Total“ hat somit über die Montage im Nachhinein die Möglichkeit, Medienereignisse für sich zu nutzen und zu vereinnahmen sowie sogar den Moderator Stefan Raab gegen eigene Fehlleistungen zu immunisieren.

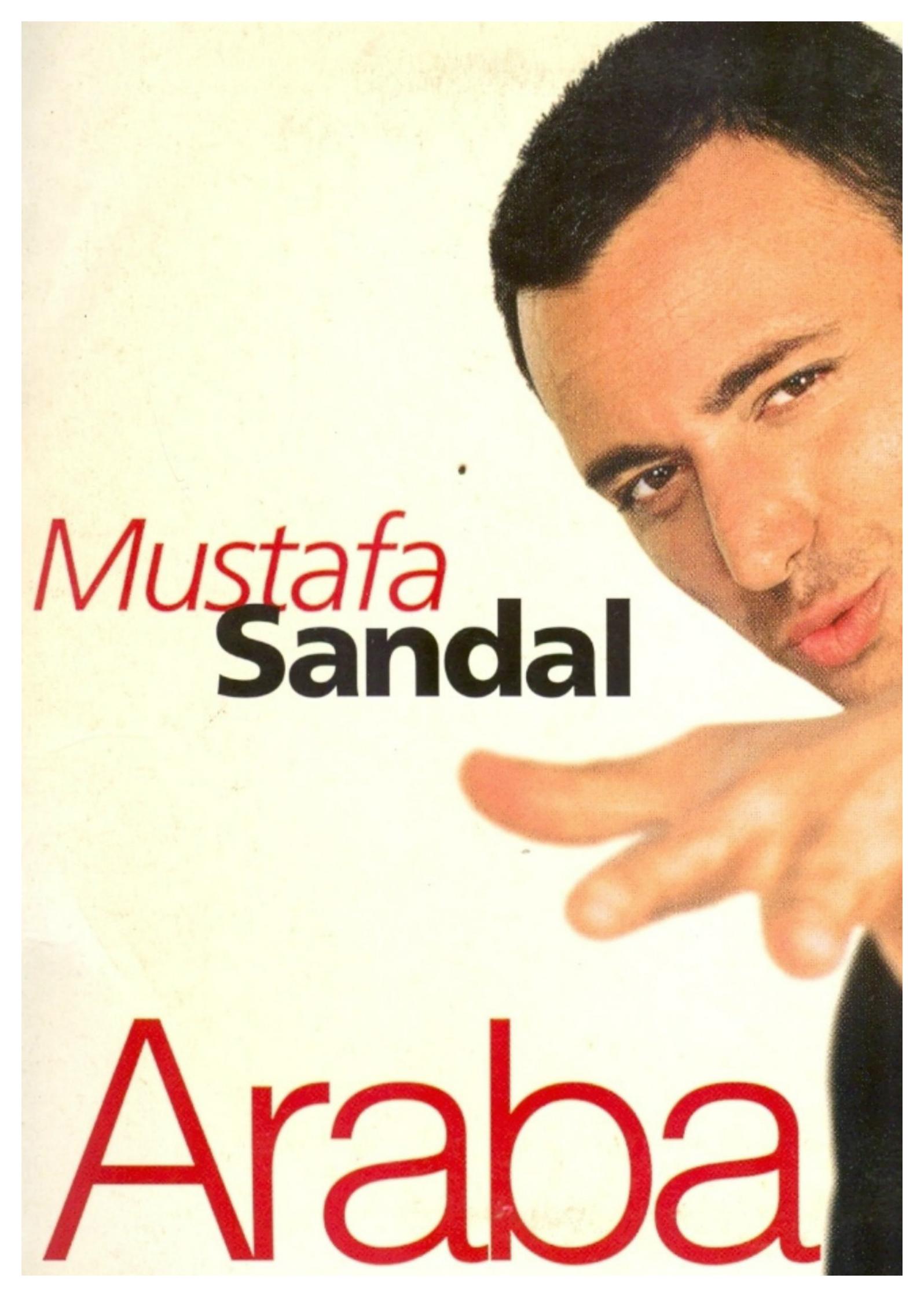
### **6.5.4 Monostrukturierung in „Istanbul Total“ und der Beyaz-Show**

Zum Prinzip der „Monostrukturierung durch den Showmaster“ (Bohnsack, 2009, S. 240) können nach Durchführung der komparativen Analyse zwei aufschlussreiche Punkte festgehalten werden. Zum einen konnte anhand der Leistungen der abgebildeten und abbildenden Bildproduzent/inn/en der türkischen Beyaz-Show gezeigt werden, dass die Monostrukturierung auch in anderen Fernsehsendungen wirksam ist – alleine der Titel „Beyaz-Show“ ist ein deutlicher Beleg dafür.

Zum anderen konnte anhand des vorliegenden Videomaterials gezeigt werden, wie auf Bedrohungen der „Monostrukturierung durch den Showmaster“ reagiert wird. Im vorliegenden Falle erfolgte die Gefährdung durch die unzureichenden Leistungen des abgebildeten Bildproduzenten Stefan Raab sowie durch den offenkundigen

Interessenskonflikt divergierender abbildender Bildproduzent/inn/en („Istanbul Total“ und Beyaz-Show). Was die Struktur der Machtverhältnisse anbelangt, so wird im Falle der Videointerpretation der Beyaz-Show deutlich, wessen Rahmungen sich letztlich durchsetzen, wenn mehrere abbildende Bildproduzent/inn/en involviert sind: Es sind die Rahmungen derer, die die Entscheidungsgewalt über die Endmontage besitzen – ungeachtet dessen, welche Leistungen durch die abgebildeten Bildproduzent/inn/en ursprünglich erbracht wurden.





*Mustafa*  
**Sandal**

**Araba**

**Abbildung 43: Schallplattencover der Single „Araba“ (1996), Sony Music France**

Quelle: <http://wikibit.net/album/Mustafa-Sandal/Araba>

## 7 Zur Videointerpretation von Musikvideos

Dieses und die beiden nachfolgenden Kapitel sind auf Musikvideos als audiovisuelle Produkte der Massenkommunikation gerichtet. Die jährlichen Medienberichte der Rundfunkanstalten (z. B. MPFS, 2012) legen Zeugnis davon ab, dass Jugendliche und Migrant/inn/en zu enthusiastischsten Anhänger/inne/n von Musikvideos gehören. Diese Gruppen lassen sich in der Terminologie des Medienpsychologen Vitouch (2007) auch als „Vielseher/innen“ bezeichnen. Durch die starke Verbindung von Mediennutzung und Alltagspraxis gelangen Musikvideos in die Position die biografischen Ereignisse und Prozesse ihrer Rezipient/inn/en zu moderieren und für sie niederschwellige Identifikations- und Entwicklungsräume zu eröffnen. Unter Berücksichtigung dieser medialen Dimensionen übersteigen Musikvideos bei weitem die ihnen oft verkürzt zugeschriebene Funktionalität einer bloßen visuellen Unterstützung oder Begleitung von Musik.

Während die historischen Wurzeln wechselseitiger kompositorischer Steigerungen von Film und Musik mehr als ein Jahrhundert bis zu den Anfängen des Kinos zurückverfolgt werden können, markieren die Einführung von im Fernsehen übertragenen Musiksendungen in den 1970ern und – noch bedeutsamer – die in den frühen 1980ern stattfindende Etablierung dezidierter Musik-Fernsehsender (z. B. MTV) wichtige Meilensteine in der Massenkommersialisierung von Musikvideos als eigenes Genre (Austerlitz, 2008). Von diesem Punkt an wurde die Produktion von Musikvideos zum wesentlichen und unverzichtbaren Bestandteil der Marketing-Mixstrategien von Musikfirmen, um maximale Zuhörerschaften zu erreichen und Lieder an ein breites Publikum zu verkaufen. Die Wirksamkeit dieser Strategie wird vom jährlichen Bericht der International Federation of the Phonographic Industry (IFPI) unterstützt. Laut IFPI wurden die Musikvideos des bisher einschaltequotenstärksten Popsängers Justin Bieber „mehr als zwei Milliarden mal angesehen“ (IFPI, 2012, S. 13). Ende 2012 stellte der koreanische Sänger Psy einen neuen Rekord für ein einzelnes Musikvideo auf. Sein internationaler Hit „Gangnam Style“ wurde über eine Milliarde Mal angesehen<sup>61</sup> (MTV, 2012). Gemäß der deutschen JIM-Studie<sup>62</sup>, einer jährlichen Umfrage zur Mediennutzung von Jugendlichen schauen 74 % der 12- bis 19-Jährigen regelmäßig Musikvideos an

---

<sup>61</sup> Zum 15. April 2013 haben laut Youtube bereits über 1,5 Milliarden Abrufe des Musikvideos „Gangnam Style“ stattgefunden. Link zum Video: <http://www.youtube.com/watch?v=9bZkp7q19f0>

<sup>62</sup> n=1201 Befragte

(MPFS, 2012). Eine zweite repräsentative Studie<sup>63</sup>, die 2007 von den deutschen Rundfunkunternehmen ARD und ZDF durchgeführt wurde, zeigt, dass Musikvideos für 77 % der in Deutschland lebenden Migrant/inn/en im Alter zwischen 14 und 29 Jahren zu den beliebtesten Fernsehübertragungen gehören (ARD/ZDF, 2007, S. 36; Walter, Schlinker & Fischer, 2007, S. 447). Wenn Menschen bestimmte Medien häufig in Anspruch nehmen, impliziert dies, dass diese Medien für ihre Nutzer/innen von großer Bedeutung sind. Häufige Mediennutzung muss praktisch in alltägliche Lebensroutinen integriert werden. Somit sind Medien nicht nur Teil der menschlichen Kommunikation, sondern selbst als eigenständige Kommunikationsmittel anzusehen. Die Idee der Kommunikation durch Musikvideos trägt der Tatsache Rechnung, dass Menschen ihre alltäglichen Erfahrungen, Träume oder Zukunftserwartungen artikulieren können, indem sie auf spezifische Videos Bezug nehmen, die von ihren Gemeinschaften geteilt werden und leicht zugänglich sind. Musikvideos stellen dabei insofern besondere Bedeutungsträger dar, als sie sich von ihrem ursprünglichen Vermittlungsmedium – dem Fernsehen – ablösen und mit scheinbarer Leichtigkeit den Übergang ins Internet vollziehen konnten<sup>64</sup>. Dieser Übergang ist anderen Medienformen, etwa Fernsehshows oder Nachrichtensendungen, bislang nicht in entsprechender Weise gelungen.

In ihrer Funktion als Kommunikationsmittel sind Musikvideos eng mit der persönlichen Entwicklung ihrer Benutzer/innen verbunden. Von daher überrascht es wenig, dass im Rahmen von Musikvideos häufig eine inhaltliche Auseinandersetzung mit Entwicklungsherausforderungen stattfindet. Eines der relevantesten – und damit auch eines der in Musikvideos am häufigsten dargestellten – Themen für Jugend- und Migrationsgruppen ist das Auf und Ab romantischer Beziehungen. Romantische Beziehungen können der Grund sein, in ein anderes Land zu ziehen, eine fremde Sprache zu lernen, neue Bücher zu lesen, den Freundeskreis zu verändern etc. Aus einer entwicklungsorientierten Perspektive spielen romantische Beziehungen daher eine signifikante Rolle im Auftreten von biografischen Brüchen und deren „Reparatur“. Damit Musikvideos erfolgreich sind, müssen sie die persönliche Erfahrung dieser Spannungszustände aktiv bewirtschaften bzw. kultivieren.

Für die nachfolgende empirische Betrachtung habe ich zwei Musikvideos ausgewählt, die auf der Basis desselben Liebesliedes produziert wurden: „Araba“ war

---

<sup>63</sup> n=3010 Befragte mit Migrationshintergrund, älter als 14 Jahre

<sup>64</sup> Laut MPFS (2012) ist Fernsehen nicht länger das alleinige Medium zur Veröffentlichung von Musikvideos. Obwohl zwei Drittel der Musikvideos noch immer online *und* im Fernsehen angesehen werden, wird mittlerweile ein Drittel ausschließlich im Internet angeschaut.

1996 einer der größten Erfolge des türkischen Popsängers Mustafa Sandal. 2004 wurde das Lied für den internationalen Markt neu verfilmt. Welche Arten von biografischen Brüchen und „Reparatur“-Formen werden in den beiden Musikvideo-Versionen visuell dargestellt? Zur Beantwortung dieser Forschungsfrage werden zunächst einige Überlegungen zur begrifflichen Bezeichnung des Fallmaterials präsentiert, bevor die visuelle Einführung in das Material mit Hilfe von Videotranskripten erfolgt. Im Zusammenhang mit der Erläuterung der zentralen methodologischen Konzepte und Begrifflichkeiten wird dann ein Überblick über das Verfahren der Videointerpretation nach der dokumentarischen Methode, mit Schwerpunktsetzung auf die Rekonstruktion von Farben und Farbkontrasten, gegeben. Abschließend erfolgt die Zusammenfassung und Diskussion der Ergebnisse der empirischen Videointerpretationen hinsichtlich biografischer Brüche und deren „Reparatur“.

### **7.1 Zwei Versionen von „Araba“**

Zur komparativen Analyse zweier Videoverversionen eines Popsongs ist es zuerst erforderlich, diese Versionen zu benennen, um eine begriffliche Bezugnahme zu ermöglichen. Nachdem die beiden Musikvideos jeweils im türkischen und im deutschen Musikfernsehen ausgestrahlt wurden, wäre es möglich, sie einfach „das türkische“ und „das deutsche“ Video zu nennen. Mit dieser Idee wären allerdings mehrere Nachteile verbunden. Erstens sind „Türkisch“ und „Deutsch“ sehr vielschichtige und weitreichende Konzepte, die eine große Bandbreite an Aspekten in sich bergen: In der Umgangssprache können „das Türkische“ und „das Deutsche“ beispielsweise auf Nation, Kultur, Sprache, Mentalität, Moral u. a. verweisen. Zweitens unterscheiden die Bezeichnungen „Türkisch“ und „Deutsch“ nicht zwischen Produzent/inn/en, Schauspieler/inne/n und Videoszenen. So ist zum Beispiel in beiden Fällen des Videomaterials der Sänger (Mustafa Sandal) türkischer Herkunft (bezogen auf die Staatsangehörigkeit bei seiner Geburt) und singt auch in türkischer Sprache. Die Produzenten der Videos (Sony Music France und Polydor) sind hingegen international operierende Produktionsgemeinschaften<sup>65</sup>. Eine weitere Möglichkeit der Benennung wäre damit: „die Sony Version“ und „die Polydor Version“. In diesem Zusammenhang könnten die beiden Videoverversionen auch durch die Einbeziehung des Jahres ihrer Veröffentlichung unterschieden werden. Denn die Sony Version wurde 1996 produziert, die Polydor Version 2004. Die Verwendung von Titeln

---

<sup>65</sup> Die erste Version von „Araba“ wurde 1996 von Sony Music France, die zweite 2004 von Polydor Germany produziert. Sony Music France gehört zu Sony Music Entertainment, Polydor zur Universal Music Group. Beide Unternehmen haben ihren Hauptsitz in New York City, USA.

wie „Sony 1996“ und „Polydor 2004“ für die zwei Videoverionen könnte möglicherweise aber zu abstrakt erscheinen und den Lesefluss stören. Abgesehen davon würde die Unterscheidung anhand von Jahreszahlen und Firmennamen möglicherweise auch vom eigentlichen Gehalt der Gegenüberstellung ablenken. Dasselbe gilt für die Option, die zwei Videos als „die erste“ und „die zweite“ oder „die alte“ und „die neue“ Version zu bezeichnen. Auch derartige Bezeichnungen wären insofern irreführend, als sie eine bestimmte Reihenfolge des Materials implizieren und damit vom inhaltlichen Fokus der Untersuchung ablenken würden. Bei dem Versuch, einen Mittelweg zwischen begrifflich implizit vorbelasteten oder voreingenommenen Bezeichnungen, aber auch zu allgemeingültigen oder vagen Benennungen zu finden, möchte ich vorschlagen die zwei Videos in Bezug auf die Märkte zu unterscheiden, für welche ihre Veröffentlichungen bestimmt waren. Märkte sind Orte, an denen verschiedene Parteien Austausch treiben. Um Absatz zu generieren, müssen Musikgesellschaften marktnah produzieren und ein praktisches Gespür dafür entwickeln, was das Publikum kaufen wird. Daher fokussiert der Begriff „Markt“ im soziokulturellen Sinn die Beziehungen zwischen Produzent/inn/en und Konsument/inn/en von Musikvideos. Zudem ist das Konzept des Marktes ein dynamisches, welches in einem kultur- und sozialgeschichtlichen Sinne vergangene Erfahrungen und Zukunftserwartungen von Marktbeteiligten bestimmter Medienumfelder mit einschließt. Dieser Ansatz passt damit gut zum medientheoretischen Charakter dieser Arbeit.

Zum Abschluss dieser Diskussion möchte ich konstatieren, dass jede Methode zur Betitelung der Musikvideos Vor- und Nachteile hat. Dennoch gibt es Begrifflichkeiten, die zum Verständnis des Fallmaterials beitragen können, noch bevor dieses interpretiert wird. Die anfänglichen Labels „Türkisch“ und „Deutsch“ oder „Araba 1996“ und „Araba 2004“ wurden aufgrund der ihnen anhaftenden impliziten Polarisierung verworfen. Derartige Konzepte neigen dazu, ein isoliertes Verständnis der beiden Videos zu begünstigen. Dabei sind diese alles andere als unabhängig voneinander. Wenngleich sich das zweite Video optisch stark vom ersten unterscheidet, so konnte es dennoch nur in Referenz zum ersten produziert werden. Indem ich die zwei Videos letzten Endes als „die Version für den internationalen Markt“ und „die Version für den türkischen Markt“ bezeichne, liegt nun die Betonung ausdrücklich auf den Gemeinsamkeiten und Dynamiken der beiden Versionen. Wenngleich das erste Video aus einer türkischen Sphäre (was auch immer diese konkret einschließt) hervorgegangen sein mag, ist das zweite Video nicht einfach nur als „internationaler“, vollkommen von der türkischen

Version unabhängiger Gegenpart anzusehen. Mit Sloterdijk (2004) in Bezug auf Latour (2001) kann die zweite Version vielmehr als „Entfaltung“ bzw. „Artikulation“ der ersten angesehen werden. Acht Jahre später produziert, schließt sie ein weiteres Publikum mit ein, das sich sowohl aus türkischstämmigen als auch aus Zusehenden anderer Länder zusammensetzen kann.<sup>66</sup>

## **7.2 Die medialen Eigenschaften von Musikvideos**

Dieses Kapitel bedient sich der empirischen Instrumente, die Bohnsack (2009) im Rahmen der Bild- und Videointerpretation nach der dokumentarischen Methode entwickelt hat, und erweitert diese entsprechend der besonderen Materialerfordernisse des Mediums „Musikvideo“. Dabei ist das Konzept einer medientheoretisch inspirierten Kulturpsychologie, welches von Slunecko (2008) eingeführt wurde, von signifikanter Bedeutung. Diese psychologische Position, die mit der psychologischen Tradition bricht, menschliches Erleben und Verhalten unabhängig von Kultur und Medien zu beforschen, wurde von Slunecko, Przyborski und Ruck theoretisch weiter ausgearbeitet (Przyborski & Slunecko, 2009, 2011; Ruck & Slunecko, 2008; Slunecko & Przyborski, 2009). Psycholog/inn/en sollten demnach Kulturen und Medien als und mit den Menschen stets verbundene Agenten (an)erkennen. Im Hinblick auf die Forschungspraxis bedeutet dies, dass neue Wege gefunden und Instrumente entwickelt werden müssen, um die nun erkennbar gewordenen Mensch-Kulturen-Medien-Phänomene empirisch zu erfassen.<sup>67</sup> In dieser Hinsicht nehme ich hier somit neben einer methodologischen auch eine metatheoretische Position ein, in welcher der gegenseitigen Beeinflussung von Menschen, Kulturen und Medien eine besondere Aufmerksamkeit geschenkt wird.

Einer von Bohnsacks Hauptbeiträgen zur qualitativen Sozialforschung kann darin gesehen werden, dass er das Feld der Kunsttheorie (Imdahl, 1996a; Panofsky, 2006) für die empirische Wissenschaft zugänglich gemacht hat. Des Weiteren ist es ihm gelungen, etablierte Konzepte der Kunsttheorie – z. B. Panofskys „Ikonografie und Ikonologie“ (2006) und Imdahls „Ikonik“ (1996a) – als empirische Instrumente der qualitativen Bild- und Videointerpretation nutzbar zu machen. Diesem Beispiel folgend, bin ich selbst im Rahmen der weiterführenden Auseinandersetzung mit Kunstliteratur auf die

---

<sup>66</sup> Diese Betrachtungsweise lässt sich mit einer Beobachtung verbinden, die ich machte, als ich beide Videoverionen Personen türkischer Herkunft zeigte. Ich hatte angenommen, dass Türken die „türkische“ Version bevorzugen würden. Es zeigte sich jedoch, dass ihnen die „internationale“ Version besser gefiel.

<sup>67</sup> Dieser Gedanke wird auch in Latours (2001) Konzept „Parlament der Dinge“ herausgearbeitet.

Arbeiten von Itten (2000, 2010) zum Thema Farben gestoßen. Paradoxerweise sind Farben ein komplexes Untersuchungsgebiet, obwohl sie nach der Wahrnehmung einer/s jeden unmittelbar erfasst werden können. Ittens Überlegungen nach existieren Farben nicht in isolierter Weise, sondern ändern ihre Beschaffenheit stets in Bezug zu anderen Farben, die sie umgeben bzw. ihnen vorausgehen oder nachfolgen. In Anbetracht dieser spezifischen Phänomenologie schlägt Itten (2000, 2010) vor, sich von der isolierten Interpretation von Farben hin zu einer Interpretation von *Farbkontrasten* zu bewegen. Dieses Konzept von miteinander in Beziehung stehenden Farben ist sowohl mit der zuvor skizzierten theoretischen Orientierung an einer medientheoretischen Kulturpsychologie als auch mit der Methodologie der dokumentarischen Methode gut zu vereinbaren. Nach Aussage von Bohnsack (2009) verdient die Untersuchung sinngenetischer Relationierungen bei der Erforschung von Film und Video besondere Beachtung: Während die Interpretation von Fotogrammen – als (repräsentative) Standbilder von Einstellungen – auf die Beziehungen der Bedeutungsebenen *innerhalb* des Bildes abstellt, kann der Montageprozess als deren nochmalige Metarelationierung verstanden werden.

Was das Konzept der dokumentarischen Methode von anderen empirischen Zugängen unterscheidet, ist sein mediensensibles Design, welches die Idee in Betracht zieht, dass Bilder „selbstreferentielle Systeme“ (Bohnsack, 2009, S. 13) sind, die nicht direkt in Text übersetzt werden können – ein Umstand, der laut Bohnsack (2009) von den textorientierten Sozialwissenschaften bisher vernachlässigt wurde. Während Texte aber nur durch sequenzielle Wortidentifizierung verstanden werden können, arbeiten Bilder in einem simultanen Modus und auf vorverbalen und präkonzeptuellen Ebenen. Imdahl (1996a) fügt dem noch seine analytische Unterscheidung von *wiedererkennendem Sehen* und *sehendem Sehen* hinzu. Ersteres bezieht sich auf den gewöhnlichen Kommunikationsmodus von Wahrnehmung, der auf die Identifizierung von Objekten, Menschen oder spezifischen Situationen gründet; Letzteres impliziert die sofortige formale Wahrnehmung eines Bildes in Bezug auf Umrisse, Farben, Größe, Gewicht und Richtungen. Während das Sehvermögen auch durch Mittel der Textinterpretation zugänglich ist, kann die spezifische Medienqualität von Bildern nur durch die empirische Rekonstruktion des sehenden Sehens offengelegt werden. Hinsichtlich differenzieller Medienanalysen berücksichtigt die dokumentarische Methode beide Interpretationsebenen. Diese Position geht auch mit McLuhans Hinweis (2003) konform, dass der Inhalt eines jeglichen Mediums nur durch die Untersuchung

seiner formalen Charakteristiken verstanden werden kann. Medien (re-)konfigurieren das menschliche Leben ständig – sie haben die Macht, bestimmte Aspekte zu erweitern und gleichzeitig andere zu reduzieren. Studien zu den Effekten von Mediennutzung zufolge ist zu beachten, dass Menschen sich zwar der Erweiterungen bewusst sind, der Einschränkungen aber nicht gewahr werden.<sup>68</sup> Das führt zu der Beobachtung, dass ein Medium immer dann aus dem Bewusstsein der Menschen verschwindet, sobald es erfolgreich in alltägliche Lebenspraktiken eingebunden ist. Latour (1998, S. 41) nennt diesen Prozess „reversibles Blackboxing“<sup>69</sup>; Sluneko und Przyborski (2009) sprechen von einem „Unsichtbar-Werden“. Im Hinblick darauf, dass Musikvideos gerne angesehen werden, bedeutet dies, dass sie ebenfalls die Erfahrung ihrer Zuschauer/innen physisch beeinflussen, und zwar unmittelbarer als die Begriffe *Fernsehen* und *Internet* vermuten lassen<sup>70</sup>. Beide Medienformen sind mittlerweile symbiotisch mit den Menschen verwachsen und zu Extensionen des physischen Organismus geworden.

### 7.3 Methodologische Herausforderungen an die Video- und Filminterpretation

Die Materialgattung *Musikvideo* stellt die Video- und Filminterpretation nach der dokumentarischen Methode (Bohnsack, 2009) vor methodologische Herausforderungen und zeigt zugleich deren bedeutsame Entwicklungspotentiale auf. Als hochkondensierte Dokumente der visuellen Massenkommunikation sind Musikvideos u. a. durch rasante Schnittfolgen und stark ineinander verschachtelte Sequenzen charakterisiert. Das gängige Prozedere der Arbeitsschritte der Film- und Videointerpretation hat sich bisher als überaus flexibles Instrument zur empirischen Rekonstruktion verschiedenster Fallmaterialien bewährt – TV-Shows, Amateur-, Unterricht-, Reisevideos, Werbespots etc. (etwa: Baltruschat, 2010; Bohnsack, 2009; Hampl, 2010; Sobotka, 2009). Aufgrund der formalen Eigensinnigkeiten auf der Ebene des Videoschnitts bzw. der *Montage* sperren sich Musikvideos aber streckenweise dagegen. Dies macht sich u. a. dadurch bemerkbar, dass zum Teil die Auswahl *repräsentativer* Sequenzen und Fotogramme misslingt.

---

<sup>68</sup> Kleidung zum Beispiel bietet dem Körper nicht nur Schutz gegen Kälte, sie verdeckt auch die Haut vor der öffentlichen Sphäre und privatisiert sie damit (McLuhan, 2003, S. 161). Das ganze soziale Konzept von „Nacktheit“ (und allem, was mit ihr verbunden ist) wäre ohne die täglichen Zeremonien des An- und Ausziehens nicht verständlich.

<sup>69</sup> Vgl. dazu auch Schäffer (2001, S. 55).

<sup>70</sup> In dieser Hinsicht mag es angemessener sein, Fernsehen als *Anchivision* („Nahsehen“) zu bezeichnen und das Internet als *Infranet* („Unter-die-Haut-gehendes Netz“).

Die empirische Auseinandersetzung mit Fallmaterial der Gattung Musikvideos macht damit auf offene Fragen der Methodenentwicklung aufmerksam. Bislang lässt sich die Montage über die (externe) *Relationierung der Einstellungen* (Bohnsack, 2009, S. 164 und S. 208) im Sinne des *Einstellungswechsels* rekonstruieren, was die Auswahl relevanter Sequenzen und Einstellungen sowie die vollständige Interpretation repräsentativer Fotogramme voraussetzt. Am Fallmaterial von Musikvideos wird sichtbar, wie dichte Montagemuster – die das komplexe Resultat vielschichtiger Kompositionsleistungen seitens der abbildenden Bildproduzent/inn/en darstellen – das Prozedere der bisher üblichen Arbeitsschritte (Bohnsack, 2009, S. 176) unterlaufen können. Zudem stellt sich die Frage, ob nicht direktere Zugänge zur Rekonstruktion der Montage möglich wären.

In der Folge möchte ich nun am Beispiel eines konkreten Musikvideos versuchen meine diesbezüglichen Beobachtungen zu erläutern. Dabei halte ich mich aus Gründen der Nachvollziehbarkeit grundsätzlich an jenen Ablauf der Arbeitsschritte, welcher Bohnsack (2009) für die dokumentarische Video- und Filminterpretation vorgestellt hat. Nur in Einzelfällen, die sich durch die besondere Struktur des Untersuchungsgegenstandes ergeben und die gut begründbar sind, weiche ich davon ab. Die in diesem Zusammenhang erforderlichen Forschungsentscheidungen und Überlegungen beruhen jedenfalls auf den metatheoretischen Grundannahmen der dokumentarischen Methode (Bohnsack, 2007, 2009).

### **7.3.1 Die Rekonstruktion der Bilddimension von Musikvideos**

Einschränkend sei an dieser Stelle erwähnt, dass der Fokus der anschließenden Videointerpretation auf der Rekonstruktion der Bilddimension liegt und dabei nicht genauer auf die Ton- bzw. Textdimension eingegangen wird. Diese Vorgangsweise mag gerade für die Betrachtung von Musikvideos unplausibel erscheinen. In Hinblick auf die Leistungsfähigkeit der dokumentarischen Video- bzw. Filminterpretation interessiert mich jedoch, wie weit die Instrumente zur Rekonstruktionen in der Bilddimension tragfähig sind. Aus methodologischen Gründen ist die Einklammerung der Tondimension erforderlich, um diese Brüche aufzuspüren. In diesem Sinne werde ich just an jenen Stellen auf die Tondimension verweisen, wo die reine Rekonstruktion der Bilddimension nicht zum Gesamtverständnis des Musikvideos ausreicht.

### 7.3.2 Die Produktionsverhältnisse von Musikvideos und Fernsehshows

Musikvideos sind in noch weit stärkerem Maße als Fernsehshows durch die Leistungen der abbildenden Bildproduzent/inn/en geprägt. Im Falle von Fernsehshow haben die abgebildeten Bildproduzent/inn/en (z. B. Moderator/inn/en und Interviewpartner/innen) – u. a. bedingt durch den Livecharakter der Aufnahmen – noch gewisse situative Handlungsspielräume, die in Musikvideos fehlen. Des Weiteren beschränkt sich im Falle von TV-Sendungen der Schnitt, wie Bohnsack (2009) erläutert, in der Regel auf die sogenannte Bildmischung. Dabei wird der zeitliche Ablauf der Aufnahmen beibehalten. Die Montagetätigkeit im Rahmen der Bildmischung besteht damit vornehmlich im Hin- und Herschalten zwischen simultan laufenden Kameras. Auch die Postproduktion, d. h. sämtliche Leistungen der abbildenden Bildproduzent/inn/en, die *nach* der Aufnahme erbracht werden, ist meist auf das Hinzufügen des Fernsehlogos und etwaiger (Namens-)Einblendungen beschränkt<sup>71</sup>. Demgegenüber sind in Musikvideos die Leistungen der abbildenden und abgebildeten Bildproduzent/inn/en vor, während und nach der Aufnahme durch die abbildenden<sup>72</sup> Bildproduzent/inn/en einer unausweichlichen und vollständigen Kontrolle unterworfen<sup>73</sup>. Im Rahmen der Videoproduktion kommt der Montage daher eine Schlüsselrolle zu. Nicht nur werden die Aufnahmen in bestimmten Sequenzen aneinandergeschnitten – das Material wird auf der Bild- und Tonebene auch noch mehrfach nachbearbeitet und korrigiert.

### 7.3.3 Farbkontraste, Montagemuster und Räumlichkeit

In Abkehr von einem isolierten Farbverständnis entwickelte der Maler, Kunsttheoretiker und Kunstpädagoge Johannes Itten (2000, 2010) seine „Theorie der sieben Farbkontraste“. Hinsichtlich der empirischen Untersuchung von Filmen und Videos haben Farbkontraste die bemerkenswerte Eigenschaft, dass sie sich sowohl innerhalb

---

<sup>71</sup> Eine Ausnahme stellt das Beispiel der Beyaz-Show dar, welches in Kapitel 6 behandelt wird. Die kontrollierenden Eingriffe der Postproduktion bestehen dabei im Abschneiden und im sogenannten Voice-over („Übersprechen“) des ursprünglichen Videomaterials.

<sup>72</sup> Im Sinne einer begrifflichen Differenzierung müsste man dies Gruppe von „abbildenden Bildproduzent/inn/en“ wohl als „montierende“ oder „postproduzierende“ Bildproduzent/inn/en bezeichnen. Flusser (1997, S. 93) unterscheidet in diesem Zusammenhang zwischen jenen Personen „welche den Filmstreifen erzeugen“ und jenen „die den Film erzeugen“. Letzteren attestiert Flusser übrigens eine weitreichendere Allmacht als Gott, denn dieser müsse sich zumindest an die zeitliche Linearität der Geschichte halten.

<sup>73</sup> Es geht mir hier lediglich um eine analytische bzw. funktionale Unterscheidung einzelner Gruppen. Die funktionale Differenzierung schließt nicht aus, dass Videos und Filme auch von einer Person in Personalunion produziert werden können.

als auch zwischen Fotogrammen dokumentieren. Fotogramme repräsentieren im Rahmen der dokumentarischen Video- und Filminterpretation Einstellungen. Zur Untersuchung von Einstellungswechseln, d.h. von Montage oder Bildmischung, werden die Fotogramme am Übergang miteinander verglichen. Somit lässt sich die Bestimmung von Farbkontrasten als Instrumente zur wechselseitigen Validierung der Analyse der Einstellung, d.h. des Fotogramms in seiner Simultanstruktur, und der Montage, d.h. der Struktur der Sequenzialität, einsetzen. In diesem Zusammenhang kann die Rekonstruktion der Farbkontraste, die beim Einstellungswechsel entstehen, auch einen eigenen Zugang zur Montage eröffnen.

Im Fallmaterial des Musikvideos „Araba“ (1996) (siehe Videotranskript: Abbildung 44) betonen Farbwechsel die Schnitte und damit die Einstellungswechsel, indem sie den visuellen Kontrast zwischen den aufeinanderfolgenden Einstellungen erhöhen<sup>74</sup>. Werden Farben über den Schnitt hinweg beibehalten, entsteht vor diesem Hintergrund dagegen eine engere Verbindung zwischen den Einstellungen. In professionellen Film- und Videoproduktionen werden laut Frost (2009) durch die Verbindung farblich homologer Einstellungen auch Räume<sup>75</sup> und Atmosphären geschaffen bzw. durch die Verbindung farblich heterogener Einstellungen Ortswechsel angezeigt. Aus der empirischen Rekonstruktion der Verhältnisse von Einstellungs- und Farbwechseln sind im Rahmen der Montage somit auch Zugänge zur spatialen und temporalen Ordnung von Filmen und Videos zu erwarten.

#### **7.4 Zur Auswahl der Sequenzen: Repräsentanz und Fokussierung**

Als Strategien für die Auswahl geeigneter Sequenzen können nach Bohnsack (2009, S. 196) das Kriterium der „Repräsentanz“ sowie die Kriterien von „Fokussierungen“ bzw. „Diskontinuitäten und Brüchen“ im Material eingesetzt werden. Hinsichtlich der Auswahlstrategie der Repräsentanz stellt die Gattung der Musikvideos eine große Herausforderung dar. Im Gegensatz zu moderierten TV-Unterhaltungssendungen (etwa: Bohnsack, 2009; Hampl, 2010) oder Amateurvideos (etwa: Baltruschat, 2010), die formal durch eine überschaubar geringe Anzahl an Einstellungen und Schnitten bestimmt sind, artikuliert sich in der rhythmisierten Bildsprache von Musikvideos zum einen eine enorm hohe Schnitt- und Einstellungsichte. Oft wird im Sekundenrhythmus oder sogar in noch kürzeren Zeitintervallen geschnitten, was zu einer starken

---

<sup>74</sup> Diese Rolle der Farben beim Schnitt wird auch von Computerprogrammen zur automatischen Schnitterkennung genutzt.

<sup>75</sup> Nicht umsonst gibt es im Deutschen den Begriff „Farbraum“.

Zergliederung des Materials führt. Zum anderen sind einzelne Einstellungen bzw. Sequenzen teilweise mehrfach kompositorisch ineinander verschränkt. Zwar gibt es Einstellungen, die häufiger und weniger häufig auftreten, doch lassen sie sich nicht ohne weiteres in einem hierarchischen Verhältnis von Haupt-, Unter- und eingelagerten Sequenzen verstehen. Diese spezifische Gattungsproblematik verdeutlicht rein visuell auch das Videotranskript (Abbildung 44) der Eingangspassage des Musikvideos „Araba“ (1996).

## **8 „Araba“ (1996): Die Version für den türkischen Markt**

Als Fallmaterial wurde im Anwendungsbereich der Interpretation von *Filmen als Alltagsdokumenten* (Bohnsack, 2009) das Musikvideo „Araba“<sup>76</sup> des türkischstämmigen Popstars *Mustafa Sandal* in der Version von 1996 ausgewählt. Laut eigenen Angaben des Sängers (in der deutschen Late-Night-Show „Istanbul Total“<sup>77</sup>) handelt es sich bei dem durch das Video propagierten Popsong um den größten Erfolg seiner Karriere. Das Album „Araba“ wurde von der Plattenfirma Sony Music France über 2.000.000 Mal verkauft<sup>78</sup>.

### **8.1 Auswahl der Sequenzen**

Dem Beispiel Bohnsacks folgend (Bohnsack, 2009, S.175) lassen sich die Fokussierungskriterien zur Auswahl der Sequenzen (Kap. 7.4) miteinander kombinieren: Im vorliegenden Fall wurden aus den beiden Musikvideos von „Araba“ die ersten vierzig Sekunden ausgewählt, da sie einerseits im Sinne einer Eingangspassage in das Musikvideo einführen und andererseits durch kompositorische Auffälligkeiten auf der Ebene des Schnitts sowie einzelner Einstellungen gekennzeichnet sind.

### **8.2 Videotranskription der Eingangspassage nach dem System MoViQ**

Das folgende Videotranskript wurde mit dem Programm MoviScript (Kap. 3) erstellt. Da der methodologische Fokus dieses Kapitels, wie eingangs analytisch begründet, in der Rekonstruktion der Bilddimension liegt, handelt es sich um ein reines Bildtranskript ohne Text- bzw. Tonspur.

---

<sup>76</sup> zu Deutsch: Auto, Karren

<sup>77</sup> Auftritt am 12. Mai 2004. Link:<http://tvtotal.prosieben.de/tvtotal/videos/player/?contentId=14163>

<sup>78</sup> Mustafa Sandal erhielt im Rahmen der türkischen *Kral Music Awards* dafür die Auszeichnung „Bester Popsänger des Jahres“. Außerdem wurde „Araba“ in der Türkei zum „besten Song des Jahres“ gekürt.

Passage (oder Sequenz): Eingangspassage  
 Film (oder Video): „Araba“ (1996) für den türkischen Markt  
 Dateiname: <http://www.youtube.com/watch?v=b0m8YPF3N70>  
 Time Code: 0-291  
 Dauer: 4:51 Min.  
 Transkription: Stefan Hampl  
 Korrektur: Stefan Hampl



Abbildung 44: MoViQ-Videotranskript des Musikvideos „Araba“ (1996) von Mustafa Sandal, Sony Music France

### 8.3 Formulierende Interpretation der Sequenzen

Entsprechend der durch die Literatur vorgegebenen Arbeitsschritte (Bohnsack, 2009, S. 198) setzt die Videointerpretation mit der formulierenden Interpretation der Sequenzen ein. Auf der Basis der ikonografischen Untersuchung der Bildinhalte gelingt dabei – trotz hoher Zergliederung des Materials – durchaus die Identifikation von zusammenhängenden Sequenzen. Probleme bereitet jedoch die Zuordnung der Sequenzen zu den üblichen Kategorien der Film- und Videointerpretation „Hauptsequenz“ (HS), „Untersequenz“ (US) und „eingelagerte Sequenz“ (ES). Wie das Videotranskript (Abbildung 44) verdeutlicht, ist diese aufgrund der hohen Schnittdichte und Verschachtelung der Sequenzen kaum möglich. Als eingelagerte Sequenzen könnten mitunter noch die Sequenzen betrachtet werden, die nur einmal im Videotranskript auftreten. In der folgenden Liste wären dies „Fischer“ (21-23), „Schiffsbug“ (29), „Filmcountdown“ (0-1) und „Hund“ (26). Genau genommen ist der „Fischer“ jedoch auch im Hintergrund der Sequenz „Ruderer“ (US: 31; 33; 35–36 Ruderer von hinten, vor Ruderboot und Fischer; Amerikanische<sup>79</sup>) zu sehen, wodurch der Versuch der Sequenzzuordnung wieder konterkariert wird. Die Sequenzen „Ferrari“ (8–9,5; 15–17; 34; 24–25; 32; 37; 39), „Ruderer“ (9,5–14; 18–20,5; 31; 33; 35–36; 38) und „Sänger“ (2–7,5; 27–28) wiederum könnten jeweils sowohl als Haupt- und Untersequenzen als auch als eingelagerte Sequenzen aufgefasst werden. Da an der Zuordnung der Sequenzen nach Bohnsack (2009, S. 201) in weiterer Folge die Auswahl repräsentativer Fotogramme hängt, entsteht hier ein Entscheidungsproblem. Für dessen Lösung war ich im vorliegenden Fall gezwungen, von den üblichen Arbeitsschnitten (Bohnsack, 2009, S. 176) abzuweichen, ohne diese völlig aufzugeben. Aus forschungspraktischer Sicht kann es als vertretbarer Kompromiss angesehen werden, vorerst von einer hierarchischen Ordnung der Sequenzen (HS, US, ES) abzusehen und lediglich von reinen Sequenzen zu sprechen. Zur Auswahl der Fotogramme lassen sich dann auf der Basis ihrer ikonografischen Identität (vgl. Kap. 6.3.1) die umfangreichsten Sequenzen der Eingangspassage heranziehen. Im vorliegenden Falle sind dies die Sequenzen „Ferrari“, „Ruderer“ und „Sänger“.

---

<sup>79</sup> Als „Amerikanische“ bzw. „amerikanische Einstellung“ wird ein Bildausschnitt bezeichnet, bei dem die Personen vom Kopf bis etwas unterhalb der Hüfte zu sehen sind.

### 8.3.1 Vorikonografische Interpretation der Sequenzen bzw. Schnitte

1. TC 0–1	Abwärtszählender Countdown: 5,4,3,2
2. TC 2–7	Sitzender Mann
3. TC 8–9,5	Teile eines gelben Autos
4. TC 9,5–14	Ein Mann geht an einem Ruderboot entlang
5. TC 15–17	Teile eines gelben Autos
6. TC 18–20,5	Mann und Ruderboot
7. TC 20,5–23	Aufwachender Mann im Ruderboot
8. TC 24–25	Teile eines gelben Autos
9. TC 26	Auf der Straße liegender Hund
10. TC 27–28	Sitzender Mann
11. TC 29	Schiffsbug mit Bugwelle
12. TC 30–31	Mann und Ruderboot
13. TC 32	Teile eines gelben Autos
14. TC 33	Mann und Ruderboot
15. TC 34	Teile eines gelben Autos
16. TC 35–36	Mann und Ruderboot
17. TC 37	Hände am Lenkrad
18. TC 38	Hände an den Fußrasten eines Ruderboots
19. TC 39	Hände am Lenkrad

### 8.3.2 Ikonografische Interpretation bzw. Auswahl der Sequenzen

Wie in Kapitel 6.3.1 erwähnt, ist zur Bestimmung von Hauptsequenzen (HS) üblicherweise die Feststellung einer „Identität in der Szenerie oder Teilen der Szenerie“ (Bohnsack, 2009, S. 198) erforderlich. Innerhalb der Hauptsequenzen stellen Untersequenzen (US) Variationen der Bildinhalte dar. Auf Basis dieser Konzeption lässt sich in weiterer Folge zumindest eine provisorische Zuordnung von Sequenzen argumentieren, auch wenn die zweifelsfreie Sequenzzuordnung beim vorliegenden Fallmaterial schwierig ist (s.o.). Dieser Schritt ist erforderlich, um nach dem Schema der üblichen Arbeitsschritte eine Auswahl repräsentativer Fotogrammen treffen zu können.

#### 8.3.2.1 Sequenz „Ferrari“

US 8–9,5	Ferrari von hinten in Detailaufnahme
US 15–17; 34	Ferrari von vorne in Detailaufnahme
US 24–25; 32	Ferrari von der Seite in Detailaufnahme
US 37; 39	Hände am Lenkrad des Ferraris in Detailaufnahme

#### 8.3.2.2 Sequenz „Ruderer“

US 9,5–14	Ruderer von hinten neben Ruderboot; Detailaufnahme
US 18–20,5	Ruderer von der Seite neben Ruderboot; Halbtotale
US 31; 33; 35–36	Ruderer von hinten, vor Ruderboot und Fischer; Amerikanische
US 38	Hände an Fußrasten des Ruderboots; Detailaufnahme

### **8.3.2.3 Sequenz „Sänger“**

US 2–7,5; 27–28 Sänger im Schneidersitz von vorn; Totale

### **8.3.2.4 Sequenz „Fischer“**

US 20,5–23 Liegender Fischer; Halbtotale

US 31; 33; 35–36 Ruderer von hinten, vor Ruderboot und Fischer; Amerikanische

### **8.3.2.5 Sequenz „Schiffsbug“**

US 27,5–30 Schiffsbug mit Bugwelle; Detailaufnahme

### **8.3.2.6 Sequenz „Filmcountdown“**

US 0–1 Filmcountdown

### **8.3.2.7 Sequenz „Hund“**

US 26 Liegender Hund; Totale

## **8.4 Auswahl der Fotogramme**

Die Auswahl der Fotogramme stellt im Rahmen der dokumentarischen Methode kein grundsätzliches Problem dar, da sich der Modus Operandi der abgebildeten und abbildenden Bildproduzent/inn/en seinem Wesen nach aus jedem beliebigen Fotogramm eines Films oder Videos rekonstruieren lassen muss. Dadurch ist jedoch nicht ausgeschlossen, dass im Fallmaterial Fotogramme zu finden sind, die aufgrund ihrer formalen und inhaltlichen Beschaffenheit geeigneter als andere zur Beantwortung des Erkenntnisinteresses sind. Der Dokumentsinn von Filmen oder Videos tritt in den einzelnen Fotogrammen unterschiedlich stark fokussiert zutage. Nicht zuletzt auch im Sinne der Erhöhung der Nachvollziehbarkeit des Auswahlprozesses und eines geringeren interpretativen Aufwandes, also aus ökonomischen Gründen, tritt Bohnsack (2009, S.201) deshalb für eine begründete Auswahl von Fotogrammen ein. Als Auswahlstrategien schlägt er neben dem Kriterium der Repräsentanz das der Fokussierung vor. Unter Repräsentanz versteht Bohnsack (2009, S.201) die Verknüpfung zwischen den im Film- oder Videomaterial identifizierten Sequenzen (Hauptsequenz, Untersequenz) und den aus ihnen ausgewählten Einzelbildern: „Fotogramme [sollten] zumindest die beiden umfangreichsten Hauptsequenzen (HS) in deren umfangreichsten Einstellungsvarianten, also in den Untersequenzen (US), repräsentieren [...]“ (Bohnsack, 2009, S.201). Fokussierungen im Film- oder Videomaterial lassen sich laut Bohnsack (2009, S.201) in Form von „dramaturgischen Steigerungen“, aber auch durch „Diskontinuitäten und Brüche“ erkennen.

Da in diesem Kapitel die Untersuchung der Montage bzw. der Einstellungswechsel von vorrangigem Interesse ist, habe ich aus den Sequenzen „Ferrari“ und „Ruderer“ jeweils ein Fotogramm aus deren Übergangsbereich gewählt: Fotogramm 16 und Fotogramm 18. Aus der Sequenz „Sänger“ wurde das Fotogramm 6 ausgewählt, das den Sänger in einer auffälligen Mimik und Körperhaltung zeigt.

## 8.5 Interpretation der Fotogramme

### 8.5.1 HS/ES I: „Ferrari“



Abbildung 45: Fotogramm 16, Sequenz „Ferrari“

#### 8.5.1.1 Formulierende Interpretation

##### Vorikonografische Interpretation

Fast vollständig bildfüllend ist ein Ausschnitt eines gelben Sportautos zu sehen. Ein Teil seiner Front sowie der linken Wagenseite erstreckt sich vom Bildvorder- bis in den Bildmittelgrund. Im Mittelgrund ist vor der dunklen Frontscheibe des Autos am linken Bildrand der Teil eines Scheibenwischers erkennbar. In der rechten Bildhälfte ragt ein rundlicher, gelber Gegenstand zur Seite. Im Bildhintergrund sind in der linken Bildhälfte hinter der Frontscheibe des Autos die Konturen von Armaturenbrett und Lenkrad erkennbar. In der rechten Bildhälfte ist brauner Schotter zu sehen.

Ober- und unterhalb des Bildes befinden sich schwarze, horizontale Balken. Links im oberen Balken steht in blauer Schrift „remixland.de“. Im Balken in der unteren Bildhälfte steht „w.remixland.de tarafindan sunulmustur“. Rechts unten über den Buchstaben „ustur“ befindet sich ein rotes Dreieck mit der weißen Inschrift „NEV TV“.

## **Ikongrafische Interpretation**

Bei dem gelben Auto handelt es sich um einen Luxusportwagen der Marke Ferrari mit getönten Scheiben und Leichtmetallfelgen. Bei den schwarzen Balken ober- und unterhalb des Bildes handelt es sich um eine sogenannte „Letterbox“<sup>80</sup>. Das obere Wort in der Letterbox kann aufgrund der Endung „.de“ als deutsche Internetseite interpretiert werden. Der untere Text in der Letterbox ist türkisch und bedeutet „präsentiert von (ww)w.remixland.de“. NEV TV ist der Name eines ehemaligen türkischen Musik-TV-Senders, über den das Musikvideo ursprünglich ausgestrahlt worden sein dürfte. Die Texte mit Bezug auf remixland.de, welche sich vor dem Logo des Senders befinden, weisen darauf hin, dass jemand das Video digital aufgezeichnet hat, um es dann im Internet über die Website [www.remixland.de](http://www.remixland.de) zu verbreiten.<sup>81</sup>

### **8.5.1.2 Reflektierende Interpretation**

#### **Planimetrische Komposition**

Der gelbe Ferrari bzw. zumindest dessen Ausschnitt ist im Bild planimetrisch fokussiert, da er fast das gesamte Bild ausfüllt. Diese Fokussierung ist durch drei Feldlinien charakterisiert, die ihren gemeinsamen Schnittpunkt am Fuße der A-Säule des Ferraris, knapp unterhalb des fahrerseitigen Rückspiegels haben (Abbildung 46).

---

<sup>80</sup> Zu der Zeit, als das Fernsehen noch ein Seitenverhältnis von 4:3 aufwies, waren Letterboxes aus technischen Gründen erforderlich, um im Seitenverhältnis 16:9 aufgenommene Kinofilme auf Fernsehgeräten darzustellen. Aufgrund dieses Zusammenhangs werden Letterboxes in Musik- oder Werbevideos oft aus ästhetischen Gründen (im Zuge der Montage bzw. Postproduktion) hinzugefügt, um Videos eine spielfilmartige Anmutung zu verleihen.

<sup>81</sup> Die Texte in den Letterboxes wurden dem Musikvideo nachträglich hinzugefügt und fließen daher nicht in die weitere Interpretation ein. Es handelt sich bei ihnen um die Leistungen sekundärer und tertiärer Bildproduzent/inn/en, in denen sich dokumentiert, dass es sich um ein populäres Musikvideo handelt. Das Erkenntnisinteresse der Arbeit ist jedoch auf die Leistungen der primären Bildproduzent/inn/en gerichtet.

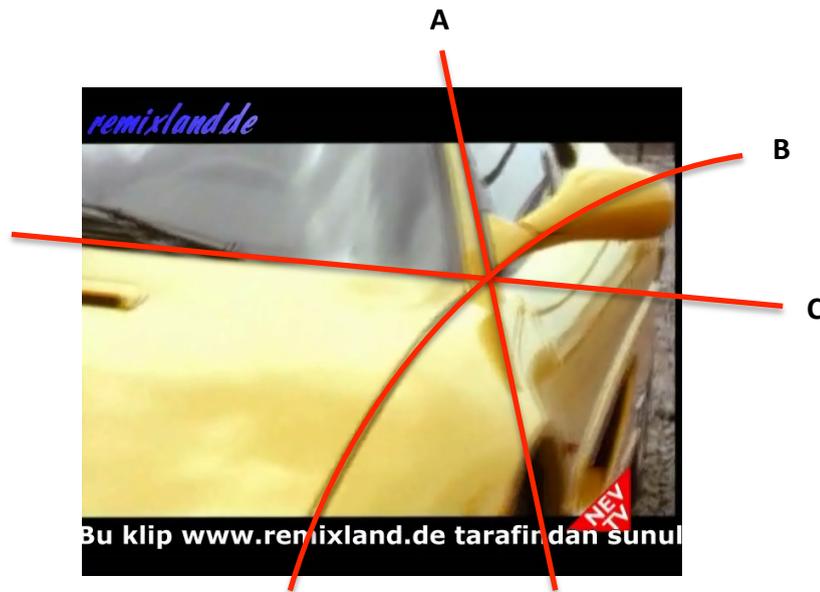


Abbildung 46: Fotogramms 16 mit Feldlinien

Die Feldlinie A verläuft gerade, aber in leichter Schräge von oben nach unten entlang der A-Säule und dann weiter über den Kotflügel und den Radkasten des Ferraris. Die Feldlinie B krümmt sich entlang des Spalts von Motorhaube und Kotflügel über den Seitenspiegel nach rechts. Die Feldlinie C verläuft von links nach rechts schräg abwärts entlang der Unterseite der Windschutzscheibe.

Des Weiteren ist der Wagen auch farblich fokussiert. Das Bild ist von einem starken Gelb-Schwarz-Kontrast geprägt. Nach Ittens (2000, 2010) System der Farbkontraste dokumentieren sich hierin drei der sieben Farbkontraste: a) der Hell-Dunkel-Kontrast (durch strahlendes Gelb und Weiß gegenüber Schwarz), b) der Qualitätskontrast (durch gesättigtes in Gegenwart von ungesättigtem bzw. abgedunkeltem Gelb) und c) der Quantitätskontrast (Gelb besitzt gegenüber allen anderen gesättigten Spektralfarben den höchsten Helligkeitswert). Unter Berücksichtigung der nachträglich hinzugefügten Logos und Texte von NEV TV und Remixland könnte auch ein Farbe-an-sich-Kontrast geltend gemacht werden (durch Vorhandensein aller drei Grundfarben im Bild: Gelb, Rot und Blau).

### **Perspektivische Konstruktion**

Der gelbe Ferrari ist in Schrägperspektive aufgenommen. Beide Fluchtpunkte befinden sich außerhalb des Bildes. Unter der Voraussetzung, dass die Fahrzeugkanten in der Realität parallel zueinander verlaufen, können deren Linien verlängert und so der erste Fluchtpunkt (FP 1) bestimmt werden. Er befindet sich vermutlich rechts oberhalb des Bildes. Der zweite Fluchtpunkt kann aus der Verlängerung der Windschutzscheiben und

des Luftschlitzes annäherungsweise erschlossen werden. Als Verbindungslinie zwischen den Fluchtpunkten ergibt sich so die Horizontlinie, die zugleich die Höhe der Kameraposition reflektiert.



Abbildung 47: Fotogramm 16 mit perspektivischen Linien

Bei diesem Bild haben wir es – in Kombination mit der starken Aufsicht – mit einer markanten Schrägperspektive zu tun, die das Bild durch die weit auseinanderliegenden Fluchtpunkte stark in der Horizontale zerdehnt (Seitensicht). Dadurch gewinnt der abgebildete Ferrari massiv an „visuellem Gewicht“ (Arnheim, 2000, S. 26). Dieser Effekt wird zusätzlich verstärkt, indem die Kamera aus erhöhter Position nach unten gerichtet ist (Aufsicht). So wird im Bild nämlich der Eindruck der Räumlichkeit betont, wodurch der Ferrari weiteres Volumen erhält und in seiner ganzen körperlichen Materialität (jenseits einer technisch-instrumentellen Funktionalität) zur Geltung kommt. Um den Eindruck zu verdeutlichen, der durch die perspektivische Konstruktion des Bildes entsteht, habe ich hier ein Vergleichsbild eingefügt. Der Ferrari wirkt wie ein aufgeblasenes „Fat Car“ des österreichischen Künstlers Erwin Wurm.



Abbildung 48: Erwin Wurm (1993). *Fat Car* (roter Porsche), Quelle: <http://www.dsgnwrld.com>

### **Szenische Choreografie**

Hinsichtlich der szenischen Choreografie ist der gelbe Ferrari, ebenso wie auf den vorhergehenden Interpretationsebenen, bildbestimmend. Hinter der verspiegelten Windschutzscheibe des Sportwagens ist kein Fahrer erkennbar. Damit ist der Ferrari das einzige Objekt im Bild. Aufgrund dieser besonderen Einzelstellung und insbesondere des fehlenden Fahrers sowie der Aufnahme im Dreiviertelprofil wird der Ferrari seiner technischen Funktionalität enthoben und wirkt letztlich wie ein Ausstellungsobjekt in einer Kunstgalerie – vergleichbar dem „fetten“, roten Porsche von Erwin Wurm.

### **Ikonologische bzw. Ikonische Interpretation**

Im Sinne der zusammenfassenden Interpretation des Fotogramms und damit dieser Sequenz kann gesagt werden, dass der gelbe Ferrari auf sämtlichen Bildebenen fokussiert ist. Die Übergegensätzlichkeit dieses Luxusportwagens besteht darin, dass hier ein technisch-funktionales Gerät zum Zwecke der Mobilität zu etwas Unlebendigem, Unbelebtem, Ornamentalem wird. Darin zeigen sich Homologien zu Wurms rotem Porsche. Dies gilt auch für die Person des Fahrers, derer es bedarf, um die technische Funktionalität und Mobilität zu sichern. Der zweite Gedanke dokumentiert sich im Fotogramm im fokussierten Bildausschnitt der Fahrerseite des Wagens.

## 8.5.2 HS/ES II: „Ruderer“



Abbildung 49: Fotogramm 18, Sequenz „Ruderer“

### 8.5.2.1 Formulierende Interpretation

#### Vorikonografische Interpretation

Im Bildvordergrund sind ein schwarz-weißes Ruderboot sowie ein graues Stück Mauer zu sehen, die sich weiter in den Bildmittelgrund erstrecken. Auf der Mauer liegen zwei schaufelartige Ruder. Im Bildmittelgrund befinden sich links mehrere weiße Boote und Masten auf einer Wasserfläche sowie rechts eine Häuserzeile. Davor ist ein stehender Mann in grauem Pullover und schwarzer Mütze zu sehen. Der Mann ist der Wasserfläche zu- und den Bildbetrachter/inne/n abgewandt. Er stützt sich mit dem rechten Arm auf den Balken. Der linke Arm ist angewinkelt Richtung Hals geführt, sodass die Hand nicht sichtbar ist. Im Bildhintergrund befindet sich ein hellblauer Himmel mit einigen weißen Wolken.

#### Ikonografische Interpretation

Das Bild zeigt den Ausschnitt einer Hafenanlage mit privaten Segel- und Motorjachten. In dieser Umgebung bereitet sich ein Ruderer darauf vor, in See zu stechen. Ruder und Ruderboot sind an der Hafenmole bereitgelegt.

## 8.5.2.2 Reflektierende Interpretation

### Planimetrische Komposition

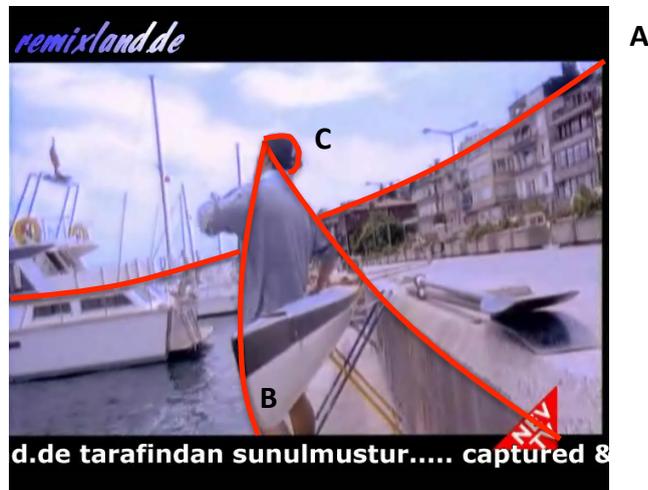


Abbildung 50: Fotogramm 18 mit Feldlinien

Im Vergleich zu Fotogramm 16 mit dem gelben Ferrari ist Fotogramm 18 weniger flächig, sondern eher kleingliedrig, d. h. durch viele kleine Details strukturiert. Dennoch können dominante Feldlinien bestimmt werden. Feldlinie A verläuft von links nach rechts oben entlang der Oberkante einer weißen Motorjacht und dann entlang der rechten Häuserzeile bis zur rechten, oberen Bildecke. Feldlinie B verläuft gekrümmt entlang des Bootrumpfs sowie entlang der Mauer der Hafentreppe nach oben spitz über den Oberkörper des Ruderers zu. Feldlinie C markiert seinen auffällig runden Hinterkopf. Der Ruderer im Bild ist dadurch fokussiert, dass sein Oberkörper aus der Feldlinie A (Abbildung 50) hervorragt und sein Hinterkopf (Feldlinie C) eine besondere Betonung der Spitze von Feldlinie B darstellt. Durch die dominanten planimetrischen Linien entsteht eine starke Dynamik nach links, infolge derer der Ruderer nach links zu kippen scheint. Die Gesamtkonstellation ragt wie ein hoher Felsen, eine Welle oder ein scharfer Zahn in den Himmel.

Farblich gesehen ist das Bild eher blass gehalten und nicht durch starke Farbkontraste bestimmt. Daher dokumentiert sich in ihm auch keiner der Farbkontraste aus dem Fotogramm mit dem gelben Ferrari. Am deutlichsten ist noch der Hell-Dunkel-Kontrast ausgeprägt. Dieser hat nach Itten (2000, 2010) eine zweifache Wirkung: Zum einen werden durch Helligkeitsunterschiede Konturen und Formen betont. So wird etwa die schwarze Mütze des Ruderers vor dem hellblauen Hintergrund im Bild hervorgehoben. Zum anderen erstreckt sich im Intervall von Hell und Dunkel eine

„ungewöhnlich große Zahl von hellen und dunklen Grautönen“ (Itten, 2010, S. 37). Neutrales Grau hat die besondere Eigenschaft, auch von schwachen Nachbarfarben zu einer komplementären Farbwirkung angeregt zu werden. Dieser zweite Effekt des Hell-Dunkel-Kontrasts ist eher subtil und beeinflusst vor allem die globale Atmosphäre des Bildes. Der in leichtes Hellblau getünchte Himmel regt die „graue“ Umgebung dazu an, in komplementärem Rot-Orange zu erscheinen. Dieser Effekt ist dafür verantwortlich, dass sich das Bild trotz seiner schwachen Farbsättigung deutlich von einem reinen Schwarz-Weiß-Bild unterscheidet.

### Perspektivische Konstruktion



Abbildung 51: Fotogramm 18 mit perspektivischen Linien

Im Vergleich zur perspektivischen Konstruktion des Ferrari-Fotogramms haben wir es im vorliegenden Fall (Abbildung 51) mit keiner Schrägperspektive sondern einer gekippten Kamerahaltung zu tun, bei der es nur einen Fluchtpunkt (FP 1) in der Bildmitte gibt. Über die Bestimmung des Fluchtpunkts lässt sich die Horizontlinie ermitteln. Diese reflektiert zugleich die Aufnahmhöhe der Kamera. Aus dem Unterschied zwischen Bildmittelsenkrechte und Wasseroberfläche lässt sich zudem feststellen, dass die Kamera bei der Aufnahme um ca. zehn Grad nach rechts gekippt wurde. Dies hat unter anderem Auswirkungen darauf, wie der Ruderer im Bild in Erscheinung tritt. Im vorliegenden Fall nimmt er eine Mittel- oder Mittlerposition zwischen Wasser und Land ein, indem er aufrecht und frei neben einem Ruderboot steht und damit Gefahr läuft ins Meer zu kippen. Dreht man das Bild im Sinne einer Kompositionsvariation nach Imdahl (1994) um ca. zehn Grad nach rechts, erscheint der Ruderer stärker an das Ruderboot angelehnt bzw. aufgestützt und zum Wasser auf

Distanz<sup>82</sup>. Indem der Fluchtpunkt im Rücken des Ruderers liegt, gleichsam in seinen Rücken „drückt“, wird der Eindruck der Kippbewegung des Ruderers verstärkt.

### **Szenische Choreografie**

Auf der Ebene der szenischen Choreografie ist festzustellen, dass sich nur eine einzige Person, nämlich der Ruderer, im Bild befindet. Dieser ist in aufrechter Körperhaltung abgebildet. Hinter einem auf einem Gestell liegenden Ruderboot sind sein Rücken, sein Hinterkopf sowie die Waden zu sehen, da er den Bildbetrachter/inne/n gegenüber abgewandt ist. Aus Kopf- und Körperhaltung kann geschlossen werden, dass der Ruderer vermutlich aufs Wasser hinausblickt. Aus seiner angewinkelten Armhaltung lässt sich mithilfe der nachfolgenden Fotogramme 19 und 20 (Abbildung 44) eine Geste rekonstruieren, die das Ausziehen des Pullovers vorbereitet. Gegenüber dem Ferrarifahrer ist der Ruderer zwar im Bild zu sehen, gibt durch sein abgewandtes Gesicht jedoch ebenso wenig wie jener seine Identität preis.

### **Ikonologische bzw. ikonische Interpretation**

Der Ruderer ist im Bild auf mehreren Ebenen durch Übergegensätzlichkeiten charakterisiert. Planimetrisch gesehen, hebt er sich farblich kaum vor dem neutral grauen Hintergrund ab und ragt doch mit seinem Hinterhaupt aus der Kulisse der umliegenden Häuser und Schiffe hervor. Aus der Differenz von perspektivischem und planimetrischem Zentrum ergibt sich eine Fokussierung seines Kopfes. Auch der ist letztlich aber unter einer Sportmütze verborgen. Damit gibt sich der Ruderer – im Gegensatz zum Ferrarifahrer – als Person zu erkennen. Aufgrund des abgewandten Gesichts ist jedoch seine genaue Identität nicht feststellbar. In der rekonstruierten Gesichtslosigkeit der beiden Fotogramme dokumentiert sich so letztlich eine bedeutsame Homologie zwischen ihnen, die aufgrund ihres starken Kontrastverhältnisses nicht auf Anhieb auffällt. Den Verdacht, dass die im Musikvideo aufgeworfenen Unterschiede mitunter eine Maskierung der Gemeinsamkeiten darstellen könnten, wollen wir uns an dieser Stelle als Suchstrategie für die Interpretation der dritten Sequenz festhalten.

---

<sup>82</sup> In der Tat erinnert der durch die Kompositionsvariation hervorgerufene Effekt stark an Imdahls (1994) Beispiel anhand der Miniatur des *Hauptmanns von Kapernaum* (Codex Egberti).

### 8.5.3 HS/ES III: „Sänger“



Abbildung 52: Fotogramm 6, Sequenz „Sänger“

#### 8.5.3.1 Formulierende Interpretation

##### Vorikonografische Interpretation

Im Bildvordergrund ist annähernd in Bildmitte ein am Boden sitzender Mann zu sehen, der den Bildbetrachter/inne/n direkt zugewandt ist. Gesicht und Oberkörper sind vollständig zu sehen. Seine Haare sind schwarz und kurz geschnitten, der Kopf ist leicht schräg – etwas nach rechts und nach hinten gebeugt, der Mund ist geöffnet, die Mundwinkel sind angezogen. Sein Blick ist (von ihm aus gesehen) nach rechts oben gerichtet. Seine Arme hängen über die angewinkelten Beine bzw. Knie herab. Er ist mit einem grauen Poloshirt, einer weißen Hose sowie schwarzen Schuhen bekleidet. An seinem rechten Handgelenk ist ein silbernes Armkettchen zu sehen, an seinem linken eine silberne Armbanduhr. Der Bildhintergrund besteht im Wesentlichen aus großen, horizontalen Farbstreifen. Sie sind blassgrün, rosarot, weiß und orange. Im oberen grünen Farbstreifen und im weißen Farbstreifen sind verschiedene Formen und Muster zu sehen. Der Hintergrund wird an der linken Bildkante durch eine weiß-grau-schwarz gemusterte Säule abgeschlossen.

##### Ikonografische Interpretation

Der Mann im Bild ist der türkische Popstar Mustafa Sandal. Er trägt am rechten Handgelenk ein Silber- oder Goldkettchen und am linken Handgelenk eine silberne oder goldene Armbanduhr. Seine Kleidung wirkt alltäglich, kann aufgrund der weißen Hose und der schwarzen Lederschuhe jedoch als sportlich-elegant bezeichnet werden. Die spezielle Körperhaltung wird als Schneidersitz bezeichnet. Sie ist in verschiedenen

Kulturkreisen sowohl bei meditativen Entspannungsübungen als auch bei manuellen Tätigkeiten (z. B. Nähen) oder beim Geschichtenerzählen in einer Gruppe (z. B. um ein Lagerfeuer sitzend) anzutreffen. In dieser Pose zeigt Mustafa Sandal ein strahlendes Lachen. Er sitzt vor einem Hintergrund auf dem in Trompe-l'Œil-Malerei verschiedene Naturmotive (Baumstämme, Steine, Gräser etc.) sowie am linken Bildrand eine antik-orientalische Säule abgebildet sind.

### 8.5.3.2 Reflektierende Interpretation

#### Planimetrische Komposition



Abbildung 53: Fotogramm 6 mit Feldlinien

In planimetrischer Hinsicht ist der Sänger im Bild zum einen durch seine zentrale Position in der Bildmitte fokussiert. Zum anderen stellt seine birnenförmige und nach links geneigte Körperhaltung einen deutlichen Kontrast zur geradlinigen Horizontal- und Vertikalstruktur des Hintergrunds dar.

Farblich gesehen wird das vorliegende Bild vor allem durch einen Kalt-Warm-Kontrast zwischen Grün- und Rottönen bestimmt. Aufgrund seiner völlig „farblosen“ Kleidung nimmt der Sänger dazwischen eine neutrale Mittelposition ein. Insgesamt wirkt das Bild eher blass und hell, was eine Homologie zum Fotogramm des Ruderers darstellt.

#### Perspektivische Konstruktion

In perspektivischer Hinsicht hat das Bild wenig Tiefenwirkung. Es handelt sich um eine „gestauchte“ Perspektive, die fotografisch durch die Verwendung starker Teleobjektive entsteht. Je stärker der Effekt der Stauchung des Raumes, desto weiter muss die Kamera von den abgebildeten Bildproduzent/inn/en entfernt sein. Aufgrund der parallel

verlaufenden Linien im Hintergrund sind keine Fluchtpunkte rekonstruierbar. Beide Aspekte tragen dazu bei, den Sänger mit dem Hintergrund zu verbinden, worin auf dieser Ebene die Fokussierung besteht.

### **Szenische Choreografie**

Der lachende Popstar Mustafa Sandal sitzt den Bildbetrachter/inne/n mit geöffneten Armen und lässig über die Knie herabhängen Händen im Schneidersitz frontal gegenüber. Sein Kopf ist leicht seitwärts und nach hinten gebeugt, seine Augen blicken nach rechts oben. Der Sänger ist im Bild zum einen dadurch fokussiert, dass er allein sitzt. Zum anderen stellt die Kombination aus seiner spezifischen Sitzhaltung und dem strahlenden Lachen eine spezielle Form der gestischen Steigerung dar.

### **Ikonologische bzw. ikonische Interpretation**

Mustafa Sandal ist dadurch im Bild fokussiert, dass er allein sitzt und gleichzeitig den Bildbetrachter/inne/n frontal zugewandt ist. Die Übergegensätzlichkeit des Fotogramms besteht darin, dass der Sänger dabei sowohl geöffnet als auch verschlossen ist. Beispielsweise öffnet er die Arme und verschließt gleichzeitig die Beine, er zeigt uns sein Gesicht und wendet zugleich den Blick ab. In diesem Zusammenhang erscheint auch sein Lachen in gewisser Weise als Hinweis auf eine Maskierung. Im konkreten Fall könnte die Maskierung durch das Lächeln darin bestehen, den Bildbetrachter/inne/n letztlich – in ähnlicher Weise wie schon beim Ferrarifahrer und Ruderer – auch den Blick auf das (wahre) Gesicht des Sängers vorzuenthalten.

## **8.6 Zur Relationierung der Fotogramme im Rahmen der Montage**

### **8.6.1 Reflektierende Interpretation von Einstellungswechsel und Montage**

Eine besondere Leistung der Montagetechnik des vorliegenden Musikvideos ist es, dass die drei untersuchten Hauptsequenzen zugleich auch eingelagerte Sequenzen darstellen und damit einander prinzipiell weder über- noch untergeordnet sind. Dieser Montagestil wird in der Filmwissenschaft als *Parallelmontage* bezeichnet. Das Wesen der Parallelmontage<sup>83</sup> ist es, zwei oder mehrere Handlungsstränge einander zeitlich gegenüberzustellen. Diese Verknüpfung kann durch die Verkürzung der Schnittfolgen

---

<sup>83</sup> Eine auf die Rezeption gerichtete Intention der Parallelmontage besteht laut Rabiger (1998, S. 184), die Zusehenden dazu zu bringen, das Gesehene miteinander zu vereinbaren. Durch diesen Mechanismus entfalte die Parallelmontage nicht nur eine ästhetische, sondern auch ihre psychologische und politische Dimension.

intensiviert werden, was meist gegen Ende passiert, bevor sich die Handlungsstränge in einem gemeinsamen Punkt treffen. Die parallele Montage von Ferrari, Ruderer und Sänger stellt diesbezüglich sogar noch einen Spezialfall dar, denn die parallel dargestellten Handlungen stehen noch zusätzlich in einem antithetischen Verhältnis zueinander. In der Typologie von Wulff (2013) handelt es sich dabei um eine sogenannte *Kontrastmontage*; die Verknüpfung der Sequenzen Ferrari und Ruderer kann in weiterer Zuspitzung sogar als *Kollisionsmontage* bezeichnet werden. Dadurch entsteht der Eindruck eines Wettrennens, das auf die direkte Konfrontation der Kontrahenten hinausläuft. Der Sänger ist von diesem Duell ausgenommen. Ihm kommt die Rolle des Barden zu, der die Geschehnisse liedhaft kommentiert, sich aber selbst nicht einmischt – und kraft der Montage auch nicht einmischen kann<sup>84</sup>! Würde man das fahrerische Wettrennen im Musikvideo als sportliches Ereignis betrachten, so wäre der Sänger eher eine Art Livereporter als ein Schiedsrichter. Letzterer hätte nämlich zumindest die Möglichkeit, das Rennen zu starten oder zu beenden, was dem Sänger nicht möglich ist. Wie kann bei den drei betrachteten Sequenzen Ferrari, Ruderer und Sänger jedoch überhaupt der Eindruck eines gemeinsamen *Spielfeldes* entstehen?

Zur Herstellung einer räumlichen Verbindung zwischen den betrachteten Sequenzen reicht der Einsatz der Parallelmontage allein nicht aus. Dazu bedarf es der Präsentation von Einstellungen, in denen im Bild die gleichzeitige Anwesenheit der sonst räumlich getrennten Handlungsstränge an einem Ort zumindest ansatzweise sichtbar wird. Die Belege dafür dokumentieren sich im Video bzw. im Videotranskript (Abbildung 44). In der Sequenz „Sänger“ (Fotogramm 5) etwa wischt das Dach eines Autos durch das Bild, was signalisiert, dass der Sänger sich an einer befahrenen Straße befindet. In der Sequenz „Ruderer“ (Fotogramme 31, 33–36, 39) sieht man den schlafenden Fischer (Fotogramme 21–23) im Hintergrund, wodurch dieser sich mit der Sphäre des Ruderers verbindet. Der Hund (Fotogramm 26) liegt auf der Straße und wird knapp eine Minute später vom vorbeifahrenden Ferrari aufgeschreckt (Fotogramm 69), das Schiff (Fotogramm 29) taucht später in der Sequenz „Ruderer“ (Fotogramme 98–99, 107–108, 140–141) wieder entfernt im Hintergrund auf.

Bohnsack (2009) bezeichnet, mit Rückgriff auf Bordwell (1985), den Schauplatz des Geschehens, der durch Filme und Videos hergestellt wird, als *montierten Raum*

---

<sup>84</sup> Hierin dokumentiert sich ein grundlegender Unterschied zwischen dem Handlungsspielraum eines Popsängers im Musikvideo und des Moderators einer TV-Show (z. B. „Istanbul Total“, Bohnsack, 2009; Hampl, 2010), der die Möglichkeit hat, den Ablauf der Ereignisse zu „dirigieren“ (Bohnsack, 2009, S. 215).

(Bordwell, 1985, S. 117). Mit Souriau (1997) kann die durch die Montage der Sequenzen hergestellte Räumlichkeit auch als *Diegese* bzw. *filmisches Universum* bezeichnet werden. Sämtliche Sequenzen des vorliegenden Videos sind auf gleichzeitige Ereignisse an einem einzigen Ort des Geschehens bezogen und werden im Musikvideo auch durch den gemeinsamen Tonraum miteinander verbunden. Durch subtile visuelle Hinweise (etwa in Fotogramm 5: Das schemenhafte, kurze Durchs-Bild-Wischen eines Autos) und den gemeinsamen Tonraum entstehen wechselseitige Bezüge zwischen den Sequenzen, die am Ende auf den gemeinsamen Ort des Geschehens verweisen: Die gesamte Handlung spielt sich am Bosphorus und in den Straßen der umliegenden Stadt Istanbul ab. Diese werden zur Arena der Wettkampfaustragung. Der einzige Ort, der außerhalb des Geschehens liegt und dennoch Bestandteil des Musikvideos ist, ist der Filmcountdown zu Beginn (Abbildung 44, Fotogramme 0 und 1). Wenngleich der Filmcountdown selbst nicht der physischen Welt entstammt, besteht seine paradoxe Eigenschaft darin anzuzeigen, dass im Folgenden nicht die Realität, sondern ein Film gezeigt wird. Im folgenden Abschnitt möchte ich nun einige Überlegungen bezüglich der farblichen Relationen der drei untersuchten Sequenzen anschließen, die meines Erachtens zu einem tieferen Verständnis der Montage des vorliegenden Musikvideos beitragen können.

### **8.6.2 Reflektierende Interpretation von Farbkontrasten und Montagemustern**

Der im vorigen Abschnitt herausgearbeitete Befund der Kontrastmontage lässt sich auf der Ebene der Farbgebung der untersuchten Sequenzen weiter planimetrisch vertiefen. Wie der Blick auf das Videotranskript (Abbildung 44) zeigt und auch die planimetrische Komposition der Fotogramme verdeutlicht hat, sind in den einzelnen Sequenzen jeweils bestimmte Farben fokussiert. In der Sequenz „Ferrari“ ist es die Farbe Gelb, in der Sequenz „Ruderer“ Hellblau-Violett und in der Sequenz „Sänger“ Braun-Grün. Auch wenn sich diese Farben nicht in jeder Teilsequenz im selben Ausmaß dokumentieren, so bleibt die Zuordnung bestimmter Farben zu bestimmten Sequenzen dennoch das gesamte Video hindurch bestehen.

Wie die weitere Interpretation zeigen wird, ist die Zuordnung bestimmter Farben zu bestimmten Sequenzen alles andere als willkürlich. Laut Frost (2009) ist eine wohlüberlegte Farbstrategie für professionelle Film- und Videoproduktionen

unabdingbar, um den intendierten Ausdruckssinn formal zu unterstützen<sup>85</sup>. Über die Auswahl der Farben für einen Film (oder ein Video) sollte laut Arnheim<sup>86</sup> (1979, S. 51) und Frost (2009) bereits beim Verfassen des Drehbuchs Klarheit bestehen. Danach wäre seitens der abbildenden Bildproduzent/inne/en<sup>87</sup> sowohl während – durch Auswahl der konkreten Beleuchtung und Kleidung der Abgebildeten – als auch nach den Aufnahmen – im Rahmen der Postproduktion bzw. Montage – auf die Einhaltung einer *a priori* festgelegten Farbstrategie zu achten. Im Vergleich zu allen anderen visuellen Gestaltungsstrategien zeichnen sich Farben durch die hohe Flexibilität ihrer Einsatzmöglichkeiten aus. Durch die Beeinflussung der Farbgebung lassen sich laut Frost (2009) sowohl unterschwellige Atmosphären und Stimmungen erzeugen als auch bestimmte Aspekte im Film (Personen, Gegenstände, Orte, Zeiten etc.) identifizierbar machen oder markieren<sup>88</sup>.

Als methodologischer Angelpunkt zur Rekonstruktion der Farben bietet sich Ittens (2000, 2010) Theorie der sieben Farbkontraste an. Ittens Auffassung nach steht keine Farbe für sich selbst, kann keine Farbe allein ein Geschehen bestimmen, sondern ist immer verbunden mit anderen Farben, die mit ihr in größerem oder kleinerem Kontrast stehen. Ittens Theorie scheint nicht nur geeignet für die Interpretation von Einzelbildern und Fotogrammen, sondern insbesondere zur Interpretation der Montage. Sie fügt sich zudem in das Denken in Relationen und Komparationen, wie es charakteristisch ist für die dokumentarische Methode.

Die Interpretation des Einstellungswechsels im vorigen Absatz hat ergeben, dass die Sequenzen „Ferrari“ (Abbildung 45) und „Ruderer“ (Abbildung 49) zueinander in einem Kontrastverhältnis stehen. Dieses Kontrastverhältnis möchte ich mithilfe der Rekonstruktion der Farbkontraste nun im Anschluss empirisch definieren und differenzieren. Im Detail dokumentieren sich im Einstellungswechsel zwischen den beiden genannten Sequenzen nach Itten (2000, 2010) folgende Kontrastrelationen:

---

<sup>85</sup> In der Filmwissenschaft wird der Vorgang der sinnunterstützenden Farbkorrektur auch als „color-grading“ bezeichnet.

<sup>86</sup> Aufsatz von 1935

<sup>87</sup> Für die Farbgebung des Films oder Videos ist in der Regel der *Kameramann* bzw. *cinematographer* oder *director of photography (DP)* gemeinsam mit dem Regisseur verantwortlich. Unter anderem aufgrund des hohen Stellenwerts der farblichen Unterstützung des intentionalen Ausdruckssinns für die professionelle Filmproduktion wechseln laut Frost (2009) berühmte Regisseure ihre Kameralente so gut wie nie und arbeiten mit ihnen oft über Jahrzehnte zusammen (z. B. Steven Spielberg und Janusz Kaminski).

<sup>88</sup> So markiert etwa das blassrote Mäntelchen eines kleinen Mädchens im Schwarz-Weiß-Film „Schindler's Liste“ (1993) eine Schlüsselszene. Filmausschnitt: <http://www.youtube.com/watch?v=j1VL-y9IHuI>

- 1) **Farbe-an-sich-Kontrast:** Durch den Schnitt zwischen den Sequenzen „Ferrari“ und „Ruderer“ wird das Gesamtgeschehen des Musikvideos „bunt“ und „lebendig“. Der Farbe-an-sich-Kontrast entsteht prinzipiell durch das Vorhandensein der drei Grundfarben Rot, Gelb, Blau, ist jedoch auch in abgeschwächter Form wirksam, wenn eine dieser Farben fehlt oder durch Mischfarben substituiert ist. Im vorliegenden Fall wird der Farbe-an-sich-Kontrast vor allem durch das satte Gelb des Ferraris lanciert, wobei die anderen Farben eher in den Hintergrund treten. Blau tritt in blasser Gestalt (Himmel, Wasser) auf. Rot kommt (mit Ausnahme des nachträglich hinzugefügten Logos des TV-Senders NEV TV) fast gar nicht vor. Bestenfalls kommt es in Form der Ferrarirücklichter (Abbildung 44, Fotogramm 24) ins Bild. Ansonsten ist der Farbe-an-sich-Kontrast im Fallmaterial nur aufgrund der blassen Sekundärfarbe Violett (Sequenz „Ruderer“) wirksam.
- 2) **Kalt-Warm-Kontrast:** Am stärksten tritt dieser Kontrast zwischen den Polen Rot – Orange und Blau – Grün bzw. Blau – Türkis in Erscheinung. Im vorliegenden Fall ist jedoch auch dieser Kontrast eher schwach ausgebildet. Zwar ist das Gelb des Ferraris sehr satt und ohne Tendenz zum Grünlichen, d. h. „Kalten“. Das Hellblau-Violett der Ruderersequenz kippt jedoch sehr wohl ins Violett-Rötliche, d. h. „Warme“. Die Farben sind somit von ihrer „Temperatur“ her eher neutral. Im Farbwechsel zwischen den Sequenzen „Ferrari“ und „Ruderer“ dokumentiert sich folglich kein Kalt-Warm-Kontrast.
- 3) **Komplementär-Kontrast:** Dieser Kontrast ist im Alltag der bekannteste. Die Farben Gelb und Violett, welche sich in den Sequenzen „Ferrari“ und „Ruderer“ wiederfinden, stellen klassische Komplementärfarben dar. Komplementärfarben schließen sich gegenseitig spektral vollständig aus und stehen deshalb auf Ittens Farbkreis direkt gegenüber. Der Komplementärkontrast tritt am stärksten zwischen gesättigten Farben in Erscheinung. Im vorliegenden Fall ist der Komplementärkontrast abgeschwächt, da dem reinen Gelb des Ferraris nur das schwache Hellblau-Violett des Ruderers gegenübersteht.
- 4) **Simultan(Sukzessiv)-Kontrast:** Dieser Kontrast ist für das Verhältnis zwischen „Ferrari“ und „Ruderer“ im vorliegenden Fallmaterial als einer der wesentlichsten anzusehen. Er verstärkt den Effekt aller bisher genannten und folgenden Kontraste. Laut Itten (2000, 2010) bewirkt der Simultan- bzw.

Sukzessivkontrast, dass bleiche und neutrale Farbtöne in Gegenwart gesättigter Farben in deren jeweiliger Komplementärfarbe erscheinen. Im vorliegenden Fall führt der Simultan- bzw. Sukzessivkontrast dazu, dass die blassen Hellblau-Violett-Töne der Sequenz „Ruderer“ gegenüber der knallgelben Sequenz „Ferrari“ visuell in der Komplementärfarbe Violett angereichert werden, was letztlich den bereits festgestellten Komplementärkontrast zwischen den beiden Sequenzen verstärkt.

- 5) **Quantitäts-Kontrast:** Dieser Kontrast hängt mit der Tatsache zusammen, dass die Farben des Spektrums unterschiedliche Helligkeitswerte besitzen. Wie Itten (2000, 2010) in Bezug auf Goethes Farbenlehre (2003) erklärt, ist Gelb die hellste, Violett die dunkelste aller Farben. Ein markanter Hell-Dunkel-Kontrast zwischen den Sequenzen wird im vorliegenden Fall vermieden, indem Violett auf das Helligkeitsniveau von Gelb angehoben ist. Da Violett durch Aufhellung jedoch seine Sättigung verliert, erscheint die Ruderer-Sequenz zwangsläufig blass und farblos.

Zusammenfassend ergibt sich anhand der analytischen Differenzierung der Farbkontraste im Zusammenhang mit den Einstellungswechseln zwischen den Sequenzen „Ferrari“ und „Ruderer“ folgender empirischer Befund: Obwohl die beiden Sequenzen fast dieselbe Helligkeit besitzen, entsteht aufgrund der Komplementärfarben Gelb-Violett ein farblicher Gegensatz zwischen ihnen. Dieses Verhältnis ist alles andere als ausgewogen. Es dominiert der satte gelbe Ferrari über den blass-violetten Ruderer. In direkter Gegenüberstellung mit der dritten Sequenz „Sänger“ wird deutlich, dass der gelbe Ferrari auch diese farblich überstrahlt. Hinsichtlich ihrer blassen Farbgebung besteht damit eine augenscheinliche, aber auch formal rekonstruierbare Homologie zwischen „Sänger“ und „Ruderer“. In puncto Sättigung sind diese beiden Sequenzen dem bildfüllend präsentierten knallgelben Ferrari eindeutig untergeordnet.

Bei differenzierter Betrachtung der Farbkontraste wird im vorliegenden Musikvideo durch die Montage folglich eine klare Hierarchie zwischen den drei untersuchten Sequenzen aufgeworfen, an deren Spitze die Sequenz „Ferrari“ steht. Die farblich gesehen blasseren Sequenzen „Ruderer“ und „Sänger“ sind dieser untergeordnet, wobei jedoch auch deren Verhältnis nicht völlig gleichwertig ist. Die Sequenz „Ruderer“ ist farblich als direkter komplementärer Gegenpart zur Sequenz „Ferrari“ konzipiert und in dieser Hinsicht fokussiert. Demgegenüber steht die Sequenz

„Sänger“ – in neutralen Farbtönen gehalten – scheinbar außerhalb dieses Farbduells. Dass dem nicht so ist, wird bei genauerer Betrachtung des Fotogramms 10 (Abbildung 44) deutlich, welches im Rahmen der Montage durch eine gleichwertig *Überblendung* der Sequenzen „Ferrari“ und „Ruderer“ zustande kommt.



Abbildung 54: Farbverhältnisse der durch die Montage verbundenen Sequenzen „Ferrari“, „Ruderer“ und „Sänger“

Durch die Überblendung der beiden Sequenzen entstehen Grün- und Brauntöne sowie eine horizontale Gliederung des Bildes. Beides – Farbgebung und Planimetrie – weisen starke Homologien zum Fotogramm der Sequenz „Sänger“ auf. Damit steht die Sequenz „Sänger“ formal gesehen keineswegs außerhalb des Geschehens, sondern kann durchaus als Synthese der aufgeworfenen Gegensätze (gelb – violett, reich – arm etc.) aufgefasst werden. Zugleich verdeutlicht der Vergleich der beiden letzten Bilder nochmals, wie die Person des Sängers aufgrund ihrer etwas eigentümlichen Birnenform aus dem Ensemble der Mischfarben herausgehoben wird.

## **8.7 Ikonologische bzw. ikonische Interpretation**

Auf den ersten Blick ist das vorliegende Musikvideo „Araba“ (1996) des türkischen Popstars Mustafa Sandal in hohem Maße durch die Polarität zwischen zwei antagonistischen Prinzipien bestimmt. Auf der Bildebene findet dies paradigmatisch durch die Komplementärfarben Gelb und Violett seinen Ausdruck, die in fester Koppelung mit den Sequenzen „Ferrari“ und „Ruderer“ verbunden sind. In Kombination mit einer hohen Schnittfrequenz (teilweise ein Schnitt pro Sekunde), die mit der Parallelmontage der Sequenzen einhergeht, vollzieht sich ein Feuerwerk an Reizwechseln, welches auch auf inhaltlicher Ebene seine Entsprechung findet: In der Eingangspassage des Musikvideos wird ein Wettrennen vorbereitet, das zwischen ungleichen Gegnern in zwei Fortbewegungsarten zu Lande und zu Wasser ausgetragen werden soll. Ein PS-starker Ferrari tritt gegen einen einfachen Ruderer an. Bedingt durch die dramaturgische Zuspitzung auf das farbliche und fahrerische Duell der beiden Kontrahenten fällt erst bei näherer Betrachtung auf, wie stark sie andere Teile des Videos maskieren und in den Hintergrund drängen – allen voran den eigentlichen Frontmann des Musikvideos, den türkischen Popstar und Sänger Mustafa Sandal.

### **8.7.1 Exklusivität und Durchschnittlichkeit**

An sich ist Mustafa Sandal prominent ins Bild gesetzt: Gleich zu Beginn wird er als erste Person im Video eingeführt; und das sogar über einen ausgedehnten Zeitraum von sechs Sekunden hinweg (Abbildung 44, Fotogramme 2–7). Auch im weiteren Verlauf des Videos kommt er wieder vor (Abbildung 44, Fotogramme 27–28). Außerdem wurde für ihn ein exklusiver Bildausschnitt gewählt: Der Sänger ist zentral, frontal, vollständig und symmetrisch ins Bild gesetzt, wodurch er auf der Ebene der Einstellung mehrfach fokussiert ist. Wenn Personen im Bild allein und mittig abgebildet sind, verleiht ihnen

dies üblicherweise eine besondere visuelle Geltung<sup>89</sup>. Hinzukommt, dass der Sänger im Tonraum nahezu ständig über seine Stimme präsent ist.

Paradoxerweise ist der Sänger aber trotz der erwähnten Mehrfachfokussierung weder auf der Ebene der Montage noch auf der Ebene der Einstellung bildbestimmend. Sein gesamter Auftritt hat den Gestus des Beiläufigen und Unauffälligen. Dies hat unter anderem mit seiner Körperhaltung, seiner Kleidung und der Umgebung im Bild zu tun: Der Sänger sitzt in Alltagskleidung am Straßenrand vor einem formlosen und verblassten Hintergrund. Im Bild herrschen verwaschene Braun-, Grün- und Grautöne vor, die jenen Farben sehr nahe kommen, die sich aus der Überblendung der gelben Ferrari-Sequenz mit der violetten Ruderer-Sequenz ergeben (Abbildung 54). Es hat den Anschein, als wäre der Sänger des Musikvideos seitens der abbildenden Bildproduzent/inn/en bewusst *wohltemperiert* und in aller (farblichen) *Durchschnittlichkeit* in Szene gesetzt. Doch worin liegt der tiefere Sinn dieses Understatements? Immerhin handelt es sich im Falle des hier abgebildeten Bildproduzenten um einen der bekanntesten und erfolgreichsten türkischen Popstars.

### 8.7.2 Teilhabe und Distanzierung

Möglicherweise erhöht die zurückhaltende Präsentation den Status des Sängers Mustafa Sandal über einen Umweg, indem sie als vornehme Schüchternheit aufgefasst werden kann: Als erfolgreicher Popstar hat man es nicht nötig, sich an selbstwerterhöhenden Wettkämpfen zu beteiligen. Anstatt sich im sportlichen Wettkampf selbst die Hände schmutzig<sup>90</sup> zu machen, kann Mustafa Sandal es sich somit leisten, von der Bühne bzw. *Tribüne* aus zuzusehen. Dies dokumentiert sich im Bild konkret darin, dass der Sänger es sich entspannt auf erhöhtem Platz *über* dem Straßenniveau<sup>91</sup> gemütlich gemacht hat. Von dort aus kann er das *bunte* Treiben am Bosphorus gut überblicken. Zusammenfassend ist die eingenommene Mittelposition des Sängers zwischen den anderen beiden Sequenzen bzw. Handlungssträngen als ein besonderes *Außen* zu verstehen, welches dem Sänger – im Sinne einer Übergegensätzlichkeit – zugleich

---

<sup>89</sup> Im Rahmen der gattungsimmanenten und gattungsübergreifenden komparativen Analyse der TV-Show „Istanbul Total“ habe ich die mittige Positionierung des Moderators im Bild als typisches Merkmal von Late-Night-Shows im Fernsehen herausgearbeitet (Kap. 5.2.1). Diese Art der Positionierung kann nach Bohnsack (2009, S. 240) sogar als Hinweis für eine „Monostrukturierung“ der Sendung durch den abgebildeten Bildproduzenten/die abgebildete Bildproduzentin und als „Hyperzentrierung“ auf dessen/deren Person gewertet werden.

<sup>90</sup> Diese hängen in Abbildung 49 locker herab, während sie Ruderer und Ferrarifahrer *an ihren (Sport-)Geräten* haben.

<sup>91</sup> Dies dokumentiert sich im Video durch die vorbeifahrenden Autos (Abbildung 44, Fotogramm 5).

Teilhabe und Distanzierung sowie ein gleichzeitiges Dabeisein und Darüberstehen ermöglicht.

Homolog zu dieser Übergegensätzlichkeit zeigt sich der Sänger nicht, wie dies etwa in seinen anderen Musikvideos<sup>92</sup> der Fall ist, auffälliger gekleidet und pompöser – d. h. für einen großen Popstar standesgemäß – ins Bild gesetzt ist: Seine spezielle Mittlerposition würde nämlich durch knalligere Farben bzw. ausgefallenerer oder eleganterer Kleidung korrumpiert. Im Falle einer kräftigeren Farbgebung würde sich automatisch der Kontrast zwischen dem Sänger und den anderen beiden Sequenzen erhöhen, wodurch der zentrale Gestus des Musikvideos, welcher in der Dramatisierung des Duells zwischen Ferrari und Ruderer liegt, zerstört würde. Spielte der Sänger eine aktivere Rolle in diesem Geschehen, stünden sich plötzlich *drei* anstatt nur zwei Parteien in der Wettkampfarena am Bosphorus gegenüber. Nach den Regeln der Kombinatorik hätte diese kleine Veränderung durchaus große Konsequenzen. Sie würde die Komplexität des antizipierbaren Wettkampfausgangs drastisch erhöhen. Während bei einem Duell prinzipiell davon auszugehen ist, dass der Bessere gewinnt, ist das Ergebnis eines Triells nicht klar vorhersehbar<sup>93</sup>. Durch seinen vornehmen Rückzug auf *die Ränge* der Austragungsarena bzw. seinen *Rang* als Popstar wird dieses drohende Szenario entschärft. Indem sich der Sänger der direkten, potentiell gefährlichen Auseinandersetzung entzieht, erhält er zugleich ein höheres Maß an Kontrolle über das Geschehen. Von seinem bequemen *Sitzplatz* aus kann der Sänger Mustafa Sandal nun die Ereignisse aus sicherer Distanz beobachten und (musikalisch) kommentieren – wohl wissend, wer letztlich als Sieger hervorgeht, wenn sich die zwei anderen vor seiner Nase streiten<sup>94</sup>.

### 8.7.3 Gesicht und Körperlichkeit

Es passt zur rekonstruierten besonderen Rolle Mustafa Sandals im Video, dass er – im Gegensatz zu Ruderer und Ferrarifahrer – den Bildbetrachter/innen stets frontal zugewandt ist. Wie ein Blick auf die betrachteten Sequenzen bzw. das Videotranskript (Abbildung 44) verdeutlicht, ist der gelbe Ferrari zwar am häufigsten und überaus prominent ins Bild gesetzt, der *Fahrer* des Ferraris ist dabei jedoch zumindest in diesem

---

<sup>92</sup> Bspw. „Pazara Kadar“, „Indir“, „All my Life“ etc.

<sup>93</sup> Beispielsweise kann auch ein schwacher Gegner ein Triell gewinnen; sogar mit relativ hoher Wahrscheinlichkeit.

<sup>94</sup> An dieser Stelle besteht noch eine interessante Homologie zur Denkfigur bzw. dem mathematischen Problem des Triells. Paradoxe Weise geht nämlich beim Triell tatsächlich der schwächste Gegner mit höchster Wahrscheinlichkeit als Sieger aus der Konfrontation hervor.

Transkriptausschnitt kein einziges Mal zu sehen. Er wird als Person nur über die Indexikalität seiner Hände am Lenkrad (Fotogramme 37 und 39) konstituiert. Erst durch diesen Kunstgriff kann in weiterer Folge (mithilfe der Montage) im Video überhaupt der rekonstruierte Wettkampf zwischen zwei menschlichen Kontrahenten aufgebaut werden<sup>95</sup>.

Im Kontrast zum Ferrarifahrer tritt der Ruderer im Bild weit sichtbarer in Erscheinung. In Parallelmontage werden auch seine Hände gezeigt (Abbildung 44, Fotogramm 38), um auf der Ebene der szenischen Choreografie den direkten Vergleich zum Ferrarifahrer herzustellen. Über weite Strecken hinweg ist aber ebenfalls der restliche (bekleidete) Körper des Ruderers in Halbtotale (Abbildung 49) zu sehen. Ähnlich wie der Ferrarifahrer entzieht er sich dabei aber dem direkten Blick der Bildbetrachter/innen: Der Ruderer ist ihnen zu keinem Zeitpunkt frontal zugewandt, sondern stets nur von hinten oder von der Seite zu sehen.

Die Person des Sängers hingegen sitzt scheinbar unverstellt vor uns und verleiht dem Musikvideo ein menschliches Gesicht (Abbildung 52). Dieses Gesicht hält der Popstar Mustafa Sandal frontal in die Kamera. Daneben ist auch sein restlicher (bekleideter) Körper in Totalaufnahme vollständig erkennbar. Echte Nahaufnahmen des Gesichts werden vermieden. Ein Grund dafür könnte sein, dass Mustafa Sandal durch eine solche Detailfokussierung seine bisher im Video rekonstruierte übergegensätzliche Mittlerposition verlieren würde. Obwohl das Gesicht des Sängers nie in Nahaufnahme zu sehen ist, demonstriert es eine große Bandbreite an Ausdrucksmöglichkeiten (vgl. Abbildung 44): von abwesend unbeteiligt (Fotogramme 2–4) über lachend (Fotogramme 5–7) bis zu nachdenklich (Fotogramme 27–28). Dieser Reichtum an Mimik haucht Mustafa Sandal – insbesondere in direkter Gegenüberstellung zu den zwei anderen Akteuren bzw. Sequenzen des Videos – Leben bzw. eine Seele ein. Seine schlichte Kleidung und die zentral-frontale Kameraeinstellung tragen weiter dazu bei, dass er menschlich und alltäglich (auf Augenhöhe) erscheint. Dadurch kann er leichter zur Identifikationsfigur für die Bildbetrachter/innen werden.

---

<sup>95</sup> Würden die Hände des Fahrers am Lenkrad fehlen, so könnte im vorliegenden Musikvideo auch der Eindruck entstehen, dass es sich beim Ferrari um einen computer- oder ferngesteuerten Gegner wie in der populären Science-Fiction-Serie *Knight Rider* handelte.

#### 8.7.4 Akteure und deren Extensionen

Die anderen untersuchten Protagonisten (Ruderer und Ferrarifahrer) sind nicht nur aufgrund ihrer körperlichen Abgewandtheit bzw. Verhüllung schwerer zugänglich<sup>96</sup>. Sowohl der Ruderer als auch der Ferrarifahrer stellen im Latourschen Sinne (2008) „Hybridakteure“<sup>97</sup> dar: Sie werden letztlich über ihre technischen Prothesen Ruderboot und Ferrari definiert, mit deren Hilfe sie sich fortbewegen bzw. präsentieren. Der Sänger kommt als Einziger ohne technische Hilfsmittel aus. Dafür ist er an seinen Sitzplatz gebunden.

Im Zusammenhang mit den Denkfiguren von Hybridakteuren und von Instrumenten stellt sich des Weiteren die Frage, ob Mustafa Sandals leere Handflächen – im Kontrast zu den beschäftigten Händen des Ferrarifahrers und Ruderers – nicht noch auf andere (fehlende) Instrumente verweisen. In Anbetracht seiner musikalischen Fähigkeiten wäre in dieser Sitzposition durchaus vorstellbar, dass Mustafa Sandal ein Musikinstrument, wie eine Gitarre oder eine kleine Trommel, hält. Das Fehlen auch dieser „Geräte“ verweist darauf, dass er keinerlei technische Hilfsmittel braucht. Sein eigentliches Instrument ist unsichtbar; es ist die Stimme. Im Video kann die Stimme des Sängers auf der Bildebene nur durch den Gesichtsausdruck repräsentiert werden, was wiederum erklärt, warum sein Gesicht – im Gegensatz zu dem der anderen Akteure – gleich zu Beginn des Musikvideos zu sehen ist. Doch mit Ausnahme des herzhaften Lachens bewegen sich die Mundwinkel des Sängers nicht. Er singt nicht, sondern bleibt stumm. Worauf wartet Mustafa Sandal?

#### 8.7.5 Warten auf die Stellvertreter

Die ersten vierzig Sekunden des Musikvideos „Araba“ haben die Funktion darzustellen, wie die untersuchten Akteure sich und ihre Geräte in Bereitschaft versetzen. Dabei werden gleichzeitig die qualitativen Unterschiede zwischen den Protagonisten herausgearbeitet. Der Sänger ist ab dem ersten Fotogramm bereit. Mit seinem unsichtbaren Instrument – der Stimme – hat er seinen Platz eingenommen und macht keinerlei Anstalten, ihn zu räumen. Er sitzt unverrückbar in seiner Position, schaut und lacht, während Ferrarifahrer und Ruderer erst *Hand* an ihre Fortbewegungsgeräte legen müssen, um sie in Bereitschaft zu versetzen. Durch die Parallelmontage wird räumliche

---

<sup>96</sup> Bei genauer Betrachtung haftet ihnen im Bild auch ein gewisses Stigma der Prothesenabhängigkeit an. Was wären sie ohne ihre maschinenartigen Krücken?

<sup>97</sup> vgl. auch Schäffer (2001, S. 54).

Nähe zwischen den Protagonisten hergestellt: Der Sänger sitzt an einer befahrenen Straße, die vermutlich am Ufer des Bosphorus liegt. Mustafa Sandal wartet also, bis die anderen so weit sind.

Der Umstand, dass Mustafa Sandal nicht einfach zu singen beginnt, bevor die anderen Handlungsakteure (Ferrarifahrer und Ruderer) entsprechend eingeführt sind, verdeutlicht, dass diese für die Präsentation seiner eigenen Person offensichtlich erforderlich sind. Der von ihnen ausgetragene Wettkampf, der gesellschaftliche Dimensionen gewinnt, tangiert ihn nicht. Er positioniert sich jenseits von *Arm* und *Reich*, jenseits des Wettkampfes von Selbstpräsentationen, die sich einerseits auf die Authentizität eigener Muskelkraft und andererseits auf PS-Stärke und Status-Symbole stützen. Dabei scheint die rekonstruierte Gesichtslosigkeit von Ferrarifahrer und Ruderer anzuzeigen, dass es sich in ihrem Falle nicht um reale, sondern imaginierte Stellvertreter handelt. In diesem Zusammenhang entwickelt das Warten des Sängers eine kontemplative Qualität und verweist – hinter der ruhigen Gesichtsfassade – auf eine lebhaft psychische Auseinandersetzung *in seinem Inneren*. Diese innere Auseinandersetzung ist in ähnlicher Weise wirksam und unsichtbar wie das ephemere Ausdrucksinstrument seiner Stimme<sup>98</sup>. Das schüchterne und zurückhaltende Auftreten des Sängers korrespondiert mit einer bunten und kontrastreichen Innenwelt, die ihn psychisch beschäftigt und von der Außenwelt trennt. Wie das Musikvideo suggeriert, kann wer es oft genug sieht und diesen komplexen Mechanismus entschlüsselt, zur Belohnung hinter Mustafa Sandals äußerer Versunkenheit einen tieferen Einblick in die Seele des Popstars erhaschen.

#### **8.7.6 Die Pose des Geschichtenerzählers**

Jemand, der eine reiche Innenwelt hat, hat auch viel zu erzählen – insbesondere dann, wenn er sitzt. Der Schneidersitz ist in vielen Kulturkreisen als konventionalisierte Körperhaltung fürs Geschichtenerzählen verbreitet. Zwar wird und wurde diese Pose aus kulturhistorischer Perspektive auch für Handarbeitstätigkeiten, zum Musizieren und Meditieren etc. eingenommen, keiner dieser Aspekte dokumentiert sich jedoch im vorliegenden Material. Auf historischen und zeitgenössischen Abbildungen sind Geschichtenerzähler häufig in sitzender Haltung abgebildet, die jener von Mustafa Sandal ähnlich sind.

---

<sup>98</sup> In diesem Sinne könnten das Gesicht des Sängers und seine leeren Hände auch als die äußeren Projektionsflächen seiner innerer Stimmungen und Konflikte gewertet werden.

Der Begriff der Pose wurde bereits früh im Rahmen der dokumentarischen Bildinterpretation aufgegriffen und hat in diesem Zusammenhang bis heute hohe Relevanz zur Interpretation der Leistungen abgebildeter Bildproduzent/inn/en (insbesondere von deren Gesten und Gebärden) gehabt (Bohnsack & Przyborski, 2014; Bohnsack, 2001b, 2009).



**Abbildung 55: Der Geschichtenerzähler, „Verstehen und Gestalten“, Oldenbourg, München, Beate Speck-Kafkoulas**

Die Pose des Schneidersitzes (vgl. Abbildung 55), bei der die Beine immobilisiert und entlastet sowie die Hände frei sind, scheint für den Geschichtenerzähler besondere Vorteile aufzuweisen. Zum einen ist diese Sitzposition komfortabel, zum anderen können die Hände gestikulativ zur Unterstreichung der Erzählung eingesetzt werden. Für Sänger ist das Sitzen eher eine untypische Haltung, da dabei nicht das volle Stimmvolumen zur Geltung gebracht werden kann. In Bezug auf Imdahl (1996b) kann der habituelle Bruch bei der Einnahme einer Körperhaltung nach Bohnsacks und Przyborskis (2014) als „De-Kontextierung“ bezeichnet werden.

Als weiterer Bruch im Sinne einer De-Kontextierung könnte gewertet werden, dass Mustafa Sandal in der Pose des Geschichtenerzählers schweigt. Anstelle von Worten bedient er sich kontrastreicher Bilder. Bisher konnten wir rekonstruieren, dass die Geschichte von Ferrarifahrer und Ruderer nicht abgekoppelt von der Person des Sängers existiert, sondern offenbar *seine* eigene Lebensgeschichte oder zumindest Lebensthemen von persönlicher Relevanz und zugleich von potentieller Relevanz für das Publikum oder Milieus innerhalb dessen porträtiert. Die bildervermittelte Narration ließe sich folgendermaßen paraphrasieren: Mustafa Sandal, der millionenschwere

Superstar, sitzt in einfacher Kleidung am Straßenrand und fragt sich tief in seinem Herzen, ob es erfolgversprechender sei, Reichtümer und Luxus zu genießen oder sich durchs Leben zu kämpfen. Zum Beweis der Aufrichtigkeit dieser Geschichte zeigt uns der Sänger sein Gesicht und seine bloßen Hände. Doch sein abgewandter Blick, die verdeckten Handflächen und das Lachen (Abbildung 52) lassen Zweifel an der vorgegebenen Aufrichtigkeit entstehen. Der abgewandte Blick und die verdeckten Handflächen können als Hinweise dafür aufgefasst werden, dass der Sänger etwas verbergen will. Dies gibt seinem Lachen den Gestus eines „wissenden Lachens“, eines Lachens, das auf einem Informationsvorsprung gegenüber den Bildbetrachter/inne/n und dem Amüsement über die geschickte Inszenierung und deren Posenhaftigkeit beruht. Der Geschichtenerzähler steht jenseits dieser Inszenierungen.

Die bemerkenswerte Botschaft des schelmischen Lachens von Mustafa Sandal scheint in diesem Zusammenhang wiederum von unglaublicher Offenheit geprägt zu sein: Traue nicht dem, was du siehst! Er steht über dem Posieren und der Selbstinszenierung dieser Milieus und über den zwischen ihnen ausgetragenen Kämpfen. Und hierin liegt möglicherweise gerade der Schlüssel für seine Popularität.

### **8.7.7 Die Übergegensätzlichkeit von Authentizität und Inszenierung**

Die zusammenfassende Übergegensätzlichkeit des Musikvideos „Araba“ (1996) besteht somit im schelmischen Spiel von Authentizität und Inszenierung. Der Sänger Mustafa Sandal entpuppt sich als sympathischer Trickster, der durch seinen diskreten Auftritt die Zusehenden einerseits von sich ablenkt und andererseits auf sich aufmerksam macht, ihre Aufmerksamkeit geradezu bannt. Als einzige Person, die ihr Gesicht zeigt und dabei vollkommen sich selbst genügt, wird er zur zentralen Identifikationsfigur des Videos. Seine äußerliche kontemplative Ruhe gegenüber den sichtbar energiegeladenen Wettkampfsequenzen der anderen Akteure löst Spekulationen über seine inneren Konflikte aus. Im Mechanismus der evozierten Anteilnahme an seinem persönlichen Befinden avanciert sein Gesicht zum Echtheitszertifikat, in dem sich die Frage nach seiner Aufrichtigkeit kristallisiert. Was sollen wir von seinem legeren und unaufgeregten Auftritt halten? Immerhin ist Mustafa Sandal ein millionenschwerer Popstar. Diese Zuschreibung wird durch die abbildenden Bildproduzent/inn/en systematisch unterbunden, indem sie alles, was mit Reichtum, Luxus und Protz verbunden ist, mit dem gelben Ferrari im Bild assoziieren. Diesem gegenüber wird Mustafa Sandal durch die mit ihm assoziierten Grün-, Braun- und Grautöne rein formal schon als bescheiden

porträtiert. Als die größte Leistung der abbildenden Bildproduzent/inn/en ist es in diesem Zusammenhang anzusehen, dass es ihnen gelingt, eine mehrdeutige Zuschreibungsspirale in Gang zu setzen, die letztendlich sicherstellt, dass für die Zusehenden kein Weg an der Person des Popstars Mustafa Sandal als abgebildeten Bildproduzenten vorbeiführt. Dieses System ist dermaßen ausgefeilt, dass selbst der mimische Hinweis des Sängers, bei dem Musikvideo würde es sich um ein inszeniertes Posieren handeln, letztendlich wieder auf ihn selbst zurückverweist. Zusammenfassend könnte man sagen, dass dies alles dazu beiträgt, die Marke Mustafa Sandal in der Popszene fest zu verankern und erfolgswirksame Folgekommunikationen in den Zielgruppen auszulösen.

Zum Abschluss dieser Videointerpretation erhellt dieser letzte Befund sogar noch zwei weitere Aspekte des Videos, die zuvor nicht recht zu deuten waren. Zum einen möchte ich abschließend noch einmal auf den Filmcountdown bzw. – in der Sprache der Filmwissenschaft – das sogenannte *Startband* zurückkommen, das gleich zu Beginn des Videos zu sehen ist. Es wurde im Zusammenhang mit der Rekonstruktion der Diegese des Films als Bruch mit dem restlichen Video rekonstruiert, ohne dass dieser Bruch bisher näher erörtert oder erklärt werden konnte. Das Startband dient bei der traditionellen Projektion von Spielfilmen in Kinosälen zur Synchronisation von Bild- und Ton sowie zur Einrichtung des Bildausschnitts. Für ein Musikvideo, das über das Fernsehen oder im Internet verbreitet wird, ist der Einsatz eines Startbandes jedoch völlig überflüssig und funktionslos. Die einzige Funktion des Filmcountdowns im vorliegenden Musikvideo besteht in dem indexikalen Hinweis, dass nun ein Film folgt. Denselben Hinweis geben im Verlauf des gesamten Videos die schwarzen Balken, die sogenannte Letterbox, an der Ober- und Unterkannte des Bildes. Auch Letterboxes erfüllen ihre technische Funktion nur bei der TV-Ausstrahlung von Kinofilmen. Zusätzlich geben sie vor, dass es sich im Falle des Musikvideos „Araba“ (1996) um einen Spielfilm handle. Wie die Interpretation des Fallmaterials jedoch deutlich gemacht hat, kann auch diese Zuschreibung ihrerseits nur als letztes Mosaiksteinchen jener komplexen Inszenierung aufgefasst werden, die das gesamte Musikvideo bestimmt.





MUSTAFA SANDAL  
ARABA

**Abbildung 56: CD-Cover der Single „Araba“ (2004) für den internationalen Markt, Polydor, Universal Music Group**

Quelle: <http://www.7digital.com/artist/mustafa-sandal/release/araba-2>

## 9 „Araba“ (2004): Die Version für den internationalen Markt

### 9.1 Auswahl des Fallmaterials zur komparativen Analyse

In einer Episode von „Istanbul Total“<sup>99</sup> (Kap. 5) wird der türkische Popsänger Mustafa Sandal eingeladen, um eine neue Version seines Musikvideos „Araba“ (vgl. Kap. 8) zu präsentieren. Während des Interviews erklärt der Sänger, dass dieses Musikvideo eigentlich auf einem alten Song, „Araba“ (zu Deutsch: „Auto“), basiere, welcher 1996 in der Türkei<sup>100</sup> einer seiner größten Erfolge gewesen sei. Nachdem er einen Vertrag mit der deutschen Musikfirma Polydor (Teil der Universal Music Group<sup>101</sup>) abschließen konnte und 2003 mit einem weiteren Lied „Aya Benzer“ („Mondlicht“) die Top Ten der deutschen Musikcharts erreichte, entschloss sich Polydor 2004 „Araba“ in einer Neuauflage auf den internationalen Markt zu bringen. Wie Mustafa Sandal weiter erklärt, beinhaltete dies auch die Produktion eines neuen Musikvideos: „Das alte Video für ‚Araba‘ hätte in Deutschland gewiss nicht bestanden. Also drehten wir ein neues in Berlin.“ Was die Auswahl des Fallmaterials für dieses Kapitel betrifft, wurden beide Versionen des Musikvideos für eine komparative Interpretation herangezogen.

Abbildung 44 zeigt das Videotranskript der ersten vierzig Sekunden der Originalversion des Musikvideos „Araba“ von 1996, das für den türkischen Binnenmarkt produziert wurde.<sup>102</sup> Abbildung 57 zeigt dieselbe Art von Videotranskript für das Musikvideo „Araba“, das 2004 für den internationalen Markt neu gedreht wurde.<sup>103</sup> Bevor ich auf die formulierende Interpretation der zwei Videoverversionen eingehe, möchte ich die Aufmerksamkeit auf deren formale Kompositionen bezüglich Schnitt bzw. Montage lenken. Dies ist ein notwendiger Schritt, um die Struktur des Mediums Video zu verstehen, welches durch seine bloße Form die Hervorbringung eines spezifischen Common Sense auf der ikonographischen Ebene unterstützt. Die Suche nach biografischen Brüchen und ihren Reparaturen auf der Inhaltsebene weist, metaphorisch gesprochen, eine Homologie zur Untersuchung von Schnitten und deren Ausbesserungen auf dem „physischen“ Niveau des Videomaterials auf. Diese Analogie kann gleichermaßen unser Verständnis für die Beziehung zwischen biografischen Brüchen und deren „Reparaturen“ erweitern: Im Prozess der Videomontage sind Brüche

---

<sup>99</sup> Ausstrahlung am 12. Mai 2004.

<sup>100</sup> Der Musikfirma Sony Music France zufolge wurden ca. 2.000.000 Kopien des Albums „Araba“ verkauft.

<sup>101</sup> Zusammen mit Sony Music Entertainment und der Warner Music Group ist die Universal Music Group eine der drei sogenannten „Hauptlabels“. Hauptlabels sind die größten Unternehmen der Musikindustrie, die den weltweiten Markt für populäre Musik kontrollieren.

<sup>102</sup> Link zur türkischen Version des Videos: <http://www.youtube.com/watch?v=b0m8YPF3N70>

<sup>103</sup> Link zur internationalen Version des Videos: <http://www.youtube.com/watch?v=QlvPQbYKLww>

und Reparaturen untrennbar miteinander verbunden. Tatsächlich macht laut Filmtheorie (z. B. Mikos, 2003) das dynamische Verhältnis von Schneiden und Ausbessern Video und Film erst zu unabhängigen Medien.

Passage (oder Sequenz): Eingangspassage  
 Film (oder Video): „Araba“ (2004) für den internationalen Markt  
 Dateiname: <http://www.youtube.com/watch?v=QlvPQbYKLww>  
 Datum: 2004  
 Time Code: 0-233  
 Dauer: 3:53 Min.  
 Transkription: Stefan Hampl  
 Korrektur: -



Abbildung 57: Videotranskript des Musikvideos „Araba“ (2004) von Mustafa Sandal, Universal Music Group

## 9.2 Die Rekonstruktion der Montage der Musikvideos „Araba“ 1996 und 2004

Wie in dieser Arbeit bereits erläutert und demonstriert, lässt sich das Videomaterial von Fernsehshows Bohnsack (2009) in eine hierarchische Struktur von Hauptsequenzen, Untersequenzen und eingelagerten Sequenzen unterteilen. Dem Segmentierungsprozess folgend können repräsentative Standbilder und Fotogramme ausgewählt und mit den Werkzeugen der komparativen Bildinterpretation detailliert analysiert werden. Aufgrund der hohen Zahl von Aufnahmen und Schnitten fordern Musikvideos die skizzierte Standardprozedur der Videointerpretation erheblich heraus. Wie die Videotranskripte der Musikvideos (Abbildung 44 und Abbildung 57) zeigen, können in beiden Videoverversionen schon während der ersten vierzig Sekunden ungefähr zwanzig Schnitte ausgemacht werden. Obwohl es möglich ist, eine wörtliche Beschreibung der einzelnen Aufnahmen sowie die Segmentierung vorzunehmen, sträubt sich das Videomaterial gegen alle Versuche, es überzeugend in Bezug auf Haupt-, Unter- oder eingelagerte Sequenzen zu beschreiben.

Bohnsack (2009, S.158) nennt zwei wesentliche Argumente für die Notwendigkeit der Entwicklung empirischer Werkzeuge zur Direktinterpretation der Montage: Wie bereits in Kapitel 2.2 erwähnt, ist nach Auffassung der Filmwissenschaften (z. B. Mikos, 2003; Pudowkin, 1996; Wiedemann, 2005) die Montage das zentrale Alleinstellungsmerkmal von Film und Video (bspw. ggü. Fotografien). Des Weiteren stellt die Montage „eindeutig das Produkt der Leistungen der abbildenden Bildproduzent/inn/en dar“ und eröffnet uns dadurch den exklusiven „Zugang zu deren modus operandi“ (Bohnsack, 2009, S. 162).

Zur Einführung in die Interpretation des Montageprozesses ist es wichtig, daran zu erinnern, dass in den ersten vierzig Sekunden beider zur Analyse herangezogener Videoverversionen ungefähr zwanzig Schnitte ausgemacht werden können. Diese Anzahl verweist auf eine hohe Dichte von Schnitten; bei genauer Betrachtung der Videotranskripte (Abbildung 44 und Abbildung 57) sind die Schnitte aber nicht gleichmäßig verteilt. Tatsächlich ist bei beiden Videos die Zeit zwischen den Schnitten am Anfang länger und gegen Ende der Sequenzen kürzer. Die zunehmende Dichte der Schnitte kann laut Mikunda (2002) durch professionelle Montagetechniken des *Preluding* erklärt werden: Als Basisbestandteile von visuellem Geschichtenerzählen müssen Orte und Charaktere zu Beginn vorgestellt werden, was Zeit erfordert. Nachdem der Begriff Prelude schon aus der Musik abgeleitet wurde, lohnt sich auch Betrachtung

der Musikspur der Videos – analog zu den visuellen Entwicklungen am Bildschirm beginnt die Musik mit einer langsamen Melodie, bevor sie energetisch und rhythmisch wird. In Abbildung 44 vollzieht sich diese Veränderung der Musikgeschwindigkeit bei 29 Sekunden, in und Abbildung 57 bei 17 Sekunden. Wie den Videotranskripten entnommen werden kann, werden die Schnitte in beiden Videoverversionen nach diesem audiovisuellen Intro häufiger. Damit können die ersten Brüche in den zwei Videoverversionen als jene identifiziert werden, die das Vorstellen der zentralen Charaktere und Orte von deren weiterer Interaktionsentwicklung trennen. In der Folge möchte ich die Intros der beiden Videoverversionen nun bezüglich ihrer formalen Kompositionen näher vergleichen.

### **9.2.1 Montage als Ausdruck von Farbkontrasten**

Um die komplexe Montage von Musikvideos zu verstehen, ist es hilfreich, die Farbbeziehungen der Aufnahmen, d.h. die Montagemuster, die sich in den „Farbkontrasten“ (Itten, 2000, 2010) der Videos ausdrücken, zu rekonstruieren. In dieser Hinsicht können Videotranskripte (wie die in Abbildung 44 und Abbildung 57 dargestellten) als systematisch arrangierte Farbpaletten verstanden werden. Um die Farbkontraste, die sich während einer Film- oder Videobearbeitung herausgebildet haben, zu rekonstruieren, wird der Fokus nicht auf den Inhalt, sondern vorerst nur auf die formalen Farbsequenzen gelegt. Farben sind dadurch, dass sie visuelle Bedeutungen in Bezug auf Farbschattierung, Helligkeit und Sättigung aussenden, das umfassendste Mittel zur Schaffung von Atmosphären in Filmstreifen.<sup>104</sup> Beim ersten, oberflächlichen Blick auf die Videotranskripte scheint Abbildung 44 viel farbenfroher und auch heller als Abbildung 57. Warum ist das so? Diese Frage führt zu einem weiteren Untersuchungsschritt, der die Farbbeziehungen der Videoverversionen mit Ittens Farbkontrasttheorie in Einklang bringt. Hier wird die komplette Liste der sieben Farbkontraste nach Itten (2000, 2010) angeführt:

1. Farbe-an-sich-Kontrast
2. Hell-Dunkel-Kontrast
3. Kalt-Warm-Kontrast
4. Komplementär-Kontrast
5. Simultan-Kontrast
6. Qualitäts-Kontrast
7. Quantitäts-Kontrast

---

<sup>104</sup> Nach Seitz (2010) sind Spielfilme nach Genre farbkodiert.

### **9.2.1.1 Farbe-an-sich-Kontrast**

Im Rückgriff auf unseren ersten Eindruck von den zwei Musikvideos kann das Konzept der Farbkontraste dazu benutzt werden, die türkische und die internationale Version von „Araba“ systematisch zu unterscheiden. Die Frage, ob ein Video bunt ist oder nicht, kann beantwortet werden, indem der Farbe-an-sich-Kontrast der Montage betrachtet wird. Nach Itten (2000, 2010) ist der Farbe-an-sich-Kontrast der einfachste Farbkontrast. Um ihn in Erscheinung treten zu lassen, müssen mindestens drei helle und voll gesättigte Farben vorliegen. Anderenfalls verschwindet der Effekt zwar nicht, wird aber verringert. Der Farbe-an-sich-Kontrast ist am ausdrucksstärksten, wenn die drei Primärfarben Rot, Gelb und Blau zum Einsatz kommen. Zusätzlich kann der Effekt dadurch verstärkt werden, dass die Farben voneinander durch weiße oder schwarze Linien getrennt werden. Diesen differenziellen Richtlinien folgend, qualifiziert sich Abbildung 44 als ein gutes Beispiel des Farbe-an-sich-Kontrastes, obwohl es nicht alle Kriterien für diesen Kontrast erfüllt. Die intensivste Einzelfarbe im Video ist helles Gelb, das durch schwarze und weiße Formen verstärkt wird. Zudem gibt es hellrote Flecken und hellblaue wie auch hellviolette Bereiche; Violett kann als Mischfarbe identifiziert werden (eine Mischung aus den Primärfarben Rot und Blau). Im Gegensatz dazu beinhaltet Abbildung 57 Kontraste zweier Sekundärfarben (Orange und Grün) mit einigen Flecken der Primärfarben Rot und Gelb, während die Primärfarbe Blau überhaupt nicht aufscheint. Demzufolge hat Abbildung 44 einen hohen Farbe-an-sich-Kontrast, Abbildung 57 einen niedrigen.

### **9.2.1.2 Hell-Dunkel-Kontrast**

Durch ihren Zusammenhang mit der Helligkeit von (bewegten)Bildern, schaffen die „Farben“ Schwarz und Weiß überaus starke Kontraste. Bei der Interpretation dieses Farbkontrasts kommt es aber nicht nur rein auf die Prüfung des Vorhandenseins von hell und dunkel an, sondern vielmehr die genaue Betrachtung der Beziehungen zwischen hellen und dunklen Aspekten. In dieser Hinsicht lenkt Itten (2000, 2010) die Aufmerksamkeit auf einen Umstand, dessen wir uns normalerweise nicht bewusst sind, der aber schon 1810 von Goethe (2003) entdeckt wurde: Jede gesättigte Farbe hat ein unterschiedliches Helligkeitsniveau (z. B. ist Gelb die hellste, Violett die dunkelste Farbe). Beim Vergleich der zwei Videotranskripte scheint das von Abbildung 44 wie bei Tag aufgenommen, das von Abbildung 57 bei Nacht. Auf globalem Montagelevel ist Abbildung 44 hell, Abbildung 57 dunkel. Trotzdem ist der Kontrast zwischen hell und

dunkel in Abbildung 57 stärker, denn in der dunklen Umgebung gibt es auch helle Lichtflecken.

### **9.2.1.3 Kalt-Warm-Kontrast**

Der Kalt-Warm-Kontrast stellt einen Zusammenhang zwischen Farbe und damit verbundenen Temperaturen oder Temperaturveränderungen her. Nach Itten (2000, 2010) kann Orange-Rot als die „wärmste“, Grün-Blau als die „kälteste“ Farbe definiert werden. Beim Vergleich der zwei Musikvideotranskripte kann konstatiert werden, dass der Kalt-Warm-Kontrast in Abbildung 44 kaum in Erscheinung tritt, während er in Abbildung 57 durch die Komplementärfarben Orange-Rot und Grün zu Tage tritt.

### **9.2.1.4 Komplementär-Kontrast**

Auf Ittens zwölfteiligem Farbkreis, welcher eine analytische Zusammenstellung von konzentrisch angeordneten Primär-, Sekundär- und Tertiärfarben ist, stehen sich Komplementärfarben direkt vis-à-vis gegenüber. Die Vermischung dieser Farben auf Papier ergibt eine neutrale grau-schwarze Farbe; ihre Synchronüberlagerung auf dem Bildschirm produziert weißes Licht (Itten, 2010, S.49). Die Kombination von Aufnahmen in Komplementärfarben durch Montage hebt die Schnitte zwischen ihnen hervor. Beide Videotranskripte sind jeweils durch einen einzigen Komplementärkontrast charakterisiert. In Abbildung 44 ist es der Kontrast zwischen Gelb und Violett; in Abbildung 57 der Kontrast zwischen Rot und Grün.

Mittlerweile dürfte deutlich geworden sein, dass mit Ittens Konzept der Farbkontraste Beziehungsprofile von Farbkompositionen gefasst werden können, die für die Rekonstruktion des Dokumentsinns von Filmen und Videos nützlich sind. Der Komplementärkontrast von Gelb und Violett (Abbildung 44) neigt immer dazu, auch einen Hell-Dunkel-Kontrast hervorzurufen, weil Gelb dreimal heller ist als Violett. Im Beispiel von Abbildung 44 kann dieser Effekt nur durch eine Desaturierung der Farbe Violett im Montageprozess verringert werden. Die Komplementärfarben Rot und Grün (Abbildung 57) befinden sich im Vergleich dazu hinsichtlich ihrer Helligkeitswerte auf derselben Stufe und sind damit gleichwertig. Rot und Grün können, unabhängig vom Grad ihrer Sättigung, potentiell einen Kalt-Warm-Kontrast initiieren. Dies ist in Abbildung 57 beobachtbar. Da alle Farbkontraste miteinander zusammenhängen, ist es nicht möglich, dass ein Bild bzw. ein Film oder Video in allen Kontrastdimensionen hohe Werte erreicht. Diese Tatsache spricht aber gerade für die differenzielle Analyse von

Farbkontrasten und ihren Einsatz als differenziertes Instrument der Bild- und Videointerpretation.

#### **9.2.1.5 Simultan-Kontrast**

Der Simultan-Kontrast kommt beim Auftreten von neutralen Graubereichen neben gesättigten Farben zum Tragen. Unter Bezugnahme auf physiologische Reaktionen im menschlichen Auge erklärt Itten (2000, 2010), dass eigentlich neutrale Graubereiche in bei der Anwesenheit gesättigter Farben in der jeweiligen Komplementärfarbe erscheinen. Dieser Effekt wird weder in Abbildung 44 noch in Abbildung 57 deutlich, kann aber in Aufnahmen gefunden werden, die zwar nicht neutral grau, aber zumindest ungesättigt sind. In Abbildung 44 zum Beispiel verursacht der Simultan-Kontrast die Verstärkung der Komplementärfarbe des ungesättigten Violetts (z. B. Einzelbilder 11–14; 18–23).

#### **9.2.1.6 Qualitäts-Kontrast**

Die Sättigung legt den Grad der Farbenreinheit und -intensität fest. Ittens (2000, 2010) Definition zufolge sind die Farben, die durch optische Lichtbrechung von weißem Licht erzeugt werden, am gesättigsten. Die Voraussetzung für das Auftreten dieses Kontrast ist, dass gesättigte Farben zusammen mit ungesättigten (dumpfen) Farben dargestellt werden. Damit der Effekt wirksam wird, müssen die gesättigten und ungesättigten (dumpfen) Farben jedoch dieselbe Farbschattierung aufweisen (z. B. muss gesättigtes Rot räumlich oder zeitlich neben ungesättigtem Rot auftreten). Wenn unterschiedliche Farbschattierungen kombiniert werden, tritt der Qualitäts-Kontrast gegenüber anderen Kontrasten in den Hintergrund. Da dieser Effekt eher dezent ist, kann er in Abbildung 44, die großteils von starken Farbkontrasten dominiert wird, kaum ausgemacht werden. Abbildung 57 hingegen demonstriert die unterschwellige Macht des Qualitäts-Effektes geradezu beispielhaft: In Abwesenheit von starken Farbkontrasten und durch die „Ein-Tönigkeit“ der bearbeiteten Aufnahmen erfolgt beispielsweise eine Verstärkung der rosa Farbe des Kleides der Frau und ihrer Lippen<sup>105</sup> vor dem ungesättigten Hintergrund.

#### **9.2.1.7 Quantitäts-Kontrast**

Dieser letzte Farbkontrast bezieht sich auf die unterschiedliche Leuchtkraft einzelner Farben, die 1810 in Goethes „Theorie der Farben“ (2010) beschrieben wurde: Gelb = 9,

---

<sup>105</sup> Wenn die Frau später im Video in einer Szene mit starker Grünfärbung (Sekunden 17–19 in Bild 2) auftritt, erscheint das pinke Kleid strahlend weiß, während die rosa Lippen auf Rot umgeschminkt werden müssen, damit sie in der komplementären Farbumgebung dunkel aussehen.

Orange = 8, Rot = 6, Violett = 3, Grün = 6. Diese Zahlen oder Farbwerte sind mit dem Raumbedarf von Farben verbunden. Gelb (Farbwert 9) braucht zum Beispiel dreimal weniger Platz als Violett (Farbwert 3); die Komplementärfarben Rot und Grün brauchen gleich viel Platz (beide mit einem Farbwert von 6). Was die von den Videotranskripten eingefangene Montage betrifft, erklärt dies das häufigere Aufkommen von Cuts in dem Musikvideo für den türkischen Markt (Abbildung 44). Im Vergleich zu den hellvioletten Aufnahmen (Sekunden 18–23) „erweitern“ die großen gelben Aufnahmen (z. B. Sekunden 17–19) den Bildschirmraum. Die roten und grünen Aufnahmen der Musikvideos für den internationalen Markt (Abbildung 57) hingegen haben den gleichen räumlichen Bedarf und sind daher gleichwertig.

### 9.2.2 Vorikonografische und ikonografische Interpretation

Mit der Orientierung an der Methodologie der dokumentarischen Methode ist es bei der Beschreibung des visuellen Inhalts von (sich bewegenden) Bildern essenziell, die vorikonografische von der ikonografischen Ebene als Teile der *formulierenden Interpretation* zu unterscheiden. Die ikonografische Ebene kommt unseren alltäglichen Kommunikationsmitteln am nächsten. Auf dieser Interpretationsebene können wir Beschreibungen von Szenarien und Menschen dadurch erschaffen, dass wir sie in unseren eigenen Bedeutungssystemen kontextualisieren. In Abbildung 57 können wir beispielsweise eine junge, förmlich angezogene Frau erkennen, die in einem feinen Restaurant wartet und die Aufmerksamkeit eines Mannes auf sich zieht, der sie von einem anderen Tisch aus beobachtet.

Wenn wir nur die vorikonografische Information, die in dieser Abbildung enthalten ist, berücksichtigen, „sehen“ wir nur eine Frau, die an einem von vielen Tischen sitzt, und einen Mann, der in die Kamera blickt. Die Differenzierung zwischen vorikonografischer und ikonografischer Information ist in zweifacher Hinsicht nützlich: Zum einen trägt sie dazu bei, dass das, was die Forschenden (persönlich) in den Interpretationsprozess einbringen, systematisch identifiziert und kontrolliert werden kann. Zum anderen hilft sie aber auch dabei, gerade diese Anteile bewusst *für* die Interpretation nutzbar zu machen. Insofern erlaubt uns die Unterscheidung zwischen vorikonografischer und ikonografischer Information, das Phänomen der „media literacy“ (Mikunda, 2002) auch empirisch zu untersuchen. Im vorliegenden Fall ist von Interesse, wie es durch Montage gelingt, den Zuschauer/inne/n einen kohärenten Sinn zu vermitteln, obwohl im Video nur kleine Bruchstücke von Ereignissen gezeigt werden.

Da ich die Vorgehensweise bei der formulierenden Interpretation (auf vorikonografischer und ikonografischer Ebene) in dieser Arbeit anhand von Fallmaterial bereits ausführlich erläutert habe, möchte ich zum Schluss nun direkt zur reflektierenden Interpretation übergehen.

### **9.2.3 Szenische Choreografie und Montage**

Laut Bohnsack (2009, S. 39) und Imdahl (1996a, S. 19) fokussiert die Logik des Konzepts der szenischen Choreografie darauf, zu erwägen, wie sich die Menschen (oder Objekte) vor der Kamera (die dargestellten Bildproduzent/inn/en) zueinander und zu den Zuschauer/inne/n verhalten. Während die Rekonstruktion von szenischer Choreografie typischerweise im Rahmen der Interpretation einzelner Standbilder (Fotogramme) oder Aufnahmen angewandt wird, kann das Konzept auch auf die Montageebene ausgeweitet werden. Im letzteren Fall werden Menschen nicht länger als sich in einem physischen Raum bewegend, sondern als sich in einem „editing space“ (Bordwell, 1985, S. 117) oder einem „montierten Raum“ bewegend verstanden. Bohnsacks (2009, S. 164) Darstellung von Hicketier (2012) zufolge ist die Kreierung eines „filmischen Raumes“ jenseits des physischen Raumes tatsächlich eines der Schlüsselmerkmale von Film und Video, welches diese Medien von anderen visuellen Medien, wie z. B. der Fotografie, absetzt. Ich werde mich nur auf diejenigen essenziellen Aspekte der szenischen Choreografie konzentrieren, die für die *ikonische Interpretation* (Bohnsack, 2009; Imdahl, 1996a) der zwei Musikvideos benötigt werden.

#### **9.2.3.1 „Araba“ für den türkischen Markt**

Wie in Abbildung 44 zu sehen ist, beginnt das „Araba“-Video für den türkischen Markt mit einem Countdown (Sekunden 0–1), gefolgt von einem sitzenden Mann (dem Sänger), der den Zuschauenden lachend zugewandt ist. In Fotogramm 5 fährt ein gelber Ferrari schnell vorbei, der in den Fotogrammen 7 bis 10 genauer anvisiert wird. Als Nächstes (11–14) sehen wir einen Mann, der von den Zuschauenden weggeht und mit seiner Hand an einem Skullboot, einem Ruderboot für *eine* Person, entlangstreicht. Von Fotogramm 15 bis Fotogramm 17 wird der gelbe Ferrari wieder in der Frontalansicht dargestellt, bevor der Ruderer sich reckt und streckt (18–20). Als Nächstes sehen wir einen Fischer aufwachen und seine Baseballmütze etwas von seinem Gesicht hochheben; dann wieder den Ferrari (24–25), einen auf der Straße liegenden Hund (26), den Sänger (27–28), den Bug eines großen Schiffes, das eine große Welle vor sich über den

Bildschirm schiebt (29–30); den Ruderer, der auf das Meer schaut und mit dem Rücken zu den Zuschauer/inne/n steht (31); den Ferrari (32), den Ruderer, wie er seine Mütze abnimmt (33); den Ferrari (34), den Ruderer, wie er seinen Pullover auszieht (35–36); zwei Männerhände am Steuerrad des Ferraris (37); den Ruderer, wie er seine Beine auf dem Skullboot ausstreckt (38); wieder zwei Hände am Steuerrad des Ferraris (39).

Die ersten vierzig Sekunden des Musikvideos „Araba“ für den türkischen Markt (Abbildung 44) können als Vorbereitung der beiden Hauptcharaktere für eine Wettfahrt zusammengefasst werden. Die zwei Protagonisten sind der Ruderer und der Fahrer des Ferraris. In Fotogramm 42 geht eine bis hierher nicht identifizierbare Frau in einem weißen Kleid durch das Bild, welche in Fotogramm 48 die Wettfahrt dadurch startet, dass sie die Hände über ihrem Kopf zusammenklatscht. Ab diesem Moment rudert der Ruderer energisch den Bosphorus entlang, während der Ferrarifahrer mehrere Hindernisse des Verkehrs in Istanbul umgehen muss. Ihr gemeinsames Ziel ist eine Insel in der Mitte des Bosphorus, wo die junge Frau verzweifelt auf das Erscheinen einer der beiden Männer wartet. Zwischen den Szenen sehen wir ausführliche Aufnahmen des tanzenden und singenden Sängers. Während des Großteils des Videos kann der Fahrer des gelben Ferraris nicht gesehen werden. Die Geschichte endet damit, dass der Ruderer die Insel erreicht und der Frau in die Arme fällt (270–272). Der Ferrari kommt am Ufer zum Stehen (273–274); der Ruderer und die Frau drehen sich um, als sich die Tür des Wagens öffnet (275–277). Die Kamera schwenkt herum und zeigt die braune Hose des Fahrers, sein weißes Hemd und am Ende sein Gesicht (278–281). Während der letzten Sekunden des Videos schaut der Ferrarifahrer in die Kamera, bevor er sich dann wendet (282–289). Am Ende wird klar, dass Fahrer, Ruderer und Sänger ein und dieselbe Person sind.

### **9.2.3.2 „Araba“ für den internationalen Markt**

In Abbildung 57 wird die szenische Choreografie des Musikvideos „Araba“ für den internationalen Markt folgendermaßen präsentiert: Zu Beginn sitzt eine junge Frau mit dunklem Haar und einem rosafarbenen Kleid alleine an einem Restauranttisch (Fotogramme 0–1). Dann sehen wir den Kopf eines Mannes (des Sängers) und den einer Frau (vielleicht seine Freundin oder Ehefrau) im Profil. Er singt, sie sieht zu ihm auf (2–7). Anschließend können wir die Frau im rosafarbenen Kleid von weiter weg sehen; die leeren Tische um sie herum werden sichtbar (8–16). Zur gleichen Zeit werden drei junge Mädchen von hinten gezeigt, die rhythmisch in einem Nachtclub tanzen. Im Restaurant

sitzt die Frau noch immer an ihrem Tisch (20). Ein Kellner kommt zu ihr und serviert ihr Wasser (22–33). In der Zwischenzeit sitzt der Sänger im selben Raum an einem anderen Tisch mit drei anderen Menschen: einer Frau (derselben wie in den Fotogrammen 2–7) und zwei männlichen Freunden (z. B. Fotogramme 21 und 24). Als der Kellner das Wasser serviert, wird der Sänger auf die einsame Frau im pastellfarbenen Kleid aufmerksam (29). Nachdem der Kellner den Tisch verlassen hat, können wir das Gesicht des Sängers und das Dekolleté (37, 42–43) sowie das Gesicht der Frau (38, 44) sehen. In Fotogramm 45 betritt der Freund der Frau die Szene und geht zu ihrem Tisch. Während sich das Gesicht des Sängers der Kamera zuwendet, setzt sich der Freund, gerade mit jemand anderem telefonierend, hin. Die Frau im pastellfarbenen Kleid ist unglücklich über die Unhöflichkeit ihres Freundes, akzeptiert sie aber und bleibt weiter sitzen (50–63), während der Sänger den Kellner bezahlt (59). Ein Szenenwechsel zeigt wieder die drei jungen, hüftenschwingenden Mädchen (64). Als Nächstes betritt der Sänger mit seiner Freundin und seinen Freunden den Nachtclub und singt (65–77). Zurück im Restaurant sehen wir das Paar gerade anstoßen (78), als eine verärgerte junge Frau an den Tisch tritt (79–81). Leute drehen sich um (82); der Freund schaut verwundert drein (83). Als die verärgerte Frau ihren Ring in sein Champagnerglas wirft, entpuppt sie sich damit als seine Verlobte (85–86). Zur gleichen Zeit nähert sich der Sänger im Nachtclub der Bar und geht an verschiedenen tanzenden Mädchen vorbei (89–92). Die Verlobte verlässt das Restaurant. Der Freund versucht die Situation zu erklären, aber die Frau im rosa Kleid steht ebenfalls auf und geht (93–96). Als Nächstes erscheint die Frau im selben Nachtclub (97–99). Sie sieht ihn singend an der Bar stehen (100–102) und geht Richtung Toiletten. Während er weiter singt, tanzen verschiedene junge Mädchen und blicken in die Kamera, die Frau wischt ihre Tränen weg (118–119), rückt ihr Dekolleté zurecht und kommt zurück in den Tanzsaal. Dort versucht sie sich dem Sänger zu nähern, doch der wird durch eine Gruppe tanzender Mädchen von ihr getrennt (130–153). Sie sieht ihn dann wieder an der Bar (154–155), aber er geht weg (156). Sie bleibt an der Bar stehen und schaut irritiert (157–158). Er schaut noch einmal zu ihr zurück (159–160), bevor er sich seiner Freundin zuwendet (161–163). Wiederum werden tanzende Mädchen gezeigt (164–171), die Frau findet den Sänger wieder und schaut ihn gebannt an (172–188). Während weiter getanzt wird und der Sänger in der Ecke – und damit in der Nähe seiner Freundin – bleibt (189–201), beginnt die Frau irgendwann auch zu tanzen (202–205). Am Ende tanzen der Sänger und die Frau beide scheinbar glücklich, aber getrennt voneinander und die Party geht weiter. In den letzten Sekunden

tauschen der Sänger (228) und die Frau (229) einen letzten Blick, beide genau in die Kamera schauend. Die Augen der Frau schließen sich, dann wird ausgeblendet.

### **9.3 Ikonologische bzw. ikonische Interpretation**

Das „Ikonische“ kann mit Imdahl (1996a) als Bedeutungssensenz, als generative Struktur des rekonstruierten Dokuments verstanden werden. In diesem letzten Interpretationsschritt werden die Ergebnisse der vorherigen Stufen miteinander in Beziehung gesetzt. Das Hauptziel der ikonologischen und ikonischen Interpretation ist es, zu einer abschließenden Schlussfolgerung zu gelangen, die von allen Interpretationsebenen unterstützt wird. Mit der komparativen Analyse der zwei Musikvideos zu „Araba“ wurde ein komplexes und herausforderndes Informationsgebilde bezüglich biografischer Brüche und ihrer „Reparaturen“ herausgearbeitet. Der dokumentarischen Methode folgend werden nun zuerst Gemeinsamkeiten diskutiert, bevor auf die Unterschiede eingegangen wird. Diese Vorgehensweise offenbart die empirischen Verbindungen zwischen den Fällen, die sonst den Alltagsinterpretationen überlassen werden.

### 9.3.1 Farbkontraste

Es gibt nur einen Farbkontrast, den die beiden Videoverversionen gemeinsam haben: den Komplementär-Kontrast. In Abbildung 44 wird er durch die Farben Gelb und Violett ausgedrückt, in Abbildung 57 durch den Kontrast von Rot und Grün. Die dazwischenliegenden Cuts werden durch komplementärfarbene Aufnahmen hervorgehoben. Während in Abbildung 44 fast jeder Schnitt mit dieser Methode hervorgehoben wird, ist das in Abbildung 57 nur bei zwei ausgewählten Cuts der Fall (Sekunden 17–19). Die geringere Anzahl von Komplementär-Kontrasten im Musikvideo für den internationalen Markt erzeugt generell eine viel homogenere Atmosphäre. Der Homogenitätsgedanke dokumentiert sich auch auf dem Niveau der Farbleuchtkraft (Itten, 2000, 2010): Im Wesentlichen ist Gelb eine viel hellere Farbe als ihre Komplementärfarbe Violett. Rot dagegen ist genauso hell wie Grün. Im Grunde macht das die Wechsel zwischen komplementärfarbenen Einstellungen der „Araba“-Version für den internationalen Markt unauffälliger als in der Version für den türkischen Markt.

Der Farbe-an-sich-Kontrast ist einer der stärksten Farbkontraste in Abbildung 44, da fast alle Primärfarben und einige Sekundärfarben anzutreffen sind. Zusammenfassend lässt sich daher sagen, dass die Version für den türkischen Markt von sehr lebendigen und hellen Farben beherrscht wird. Die Farben sind sehr „laut“, was noch durch die Dominanz des Quantitätseffekts unterstrichen wird. Durch die Montage kann kaum ein stärkerer „Farbschock“ (Mikunda, 2002) ausgelöst werden als durch die Vollbildansicht des knallgelben Ferraris. Dies erzeugt ein signifikantes strukturelles Ungleichgewicht im Video, das durch die komplementärfarbenen Gegenaufnahmen und den ständigen Wechsel zwischen den Aufnahmen ausbalanciert wird. Im Sinne Imdahls (1996a) kann dieser Mechanismus als Übergegensätzlichkeit bezeichnet werden. Auf der einen Seite wird der Effekt der gelben Farbe durch das Aufscheinen ihrer Komplementärfarbe reduziert; auf der anderen Seite verstärkt in der darauffolgenden Aufnahme das Vorhandensein der Komplementärfarbe die Dominanz des Gelben sogar noch. Auf ikonischer Ebene repräsentiert dies den Kampf gegen einen fast unbesiegbaren Feind.

Auf der Ebene von Farbkontrasten illustriert Abbildung 57 einen anderen Zugang: Keine Farbe wird hier aktiv und unübersehbar in den Vordergrund gestellt. Dies lässt eine viel dezentere, aber sogar noch effektivere Farbeinstellung zu. Eine allgemeine Lichtreduzierung ist der effektivste Weg, dieses Ergebnis zu erreichen. Indem das

Setting vom hellen Tageslicht nach drinnen verlegt wird, kann bei der Bearbeitung die Verstärkung und Reduzierung von Licht effektiv eingesetzt werden. Anstatt nur den Farbe-an-sich- und den Qualitäts-Kontrast in einer hellen, sonnenbeleuchteten Atmosphäre differenzierend einsetzen zu können, war es den Herausgebern der „Araba“-Version für den internationalen Markt durch die Lichtreduzierung möglich, die Scheinwerfer in einer vielseitigeren und weniger auffälligen Art und Weise auf spezifische Farben zu richten und bestimmte Aspekte des Szenarios in den Mittelpunkt zu stellen.

### **9.3.2 Szenische Choreografie im Verhältnis zu Farbkontrasten**

Die Protagonisten der „Araba“-Version für den türkischen Markt (Abbildung 44) sind mit einer spezifischen Farbanordnung assoziiert: Der Ruderer ist in eine hellviolette oder hellblaue Umgebung eingebettet, während der Ferrarifahrer von der gelben Farbe seines Autos umgeben und repräsentiert wird. Zwischen diesen beiden Umgebungen hin und her zu wechseln, ist – nicht nur auf der Farbebene, sondern auch auf dem Level der sozialen Klasse und des materiellen Wohlstandes – möglich. Im Vergleich dazu können sich die Protagonist/inn/en der „Araba“-Version für den internationalen Markt (Abbildung 57) an ihre (Farb-)Umgebungen anpassen und auch zwischen ihnen wechseln. Dadurch werden rein formal – im Gegensatz zur türkischen Version – keine expliziten Klassenunterschiede aufgeworfen. Die Grenzen zwischen den „sozialen Welten“ des Sängers und der Frau im pastellfarbenen Kleid sind durchlässig und machen eine Begegnung auf Augenhöhe möglich. Auf farblicher Ebene findet dieser Aspekt durch die ausgewogene Mischung der Komplementärfarben Grün und Rot seinen Ausdruck.

### **9.3.3 Biografische Brüche von Männern und Frauen**

Im Musikvideo „Araba“ für den türkischen Markt beruht der biografische Bruch für die Frau in ihrem Hin- und Hergerissensein zwischen zwei miteinander konkurrierenden Männern oder zwei konkurrierenden Seiten *eines* Mannes.<sup>106</sup> Von einer Insel im Bosphorus aus initiiert die Frau das Schlussrennen zwischen den Männern. Ein Konkurrent ist durch den gelben Ferrari, einem Merkmal für Wohlstand und Luxus, charakterisiert. Der andere ist durch das Rudern über den Bosphorus gekennzeichnet, welches harte Arbeit und physische Kraft symbolisiert. Der Bruch geschieht vor allem

---

<sup>106</sup> Wenn der Sänger noch mitgerechnet wird, könnte sogar behauptet werden, dass der Hauptdarsteller *drei* miteinander konkurrierende Seiten hat.

auf der Ebene der formalen Komposition (z. B. durch starke Farbkontraste). Dennoch gibt es auch einen explizit artikulierten Bruch zwischen den zwei Rivalen, nämlich den bezüglich Klasse und Einkommen. Brüche in der „Araba“-Version für den türkischen Markt sind daher universale und akzeptierte Angelegenheiten hinsichtlich männlicher Charakterkonstitutionen und sozialer Klasse. Nur für die Frau, die erwartet, dass es zu einer Entscheidung kommt, ist der Bruch biografisch.

In der „Araba“-Version für den internationalen Markt gibt es keinen expliziten Bruch in der Gesellschaft. Obwohl argumentiert werden könnte, dass es einen Bruch in der Persönlichkeit des untreuen Freundes gibt, kann das nicht für alle Männer im Video generalisiert werden. Der Sänger beispielsweise wird als Mann mit einer integren Persönlichkeit porträtiert. Im Gegensatz zur „Araba“-Version für den türkischen Markt ist er vom biografischen Bruch der Frau ebenfalls betroffen.

#### **9.3.4 Reparatur als Naturgesetz**

In der „Araba“-Version für den türkischen Markt rettet der Mann, der sich mit ganzer Kraft für die Frau einsetzt, die Beziehung. Indem er harte Arbeit und Schweiß wühlt, um seine Liebste auf der Insel zu erreichen, entscheidet er sich gegen Luxus und Oberflächlichkeit, gegen schicke Autos und feine Kleidung. Seine Entscheidung stellt eine Konsolidierung der romantischen Liebesbeziehung dar, die aufgrund der demonstrierten Ambiguität seines Charakters auf dem Spiel stand. Am Ende fällt die Frau ihrem Liebsten in die Arme. Als sie dem inneren Rivalen des Mannes in den Abschlusszenen gegenüberstehen, kommt ein Funken Zweifel im Paar auf, aber indem sich der Ferrarifahrer von ihnen wegdreht, wird die Situation besiegelt und die Beziehung gerettet.

Was die Darstellung der Frau im Musikvideo betrifft, ist erwähnenswert, dass sie den Charakterwettkampf initiiert, aber keine aktive Rolle bei der Schicksalsentscheidung spielt. Die Beziehung des Paares steht im Grunde nie wirklich auf dem Spiel. Sie wird weitergehen, unabhängig davon, welcher der konkurrierenden Persönlichkeitsmerkmale des Mannes gewinnt. Die Frau würde sich jedem möglichen Ergebnis fügen, solange nur die Ambiguität des Mannes ein Ende findet. Ihre Rolle der machtlosen Zuschauerin wird durch das Auftreten des Sängers Mustafa Sandal ergänzt, der einerseits außerhalb des Ereignisablaufs agiert (indem er nur über sie singt), andererseits aber auch noch die Rolle der zwei männlichen Hauptfiguren spielt und daher das gesamte Video beherrscht. Abschließend lässt sich sagen, dass dieses

indirekte Mittel der Situationskontrolle von einer Außenposition aus als bemerkenswerte Analogie zu der im Video skizzierten Rolle der Frau angesehen werden kann. Die Hauptaussage zu Reparaturmöglichkeiten ist hier, dass Männer aktive Entscheidungen treffen sollten, um ihren Partnerinnen bei der Lösung ihrer Probleme zur Seite zu stehen. Den Frauen ihrerseits kommt der Part zu, den männlichen Entscheidungsprozess indirekt zu initiieren, indem der emotionale Druck (z. B. durch die Zurschaustellung von Unglücklichsein und Leiden) erhöht wird, obwohl es ihnen aber nicht möglich ist, eine *aktive* Rolle im Entscheidungsprozess einzunehmen. Die Beziehung in die Brüche gehen zu lassen und allein zu sein, ist weder für die Männer noch die Frauen eine Option. Dies ist ein signifikanter Unterschied zu Reparaturvorschlägen in der internationalen Version des Musikvideos.

### **9.3.5 Reparatur als Angelegenheit individueller Freiheit**

In der „Araba“-Version für den internationalen Markt ist es die Frau, die die Beziehung beendet, nachdem sie von der Untreue ihres Geliebten erfährt. Sie entscheidet sich dafür, allein zu sein, anstatt mit jemandem zusammen zu sein, der sie betrogen hat. Dieser Bruch verlangt von ihr emotionale Stärke und Selbstvertrauen, obwohl sie die Entscheidung nicht rational und nach langer Überlegung, sondern emotional und impulsiv trifft.

Die „angeknackste“ Beziehung ist zu Ende, der Mann verloren. Jetzt muss die Frau mit der niederschmetternden Erfahrung des Betrugs durch ihren Geliebten ins Reine kommen. Zunächst flieht sie vor der Situation, betritt aber nur wenige Momente später einen Nachtclub und wird vom Blick des türkischen Sängers gefangen genommen, der – im Gegensatz zur „türkischen“ Version von „Araba“ – hier einen aktiven Part in der Geschichte einnimmt. Die Frau läuft zu den Toiletten, nimmt ihren Schal ab und wischt ihre Tränen weg. Dann rückt sie ihr Dekolleté zurecht, öffnet ihre Haare und betritt die Nachtclubszene erneut. Ihr helles Kleid und die helle Haut heben sich von der dunklen Umgebung der Bar und der Kleidung der anderen ab. Es hat den Anschein, als wolle sie sich an ihrem Exfreund rächen und den türkischen Sänger für ein sexuelles Abenteuer gewinnen. Der ist zwar von ihr angezogen, bleibt aber seiner Partnerin treu; sodass die Frau ihn nur aus der Ferne beobachten kann und akzeptieren muss, dass er nicht zu haben ist. Dennoch scheint sie im Musikvideo durch die Vorstellung beruhigt zu werden, dass es möglich ist, einen anderen Menschen zu lieben und von dieser Person wiedergeliebt zu werden. Indem sie sich mit einer platonischen Beziehung zu dem

türkischen Sänger zufriedengibt, kann die Frau ihr ursprüngliches Bedürfnis, sich an ihrem Exfreund zu rächen, überwinden. Am Ende fühlt sie sich wohl damit, allein zu sein, und kann die Atmosphäre des Nachtclubs einfach genießen. Die Anziehung zwischen dem türkischen Sänger und der Frau bleibt bestehen, als sich ihre Blicke ein letztes Mal treffen. Am Ende schließt sie aber dennoch die Augen vor diesem Kapitel in ihrem Leben.

Die Hauptaussage zu Reparaturoptionen, die in der „Araba“-Version für den internationalen Markt entwickelt wird, ist, dass Frauen, deren Partner untreu sind, die Entscheidung treffen sollten, ihre Beziehung zu beenden. Sie sollten sich auf sich selbst besinnen und sich auf die Jagd nach ihrem nächsten Partner begeben, indem sie ihre sexuellen Reize zur Schau stellen. In stereotypen Alltagsvorstellungen halten türkische Männer am Glauben an kurze sexuelle Abenteuer fest und versuchen ständig, sich in zerrüttete Beziehungen von Frauen zu drängen. Der türkische Sänger verhält sich in dieser Hinsicht nicht stereotypisch, da er seiner Freundin treu bleibt. Dies veranschaulicht, dass die „Araba“-Version für den internationalen Markt aktiv das Thema der nationalen Herkunft oder des kulturellen Hintergrundes adressiert, was in der Version für den türkischen Markt nicht der Fall ist. Dies kann möglicherweise auch erklären, warum sich die sexuelle Intrige am Ende auflöst. Im Wesentlichen sagt das Video aus, dass es als Antwort auf einen Vertrauensbruch akzeptabel ist, die eigenen sexuellen Hemmungen zu überwinden. Letzten Endes kann Untreue aber dennoch nur durch Treue repariert werden.

#### **9.4 Zusammenfassung**

Auf den ersten Blick erscheint die Frau in der „Araba“-Version für den internationalen Markt sehr viel selbstbewusster als die Frau im Video für den türkischen Markt. Dies ist größtenteils dem Umstand zu verdanken, dass es ihr möglich ist, einen biografischen Bruch in ihrem eigenen Leben aktiv herbeizuführen, indem sie eine unehrliche Beziehung beendet. Als Reaktion auf die Untreue ihres (Ex-)Freundes wird der Bruch zur Legitimierung von rachsüchtiger und sexualisierender „Reparatur“. Es scheint, als wäre die Frau moralisch dazu berechtigt, eine sexuelle Beziehung mit dem nächstbesten Mann einzugehen. Im Video für den internationalen Markt regt der exotische Fremde, der türkische Sänger, den sexuellen Appetit der Frau an. Nach seiner unerwarteten Weigerung, sich mit ihr einzulassen, kann sie aber die Situation akzeptieren und ihre Begierde im Tanz sublimieren.

Im Vergleich dazu ist der Beziehungsabbruch in der türkischen Version des Videos weder für die Frau noch für den Mann/die Männer eine Option; ebenso wenig werden Horizonte von Rache oder Sexualisierung aufgeworfen. Von signifikantem Stellenwert für die Frau im Video für den türkischen Markt ist jedoch die Rivalität zweier Männer, die um sie im Wettkampf buhlen. In dem Umstand, dass dieser Wettkampf ohne Tricks und Betrügereien abläuft, dokumentiert sich, dass aus moralischer Sicht Fairplay offensichtlich hohe Relevanz für die Bildproduzent/inn/en besitzt. Das Musikvideo erhält dadurch fast den Charakter einer guten Sportübertragung: Möge der bessere gewinnen! Gewinner und Verlierer werden beide das Resultat anerkennen. Da – wie sich am Ende des Musikvideos herausstellt – beide Rivalen sowie der Sänger ein und dieselbe Person sind, beruht der dargestellte Bruch vermutlich auf der gespaltenen Persönlichkeit dieser facettenreichen männlichen Hauptfigur. Der hart erkämpfte Sieg des einfachen Ruderers stellt damit auch eine Integration der Persönlichkeit dar. Dies ist der Lohn für seine unnachgiebige Willensstärke und Enthaltensamkeit, die er in der Abkehr von Luxus, Bequemlichkeit und Oberflächlichkeit aufzubringen vermag.

Demnach dreht sich der dargestellte Ablauf der romantischen Beziehung in der Version für den türkischen Markt grundsätzlich um die persönliche Entwicklung des Mannes, während es in der Version für den internationalen Markt um die persönliche Entwicklung der Frau geht. In diesem Zusammenhang verweist das Ergebnis der komparativen Analyse auf kultur- und geschlechtsspezifische Besonderheiten. In beiden Versionen des Musikvideos „Araba“ besteht ein prinzipieller Bruch in der Persönlichkeit des Mannes. In der Version für den türkischen Markt wird dieser offen nach außen getragen und in Form und Ästhetik der Parallelmontage komplementärer Einstellungen gestalterisch umgesetzt. Erst am Ende des Videos lösen sich die Farbkontraste und Rivalitäten auf, nachdem der männliche Protagonist zu sich selbst und *seiner* Frau gefunden hat.

In der Version für den internationalen Markt hingegen ist der männliche Persönlichkeitsbruch formal unauffällig und im Wesentlichen unsichtbar. Er wird nur *ex negativo* durch den Auftritt der weiblichen Nebenbuhlerin evident. Formale Konsequenzen für Bildgestaltung und Montage hat dabei nicht dieser Bruch an sich, sondern erst die aktive Entscheidung der betrogenen Hauptakteurin, die Szene zu verlassen. In diesem Sinne haben wir es in der internationalen Version des Videos mit einem impulsiven und selbstinduzierten Bruch der Frau zu tun. Diese hätte schließlich auch sitzen bleiben und die unerfreuliche Angelegenheit mit ihrem Freund bzw. Mann in

Ruhe besprechen können. In der internationalen Version wird auch die „Reparatur“ des selbstinduzierten Bruchs der Frau als impulsive und persönliche Angelegenheit dargestellt. Diese ist mit Gelüsten nach Rache und ungehemmter Sexualität als notwendige Übergangsstadien verbunden.

Während die im Video für den internationalen Markt dargebotenen Reparaturoptionen auf die persönlichen Kapazitäten der einzelnen Frau setzen, ist Reparatur in der türkischen Version ein automatisches Ergebnis des männlichen Konkurrenzkampfes. In der internationalen Version wird der Bruch der Beziehung als dramatischer und existenzieller Moment mit unsicherem Ausgang dargestellt. In der türkischen Version ist der Bruch der männlichen Persönlichkeit als zugrundegelegte Vorbedingung des Videos angelegt, dessen „Reparatur“ aber nur eine Frage der Zeit ist: Derjenige, der zuerst bei der Frau ankommt, entscheidet den Wettbewerb. Die Beziehung ist dabei niemals wirklich bedroht; sie wird *in beiden Fällen* gerettet.

Wie durch die komparative Videointerpretation der türkischen und internationalen Version des Musikvideos „Araba“ gezeigt werden konnte, sind biografische Brüche und ihre Reparaturformen nicht universell, sondern verweisen jeweils auf einen soziohistorischen Kontext vermittelter Praktiken und habitueller Orientierungen. Unter Verwendung von Bourdieus Konzeptualisierung der *strukturierenden Strukturen* (1984) können beliebte Musikvideos sowohl als Dokumente als auch als Ursachen kultureller Praktiken angesehen werden: Auf der einen Seite müssen Musikvideos an die täglichen Lebensereignisse ihrer Rezipient/inn/en anknüpfen, um innerhalb von deren Relevanzsystemen verstanden zu werden. Auf der anderen Seite müssen Musikvideos auch die *reine* Wiedergabe des Alltagslebens überschreiten, um aus diesem herauszuragen und Popularität zu erlangen. Dazu müssen Musikvideos vereinfachen und idealisieren. Im Mechanismus ihrer wechselseitigen Abhängigkeit moderieren Medien und Alltagspraktiken gemeinsam die Selbstkonzepte und Gewohnheiten der involvierten Rezipient/inn/en bzw. Medienkulturen. Latour (1996) beschreibt diese reziproke Beziehung zwischen Medien und Alltagspraktiken mit dem Konzept der *Artikulation*. Laut Sloterdijk (2004) bildet die Artikulation „den bewegten Horizont aller ‚Vorschläge‘, in denen mögliches und wirkliches Seiendes sich in propositioneller oder provokativer Weise dem menschlichen Bemerkten anbietet“ (2004, S. 219). Anders gesprochen können Artikulationen auch wie Falten im Stoff der menschlichen Existenz aufgefasst werden. Ihr Inhalt kann niemals als Ganzes expliziert werden, sondern immer nur lokal durch das Entfernen einer bestimmten Falte

(Sloterdijk, 2004, S. 220). Die vorliegende Interpretation der beiden Musikvideoverionen von „Araba“ kann als Beispiel einer kultur- und medienpsychologischen Explikation verstanden werden. In diesem Sinne können populäre Alltagsmedien wie Musikvideos auch als Katalysatoren von Artikulationsprozessen angesehen werden. Sie bieten sich ihren Nutzer/inne/n nicht nur als Inhalte, sondern auch als Mittel wechselseitiger Kommunikationsprozesse an.





**Abbildung 58: „Die Freilegung eines bedeutenden Mosaiks in der Türkei, Sommer 2012 “**

Quelle:

<http://newsroom.unl.edu/releases/2012/09/18/UNL+archaeological+team+unearths+giant+Roman+mosaic+in+southern+Turkey>

Ausschnitte aus der Meldung des Fundes auf der Homepage der University of Nebraska-Lincoln:

“A University of Nebraska-Lincoln archaeological team has uncovered a massive Roman mosaic in southern Turkey – a meticulously crafted, 1,600-square-foot work of decorative handiwork built during the region’s imperial zenith. It’s believed to be the largest mosaic of its type in the region and demonstrates the reach and cultural influence of the Roman Empire in the area in the third and fourth centuries A.D., said Michael Hoff, Hixson-Lied professor of art history at UNL and the director of the excavation. ...

‘We were surprised to have found a mosaic of such size and of such caliber in this region – it’s an area that had usually been off the radar screens of most ancient historians and archeologists, and suddenly this mosaic comes into view and causes us to change our focus about what we think (the region) was like.’ ...

Ben Kreimer, a senior journalism major, agreed: ‘(Working on) the mosaic was great because the more soil you removed, the more mosaic there was,’ he said. ‘Visually, it was also stunning, especially once it got cleaned off. It wasn’t very deep under the surface of the soil, either, so ... we had to be careful not to swing the handpick too hard so as not to damage the priceless mosaic that lay just inches beneath us.’” (Andersen & Smith, 2012)

## 10 Abschließende Diskussion

Es war der Anspruch dieser Arbeit, die Leistungsfähigkeit sowie die methodologischen Erweiterungspotentiale des empirischen Verfahrens der Film- und Videointerpretation nach der dokumentarischen Methode anhand konkreter Fallbeispiele vorzustellen. Dabei konnte ich mich im Itten'schen Sinne (siehe Einleitung) über weite Strecken hinweg auf den durch meine Lehrer und Lehrerinnen vorbereiteten sicheren „Fahrwegen“ wissen. Außerdem war es mir gestattet, für mein Vorankommen auch deren bewährte „Wägen“, d. h. ihre erprobten Instrumente und Denkweisen, zu verwenden. So haben mir die vielfältigen Vorarbeiten anderer in vielerlei Hinsicht erspart „das Rad neu zu erfinden“. Ich hoffe, dass ich im Gegenzug im Rahmen dieser Arbeit einige Ansätze entwickeln konnte, die wiederum für andere Forscher/innen im Bereich der Bild- und Videointerpretation hilfreich sind.

Von ihrem Aufbau her ist diese Arbeit stark durch empirische Fallbeispiele strukturiert, die auf anschauliche Weise praktische Erweiterungsansätze zur Methodologie der Video- und Filminterpretation nach der dokumentarischen Methode vermitteln sollten. Konkrete Ansatzpunkte zur Methodologieentwicklung wurden dabei – über mehrere Kapitel verteilt – auf verschiedenen Ebenen vorgebracht (methodisch, technisch, metatheoretisch, inhaltlich etc.). In der abschließenden Betrachtung dieser Zusammenfassung und Diskussion möchte ich nochmals explizit den gemeinsamen Gesamtbeitrag der einzelnen Teile verdichten und die daraus resultierenden methodologischen Konsequenzen aufzuzeigen.

Das Videotranskriptionssystem MoViQ und die Transkriptionssoftware MoviScript können als technische Instrumente der Video- und Filminterpretation angesehen werden. Zugleich wird dadurch jedoch die Einnahme einer neuen Einstellung der Betrachter/innen zum Video- bzw. Filmmaterial möglich, die ohne diese Hilfsmittel nicht zu erreichen wäre. Diese neue Perspektive schafft in erste Linie einen Überblick und erlaubt es den Forscher/inne/n, sich ihre audiovisuellen „Fälle“ in aller Ruhe anzusehen. Im Gegensatz zum laufenden Videobild, das sich in einer Sequenzialität ständiger Veränderung befindetet, erlaubt das Videotranskript den simultanen Zugriff auf die unablässig parallel ablaufenden Bild- und Tonereignisse. Dadurch werden zuvor verborgene Spuren des Produktionsprozesses plötzlich sichtbar, die sich in den wechselseitigen Mustern der Bild- und Tonkomposition dokumentieren. Auf der Ebene der Einstellungen etwa lässt sich so im Falle von Fernsehshows eine Hierarchie von

Haupt-, Unter- und eingelagerten Sequenzen rekonstruieren, was wiederum auf einen bestimmten Montagestil verweist. Außerdem können Gesten und Gebärden als Leistungen der abgebildeten Bildproduzent/inn/en in ihrem Vollzug und Ablauf untersucht sowie mit anderen Gesten und Gebärden verglichen werden. Des Weiteren erlauben Videotranskripte eine Verständigung über die an einem Video bzw. Film beteiligten Farben und Farbkontraste, denen zumindest in professionellen Produktionen eine dezidierte Hauptfunktion zur unterschweligen Unterstützung des intentionalen Ausdruckssinns zukommt (Frost, 2009). Wie das Fallbeispiel der Beyaz-Show gezeigt hat, können Videotranskripte auch dazu eingesetzt werden, das Verhältnis zwischen Bild- und Tondimension systematisch zu verstehen und etwaige Diskrepanzen zwischen diesen Ebenen sichtbar zu machen. Im Kapitel „Istanbul Total“ werden Videotranskripte auch im Rahmen von Gattungsanalysen verwendet: Durch den Vergleich von Montagemustern, die sich wie ein mosaikartiger Fingerabdruck in Videotranskripten dokumentieren, lassen sich Typen – d.h. Genres oder Gattungen – von Videos identifizieren. Abschließend können Videotranskripte sowohl im Rahmen von Forschungswerkstätten als auch von Publikationen in besonderem Maße die intersubjektive Nachvollziehbarkeit der durchgeführten Arbeits- und Interpretationsschritte verbessern. In diesem Zusammenhang schaffen sie auch verbindliche Referenzpunkte zur wissenschaftlichen Kommunikation über Video- und Filmmaterial.

Bei aller Begeisterung für die Anschauungsform der Videotranskripte möchte ich den Leser/inne/n an dieser Stelle jedoch nicht vorenthalten, dass deren Erstellung und Einsatz aus erkenntnistheoretischer Sicht auch problematisiert werden kann. Letztlich stellen Videotranskripte nicht nur Forschungsinstrumente, sondern zugleich auch Artefakte von Forscher/innen/handeln dar. Wenngleich an dieser Stelle keine elaborierte Dekonstruktion dieses Handelns zu erwarten ist, so sei zumindest der Hinweis auf die weiterführende Lektüre Latours (2002) gegeben. Hinsichtlich ihrer medialen Möglichkeiten erinnern Videotranskripte in gewisser Weise durchaus an die Setzkästen (Latour, 2002, Abb. 2.13, S. 65 und Abb. 2.14, S. 68), mithilfe derer Latours Urwaldforscher/innen Bodenproben aus der ganzen Welt miteinander vergleichen. Die durch Standardisierung angestrebte Erhöhung der Vergleichbarkeit des Fallmaterials geht dabei mit dessen Dekontextualisierung einher. An Latour erinnert in weitere Folge auch der Einsatz von Farben bzw. Farbkontrasten, der in dieser Arbeit zu Interpretationszwecken entwickelt wurde. „Latours Setzkasten“ ermöglicht in

Kombination mit einer standardisierten Farbkarte nämlich auch die systematische Klassifizierung der Bodenproben (Latour, 2002, Abb. 2.16, S. 73). Die Theorie der Farbkontraste nach Itten (Itten, 2000, 2010) wurde in dieser Arbeit aus mehreren Gründen eingeführt. Wohl am zentralsten ist dabei die Überlegung, dass die Theorie der Farbkontraste einen wesentlichen Beitrag zum tieferen Verständnis des „sehenden Sehens“ *sensu* Imdahl (1996a) leisten kann, indem sie eine Ergänzung zum etablierten Feldlinienansatz darstellt. Die Rekonstruktion von Feldlinien war bislang die einzige Möglichkeit zur wissenschaftlichen Verständigung über die Bildebene der planimetrischen Komposition, die für die Interpretation der Ikonik, d. h. des bildlichen Dokumentsinns, von maßgeblicher Bedeutung ist. Wie anhand des Fallmaterials von „Araba“ für den türkischen Markt (1996) gezeigt werden konnte, kann die Interpretation der Farbkontraste in methodologischer Hinsicht auch zur wechselseitigen Validierung mit Feldlinien eingesetzt werden. Ihr eigentliches Potential entfaltet die Farbkontrasttheorie aber in der Anwendung auf die Rekonstruktion der Montage. Anhand der Musikvideos „Araba“ (1996) und „Araba“ (2004) habe ich den empirischen Versuch unternommen, Einstellungswechsel als Farbwechsel aufzufassen und als solche zu interpretieren. Formal gesehen bietet jeder Einstellungswechsel den abbildenden Bildproduzent/inn/en die Möglichkeit, diesen durch Farbkontraste zu betonen oder durch Farbähnlichkeiten zu kaschieren. Damit lassen sich Farbkontraste als Funktion von Montagemustern verstehen. Die Eleganz dieses Zugangs liegt in der Möglichkeit eines direkten Zugangs zur Rekonstruktion der Montage. Dieser kann so ohne Umwege über die Detailinterpretation von Einstellungen geleistet werden. Wie aus dem empirischen Gattungsvergleich dieser Arbeit abgeleitet werden kann, mag dieser Vorteil bei der Interpretation von Fernsehshows und vergleichbarem Videomaterial nicht allzu sehr ins Gewicht fallen. Der Dokumentsinn von Fernsehshows ist in der Regel anhand nur weniger Einzeleinstellungen bestimmbar. Im Gegensatz dazu stellt sich bei Musikvideos aber die Frage, inwieweit dem Fallmaterial ohne einen direkten Rekonstruktionsansatz hinsichtlich der Montage entsprochen werden kann. Zwar konnte am Fallmaterial von „Araba“ (1996) gezeigt werden, dass auf der Basis der bestehenden Methodologie der dokumentarischen Video- und Filminterpretation die Interpretation von Musikvideos prinzipiell möglich ist. Die spezifische Beschaffenheit des Fallmaterials sorgte jedoch dafür, dass bei der Auswahl der Sequenzen und Einstellungen von der konsequenten Einhaltung der vorgegebenen Arbeitsschritte abgewichen werden musste. Demgegenüber sind von der vorrangigen Interpretation

von Montagemustern im Falle von Musikvideos (und mitunter auch weiterer Gattungen) einige Vorteile zu erwarten. Ein wichtiger Gedanke hierbei ist, dass ein direkter Zugang zur formalen Rekonstruktion der Montage die forschungsökonomische Entwicklung von Interpretationsansätzen erleichtern kann. Auf der Basis dieser ersten empirischen Ansätze lässt sich in weiterer Folge und in methodisch kontrollierter Weise eine Auswahl repräsentativer bzw. fokussierter Einstellungen sowie von Fotogrammen treffen. Anschließend können die Fotogramme ihrerseits einer detaillierten Bildinterpretation unterzogen werden. Außerdem könnte durch einen direkten Zugang zur Rekonstruktion der Montage letztlich auch seitens der Psychologie sowie der Sozialwissenschaften dem filmwissenschaftlichen Standpunkt Rechnung getragen werden, wonach die Montage das eigentliche Alleinstellungsmerkmal von Videos und Filmen darstellt.

Als Vorzug des in dieser Arbeit vorgestellten Methodeninventars können abschließend seine Flexibilität und vielfältigen Einsatzmöglichkeiten verstanden werden. So lassen sich etwa Fotogramme, die im Zuge der Rekonstruktion der Montage ausgewählt und aus Videotranskripten gewonnen wurden, ihrerseits wieder als Suchkriterien für eine komparative Analyse nutzen. Während es früher aus technologischen Gründen lediglich möglich war, nach (vor-)ikonografischen Kontrasten zu bestehendem Bildmaterial zu suchen, erlauben moderne Suchmaschinen im Internet heute auch die bildgestützte Suche nach Vergleichsbildern, die als planimetrische Kontraste dienen können. Wie der diesbezügliche Beitrag im Kapitel „Istanbul Total“ gezeigt hat, sind diese Verfahren nach wie vor durch den Umfang der Bilddatenbanken sowie teilweise durch mangelnde Transparenz hinsichtlich der bildlichen Auswahlkriterien der Suchanbieter begrenzt. Aus kultur- und medientheoretischer Sicht verweisen solche technologischen Möglichkeiten jedoch deutlich auf zukünftige Entwicklungspotentiale der Bild- und Videointerpretation nach der dokumentarischen Methode. Wie die in dieser Arbeit anhand von Fernsehshows und Musikvideos vorgestellten methodologischen Erweiterungen gezeigt haben, bewährt sich das Verfahren nicht nur zur Rekonstruktion des *sozialen Sinns* von Video- und Filmmaterial. Aufgrund ihrer besonderen methodologischen Orientierung an der Eigensinnigkeit der (bewegten) Bilder eignet sich die dokumentarische Bild- und Videointerpretation auch zur Rekonstruktion des *medialen Sinns* in der Bilddimension. Will man der Behauptung eines bekannten Medientheoretikers Glauben schenken, so stellt das Medium die

Botschaft dar. Mithilfe des in dieser Arbeit vorgestellten Methodenrepertoires lässt sich diese Aussage nun auf ihren empirischen Gehalt hin prüfen.



## 11 Literaturverzeichnis

- Andersen, K., & Smith, S. (2012). UNL archaeological team unearths giant Roman mosaic in southern Turkey. Abgerufen 30. April 2013, von <http://newsroom.unl.edu/releases/2012/09/18/UNL+archaeological+team+unearths+giant+Roman+mosaic+in+southern+Turkey>
- ARD/ZDF. (2007). *Migranten und Medien*. Köln: Westdeutscher Rundfunk Köln.
- Arnheim, R. (1979). Bemerkungen zum Farbfilm (1935). In H. H. Diederichs (Hrsg.), *Kritiken und Aufsätze zum Film*. (S. 47–52). Frankfurt am Main: Fischer Taschenbuch Verlag.
- Arnheim, R. (2000). *Kunst und Sehen. Eine Psychologie des schöpferischen Auges*. (H. Hermann, Übers.) (Titel der engl. Originalausgabe von 1954: „Art and Visual Perception - A psychology of the creative eye“; 3., unveränd. Aufl.). Berlin; New York: De Gruyter Verlag.
- Austerlitz, S. (2008). *Money for Nothing: A History of the Music Video from the Beatles to the White Stripes* (1. Aufl.). London; New York City: Continuum International Publishing Group.
- Baltruschat, A. (2010). *Die Dekoration der Institution Schule: Filminterpretationen nach der dokumentarischen Methode*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Baltruschat, A., Bohnsack, R., & Pfaff, N. (Hrsg.). (2010). Die dokumentarische Methode in der Text-, Bild- und Videointerpretation. In *Enzyklopädie Erziehungswissenschaft Online (EEO)*. Weinheim: Juventa Verlag.
- Baltruschat, A., & Hampl, S. (2013). Das Bild im Film - der Film im Bild. Zur Interpretation von Filmen nach der dokumentarischen Methode. In P. Loos, A.-M. Nohl, A. Przyborski, & B. Schäffer (Hrsg.), *Dokumentarische Methode. Grundlagen - Entwicklungen - Anwendungen*. (S. 243–267). Leverkusen, Berlin, Toronto: Verlag Barbara Budrich.
- Bohnsack, R. (2001a). Die dokumentarische Methode in der Bild- und Fotointerpretation. In R. Bohnsack, I. Nentwig-Gesemann, & A.-M. Nohl (Hrsg.), *Die dokumentarische Methode und ihre Forschungspraxis: Grundlagen qualitativer Sozialforschung* (S. 67–90). Opladen: Leske + Budrich Verlag.
- Bohnsack, R. (2001b). „Heidi“: Eine exemplarische Bildinterpretation auf der Basis der dokumentarischen Methode. In R. Bohnsack, I. Nentwig-Gesemann, & A.-M. Nohl (Hrsg.), *Die dokumentarische Methode und ihre Forschungspraxis: Grundlagen qualitativer Sozialforschung* (S. 323–337). Opladen: Leske + Budrich Verlag.
- Bohnsack, R. (2003). Die dokumentarische Methode in der Bild- und Fotointerpretation. In Y. Ehrenspeck & B. Schäffer (Hrsg.), *Film- und Fotoanalyse in der Erziehungswissenschaft. Ein Handbuch* (1. Aufl., S. 87–107). Opladen: Leske + Budrich Verlag.
- Bohnsack, R. (2007). *Rekonstruktive Sozialforschung. Einführung in qualitative Methoden* (6., durchges. und aktualisierte Aufl.). Opladen & Farmington Hills: Verlag Barbara Budrich.
- Bohnsack, R. (2009). *Qualitative Bild- und Videointerpretation. Die dokumentarische Methode*. Opladen & Farmington Hills: Verlag Barbara Budrich.

- Bohnsack, R., & Przyborski, A. (2014). Pose, Lifestyle und Habitus in der Ikonik. In R. Bohnsack, B. Michel, & A. Przyborski (Hrsg.), *Dokumentarische Bildinterpretation. Methodologie und Praxis*. Leverkusen, Berlin, Toronto: Verlag Barbara Budrich.
- Bordwell, D. (1985). *Narration in the Fiction Film*. Madison: University of Wisconsin Press.
- Bortz, J., & Döring, N. (2002). *Forschungsmethoden und Evaluation: für Human- und Sozialwissenschaftler* (3. überarb. Aufl.). Berlin [u.a.]: Springer Verlag.
- Breckner, R. (2010). *Sozialtheorie des Bildes : Zur interpretativen Analyse von Bildern und Fotografien*. Bielefeld: transcript Verlag.
- Dawes, B. (2004). Cinema Redux. Abgerufen 30. April 2013, von <http://brendandawes.com/projects/cinemaredux>
- Deleuze, G. (1996). *Das Bewegungs-Bild: Kino 1* (5. Aufl.). Frankfurt am Main: Suhrkamp Verlag.
- Dinkelaker, J., & Herrle, M. (2009). *Erziehungswissenschaftliche Videographie: eine Einführung* (1. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Flusser, V. (1997). *Medienkultur* (5. Aufl.). Frankfurt am Main: Fischer Taschenbuch Verlag.
- Frost, J. B. (2009). *Cinematography for Directors: A Guide for Creative Collaboration*. Studio City: Michael Wiese Productions.
- Goethe, J. W. von. (2003). *Farbenlehre: 5 Bde.* (7. Aufl.). Stuttgart: Freies Geistesleben.
- Hampl, S. (2005). *Exemplarische Videotranskription von „Istanbul Total“ in Partiturschreibweise*. Videotranskript, Interpretationsvorlage und Vortrag im Rahmen der Forschungswerkstatt für qualitative Methoden (WS 2005/06). Institut für Publizistik und Kommunikationswissenschaft, Universität Wien.
- Hampl, S. (2006, Jänner). „Istanbul Total“. *Die mediale Artikulation eines transnationalen Kultur- und Erfahrungsraumes. Dokumentarische Methode/Videointerpretation, Video bzw. Video-Transkription*. Vortrag beim 9. Bundesweiten ZBBS-Workshop zur Qualitativen Bildungs- und Sozialforschung, Otto-von-Guericke-Universität Magdeburg.
- Hampl, S. (2008). *MoviScript - Software zur Videotranskription. (Computerprogramm)*. Wien. Abgerufen von [www.moviscript.net](http://www.moviscript.net) am 30. April 2013.
- Hampl, S. (2010). Videos interpretieren und darstellen. Die dokumentarische Methode. In F. Corsten, M. Krug, & C. Moritz (Hrsg.), *Videographie praktizieren. Herangehensweisen, Möglichkeiten und Grenzen* (S. 53–88). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Hara, K. (2009). *Weiss* (1. Aufl.). Baden: Lars Müller Publishers.
- Herrle, M. (2007). *Selektive Kontextvariation: die Rekonstruktion von Interaktion in Kursen der Erwachsenenbildung auf der Basis audiovisueller Daten*. Frankfurt am Main: Fachbereich Erziehungswissenschaft der Johann-Wolfgang-Goethe-Universität.
- Hickethier, K. (2012). *Film- und Fernsehanalyse* (5., aktualisierte und erweiterte Auflage.). Stuttgart; Weimar: Metzler Verlag.

- IFPI. (2012). *Digital Music Report*. London: International Federation of the Phonographic Industry.
- Imdahl, M. (1994). Ikonik. Bilder und ihre Anschauung. In G. Boehm (Hrsg.), *Was ist ein Bild?* (S. 300–324). München: Wilhelm Fink Verlag.
- Imdahl, M. (1996a). *Giotto. Arenafresken: Ikonographie - Ikonologie - Ikonik* (3. Aufl.). München: Wilhelm Fink Verlag.
- Imdahl, M. (1996b). *Gesammelte Schriften 3. Reflexion - Theorie - Methode*. (G. Boehm, Hrsg.) (Bd. 1-3, Bd. 3). Frankfurt am Main: Suhrkamp Verlag.
- Imdahl, M. (2003). *Farbe* (3., unveränderte Aufl.). München: Wilhelm Fink Verlag.
- Itten, J. (2000). *Kunst der Farbe: Subjektives Erleben und objektives Erkennen als Wege zur Kunst* (Erstauf. 1961; 13., Aufl.). Ravensburg: Otto Maier Verlag.
- Itten, J. (2010). *Kunst der Farbe. Studienausgabe: Subjektives Erleben und objektives Erkennen als Wege zur Kunst* (Neuaufgabe.). Stuttgart: Kreuz Verlag.
- Juen, M. (2012). *Die ersten Minuten des Unterrichts: Skizzen einer Kairologie des Anfangs aus kommunikativ-theologischer Perspektive* (unveröff. Dissertation). Universität Innsbruck.
- Kandinsky, W. (1986). *Punkt und Linie zu Fläche. Beitrag zur Analyse der malerischen Elemente* (8. Aufl.). Sulgen: Benteli Verlag.
- Keppeler, A. (2006). *Mediale Gegenwart: eine Theorie des Fernsehens am Beispiel der Darstellung von Gewalt* (1. Aufl.). Frankfurt am Main: Suhrkamp Verlag.
- Klee, P. (1990). *Form- und Gestaltungslehre: Das bildnerische Denken: Form- und Gestaltungslehre, 1: BD 1* (5. Aufl.). Basel: Schwabe Verlag.
- Knoblauch, H. (2004). Die Video-Interaktions-Analyse. *Sozialer Sinn*, 5(1), 123–138.
- Krämer, S. (1998). Medium als Spur und als Apparat. In S. Krämer (Hrsg.), *Medien - Computer - Realität: Wirklichkeitsvorstellungen und Neue Medien* (4. Aufl., S. 73–94). Frankfurt am Main: Suhrkamp Verlag.
- Latour, B. (1998). Über technische Vermittlung. Philosophie, Soziologie, Genealogie. In W. Rammert (Hrsg.), *Technik und Sozialtheorie* (S. 29–81). Frankfurt/Main; New York: Campus Verlag.
- Latour, B. (2001). *Das Parlament der Dinge: Für eine politische Ökologie* (1. Aufl.). Frankfurt am Main: Suhrkamp Verlag.
- Latour, B. (2002). *Die Hoffnung der Pandora: Untersuchungen zur Wirklichkeit der Wissenschaft* (4. Aufl.). Frankfurt am Main: Suhrkamp Verlag.
- Latour, B. (2008). *Wir sind nie modern gewesen - Versuch einer symmetrischen Anthropologie* (1. Aufl.). Frankfurt am Main: Suhrkamp Verlag.
- Malewitsch, K. (1994). *Suprematismus. Die gegenstandslose Welt*. (Erstveröffentlichung: 1927.). Köln: DuMont Literatur und Kunst Verlag.
- Mannheim, K. (2003). *Strukturen des Denkens*. Frankfurt am Main: Suhrkamp Verlag.
- McLuhan, H. M. (1992). *Die magischen Kanäle. Understanding Media* (Neuausgabe.). Düsseldorf; Wien; New York; Moskau: ECON Verlag.
- Mikos, L. (2003). *Film- und Fernsehanalyse*. Konstanz; Stuttgart: UVK Verlag; UTB Verlag.

- Mikunda, C. (2002). *Kino spüren: Strategien der emotionalen Filmgestaltung* (Neuauf.). Wien: WUV Verlag.
- Mohn, B. E. (2011). Methodologie des forschenden Blicks. Die vier Spielarten des Dokumentierens beim ethnographischen Forschen. In P. Cloos & M. Schultz (Hrsg.), *Kindliches Tun beobachten und dokumentieren* (S. 79–98). Weinheim: Juventa Verlag.
- Monaco, J. (2009). *Film verstehen: Kunst, Technik, Sprache, Geschichte und Theorie des Films und der Neuen Medien* (2. Aufl.). Reinbek: rororo Verlag.
- MPFS. (2012). *JIM-Studie 2012. Jugend, Information, (Multi-)Media. Basisstudie zum Medienumgang 12- bis 19-Jähriger in Deutschland*. Stuttgart: Medienpädagogischer Forschungsverbund Südwest (MPFS).
- Müller, M. R. (2012). Figurative Hermeneutik. Zur methodologischen Konzeption einer Wissenssoziologie des Bildes. *Sozialer Sinn*, 13(1), 129–163.
- Muybridge, E. (2009). Plate 44. In *Animal locomotion: an electro-photographic investigation of consecutive phases of animal movements (1872-1885)*. Boston: Boston Public Library. Abgerufen von [http://www.flickr.com/photos/boston\\_public\\_library/4325570997](http://www.flickr.com/photos/boston_public_library/4325570997) am 30. April 2013.
- Panofsky, E. (2006). *Ikonographie und Ikonologie* (1. Aufl.). Köln: DuMont Literatur und Kunst Verlag.
- Przyborski, A., & Slunecko, T. (2009). Against Reification! Praxeological Methodology and its Benefits. In J. Valsiner, P. C. M. Molenaar, M. C. D. P. Lyra, & N. Chaudhary (Hrsg.), *Dynamic Process Methodology in the Social and Developmental Sciences* (S. 141–170). Springer US. Abgerufen von <http://www.springerlink.com/content/w5731h765ul140tw/abstract/>
- Przyborski, A., & Slunecko, T. (2011). Learning to think iconically in the human and social sciences: iconic standards of understanding as a pivotal challenge for method development. *Integrative Psychological and Behavioral Science*. doi:10.1007/s12124-011-9159-6
- Przyborski, A., & Slunecko, T. (2013). Ikonizität – medientheoretisch gedacht und empirisch beleuchtet. In P. Loos, A.-M. Nohl, A. Przyborski, & B. Schäffer (Hrsg.), *Dokumentarische Methode. Grundlagen - Entwicklungen - Anwendungen*. (S. 189–212). Leverkusen, Berlin, Toronto: Verlag Barbara Budrich.
- Przyborski, A., & Wohlrab-Sahr, M. (2009). *Qualitative Sozialforschung. Ein Arbeitsbuch*. (2., korrigierte Aufl.). München: Oldenbourg Wissenschaftsverlag.
- Pudowkin, W. I. (1999). Filmregie und Filmmanuskript. In F.-J. Albertsmeier (Hrsg.), *Texte zur Theorie des Films* (5., durchges. u. erw. Aufl., S. 70–73). Stuttgart: Reclam Verlag.
- Qantara. (2013). Representation of Irene and John II Comnenus on a mosaic in Hagia Sophia, Constantinople. *Qantara. Mediterranean Heritage*. Abgerufen 30. April 2013, von [http://www.qantara-med.org/qantara4/public/show\\_document.php?do\\_id=727&lang=en](http://www.qantara-med.org/qantara4/public/show_document.php?do_id=727&lang=en)
- Quotenmeter. (2004, Mai 18). Quotenmeter.de - Das Online-Fernsehmagazin. «TV Total» mit Top-Quoten. Abgerufen 24. August 2007, von <http://www.quotenmeter.de/>

- Raab, J. (2008). *Visuelle Wissenssoziologie. Theoretische Konzeption und materiale Analysen*. Konstanz: UVK Verlag.
- Rabiger, M. (1998). *Directing the documentary* (3rd ed.). Boston: Focal Press.
- Ruck, N., & Slunecko, T. (2008). A Portrait of a Dialogical Self: Image Science and the Dialogical Self. *International Journal for Dialogical Self*, 3(1), 261–290.
- Salavon, J. (2012). One Week Skin (ESPN-Hs). Abgerufen 30. April 2013, von <http://salavon.com/work/one-week-skin/image/405/>
- Schäffer, B. (2001). „Kontagion“ mit dem Technischen. Zur generationsspezifischen Einbindung in die Welt medientechnischer Dinge. In R. Bohnsack, I. Nentwig-Gesemann, & A.-M. Nohl (Hrsg.), *Die dokumentarische Methode und ihre Forschungspraxis. Grundlagen qualitativer Sozialforschung* (S. 43–64). Opladen: Leske + Budrich Verlag.
- Schnettler, B. (2007). Auf dem Weg zu einer Soziologie visuellen Wissens. *Sozialer Sinn*, 8(2), 189–210.
- Schnettler, B., & Raab, J. (2008). Interpretative Visual Analysis. Developments, State of the Art and Pending Problems. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 9(3). Abgerufen von <http://www.qualitative-research.net/index.php/fqs/article/view/1149/2555> am 30. April 2013.
- Schnettler, B., Tuma, R., & Schreiber, S. (2010). Präsentation – Demonstration – Rezeption: Visualisierung der Wissenskommunikation. In *Unsichere Zeiten. Herausforderungen gesellschaftlicher Transformationen. Verhandlungen des 34. Kongresses der Deutschen Gesellschaft für Soziologie in Jena 2008*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schumm, G. (2004). Montage. In H. Schättle & D. Wiedemann (Hrsg.), *Bewegte Bilder - Bewegte Zeit. Mit CD-ROM* (1., Aufl., S. 167–169). Berlin: Vistas Verlag.
- Schütze, F., Meinefeld, W., & Weymann, A. (1973). Grundlagentheoretische Voraussetzungen methodisch kontrollierten Fremdverstehens. In Arbeitsgruppe Bielefelder Soziologen (Hrsg.), *Alltagswissen, Interaktion und gesellschaftliche Wirklichkeit*. (Bd. 2, S. 433–495). Reinbek: Rowohlt.
- Seitz, D. (2010, August 5). 5 Annoying Trends That Make Every Movie Look the Same. *Cracked.com*. Abgerufen 15. Dezember 2012, von [http://www.cracked.com/article\\_18664\\_5-annoying-trends-that-make-every-movie-look-same.html](http://www.cracked.com/article_18664_5-annoying-trends-that-make-every-movie-look-same.html)
- Sloterdijk, P. (2004). *Sphären. Plurale Sphärologie: Band III: Schäume: BD III* (1. Aufl.). Frankfurt am Main: Suhrkamp Verlag.
- Slunecko, T. (2008). *Von der Konstruktion zur dynamischen Konstitution: Beobachtungen auf der eigenen Spur* (2., überarb. Aufl.). Wien: Facultas WUV Verlag.
- Slunecko, T., & Przyborski, A. (2009). Kulturdialog als Mediendialog. *Journal für Psychologie*, 17(2). Abgerufen von <http://www.journal-fuer-psychologie.de/jfp-2-2009-05.html> am 30. April 2013.
- Sobotka, M.-L. (2009). *Humor und Geschlechterverhältnis im Unterhaltungsformat „Dancing Stars“: eine rekonstruktive Studie* (unveröff. Diplomarbeit). Universität Wien.

- Souriau, E. (1997). Die Struktur des filmischen Universums und das Vokabular der Filmologie [frz. 1951]. *Montage AV*, 6(2), 140–157.
- Valsiner, J., Molenaar, P. C. M., Lyra, M. C. D. P., & Nandita, C. (2009). Dynamic Nature of Phenomena and its Study. In J. Valsiner, P. C. M. Molenaar, M. C. D. P. Lyra, & N. Chaudhary (Hrsg.), *Dynamic Process Methodology in the Social and Developmental Sciences* (S. v–xii). Springer US. Abgerufen von <http://www.springerlink.com/content/w5731h765ul140tw/abstract/>
- Vitouch, P. (2007). *Fernsehen und Angstbewältigung: zur Typologie des Zuschauerhaltens* (3. Auflage.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Walter, M., Schlinker, U., & Fischer, C. (2007). Fernsehnutzung von Migranten. Ergebnisse der ARD/ZDF-Studie „Migranten und Medien 2007“. *Media Perspektiven*, 9, 436–451.
- Wiedemann, D. (2005). Film und Fernsehen. In *Bildwissenschaft: Disziplinen, Themen, Methoden* (1. Aufl., Originalausg., S. 365–380). Frankfurt am Main: Suhrkamp Verlag.
- Wulff, J. (2013). Lexikon der Filmbegriffe. Kiel: Christian-Albrechts-Universität. Abgerufen von <http://filmlexikon.uni-kiel.de> am 30. April 2013.

## **Anhang: Vorarbeiten zur inhaltlichen Interpretation von „Istanbul Total“**

### **Interpretationsansatz 1: Kommerzieller Verwertungskontext**

Bereits bei einem flüchtigen Blick auf die häufigsten Einstellungen sowie die formale Abfolge der Gesten und Einstellungswechsel in der Eingangspassage der Fernsehshow „Istanbul Total“ (Abbildung 6) lassen sich drei zentrale Merkmale der Sendung erkennen: 1) In den meisten Einstellungen (am Balkon und im Studio) ist der Moderator Stefan Raab zu sehen, 2) eingelagerten Sequenzen geht regelmäßig eine Geste des Zeigens von Stefan Raab voraus, außer 3) es handelt sich um Sequenzen, auf denen entweder das Publikum oder die Studiogäste Elton und Gülcan zu sehen sind (Letztere werden durch Close-ups identifizierbar gemacht, wenn Stefan Raab sich ihnen zuwendet). Der Moderator ist beinahe ständig im Bild zu sehen. Bildwechsel ereignen sich nur auf seine Anweisung hin oder um akklamierendes Publikum ins Bild zu holen. Bohnsack (2009) fasst diese Beobachtungen im Rahmen seiner Rekonstruktion der Eingangspassage unter dem Begriff der „Monostrukturierung durch den Showmaster“ zusammen. Das Sendungsgeschehen wird von Stefan Raab durchgehend kontrolliert und dirigiert. Homologien dazu zeigen sich auf der Textebene darin, dass Raab stets die Gesprächsthemen bestimmt. Wie die komparative Analyse mit einem Ausschnitt der Beyaz-Show (Kap. 6) zeigt, der am zweiten Sendetag in „Istanbul Total“ präsentiert wurde, werden sogar Mittel der Postproduktion eingesetzt, um im Nachhinein Unsicherheiten und Unbeholfenheit des Moderators zu kaschieren und ihn in einem sendungskonformen Licht erscheinen zu lassen. In dieser Hinsicht ist Bohnsack (2009) zuzustimmen, dass die „Struktur der Sendung insgesamt und der individuelle Habitus des Moderators“ nicht mehr zu trennen sind. Daraus ergibt sich die Frage nach den Produktionsverhältnissen der Fernsehproduktion, die ihrerseits wiederum mit spezifischen Erfahrungsräumen außerhalb der TV-Show „Istanbul Total“ verknüpft sind. Was lässt sich formal über das Milieu aussagen, das von TV-Shows lebt?

Bohnsack (2009) macht hinsichtlich der Fernsehsendung „Istanbul Total“ auf die Rahmenkomponente des „parasitären Aufmerksamkeitsgewinnes“ aufmerksam, wonach sich die TV-Sendung im Kern insbesondere durch Bezugnahme auf andere mediale Großereignisse speise. Wie die Analyse der Interviews zeigt, die Stefan Raab in seinem Studio am Bosphorus mit zahlreichen Gästen führt, lässt sich dieses Konzept auf einflussreiche Einzelpersonen erweitern: seien es der Songcontest-Kandidat Max

Mutzke, Nina Öger (die Geschäftsführerin von Öger-Tours), der türkische Popsänger Mustafa Sandal, der Fußballtrainer Christoph Daum oder andere.

Jedoch dokumentieren sich in den Interviews nicht nur unilaterale Austauschrelationen. Es sind auch deutliche Synergien zwischen den Beteiligten erkennbar. Entsprechend der oben erwähnten Monostrukturierung durch den Moderator wird den Interviewpartner/inne/n dezidiertes Raum zur Selbstpräsentation eingeräumt, sei es, um ihr aktuelles Musikvideo, ihre neue CD oder ihre neue Modedkollektion vorzustellen bzw. schlicht für sich als Marke und Firma zu werben. Mitunter wäre somit öffentliche Vermarktung tatsächlich ein plausibles Argument für die unternehmerische Relevanz von Mono- bzw. schlicht Machtstrukturierung (s. o.). Flusser (1997, S.105) spricht in diesem Zusammenhang auch vom „ethischen, Missbrauch“ der Zusehenden durch das Fernsehen:

Gegenwärtig wird das Fernsehen verwendet, um dem Empfänger seiner Botschaften zu einem spezifischen Verhalten zu führen, nämlich zum Verbrauch jener physischen und geistigen Güter, an denen die Besitzer des Fernsehsystems interessiert sind (Flusser, 1997, S. 105).

Unbeantwortet bleibt dabei vorerst die Frage, wie es gelingt, Nachfrage zu generieren. Stefan Raab und seine Gäste können in diesem Zusammenhang jedoch als Marken betrachtet werden, die sich durch die gemeinsame Performance in einer Fernsehshow gegenseitige Popularitätsgewinne, Reichweite und Marktanteile verschaffen.

Die genauere Betrachtung der Unternehmensverflechtungen legt diesen Verdacht nahe. Die Sendungen „Istanbul Total“ und „TV Total“ sind beides Marken der Produktionsfirmen „Raab TV“ und „Brainpool TV GmbH“. Eigentümer der Firma „Raab TV“ waren zum Zeitpunkt der Sendungen zu 50% Stefan Raab selbst und zu 50% die „Brainpool TV GmbH“<sup>107</sup>. Zugleich war die „Brainpool TV GmbH“ ihrerseits ein 100%iges Tochterunternehmen der „Viva Media GmbH“<sup>108</sup>. Die „Viva Media GmbH“ produzierte über eine andere 100%ige Tochter, die „Viva Fernsehen GmbH“, wiederum das Musikfernsehen Viva, auf dem etwa das Musikvideo „Araba“ (2004) von Mustafa Sandal lief und promoted wurde. Im August 2004 übernahm das amerikanische Konkurrenzunternehmen „Viacom“ die „Viva Media GmbH“ zu 98%. Zu „Viacom“ gehört auch der weltweit bekannteste Musiksender „MTV“. Durch den Kauf von Viva konnte

---

<sup>107</sup> Mit 31. Dezember 2008 wurde die „Brainpool TV GmbH“ 100%ige Eigentümerin von „Raab TV“.

<sup>108</sup> Mit 1. Jänner 2007 kam es zum Management-Buy-Out, wodurch die „Brainpool TV GmbH“ heute wieder ein eigenständiges Unternehmen ist.

„Viacom“ damit seine angestrebte weltweite Monopolstellung für Musiksender weiter ausbauen.

So wirken sich die Songcontestpräsentation von Max Mutzke sowie die Präsentation des Musikvideos „Araba“ von Mustafa Sandal in der TV-Show „Istanbul Total“ nicht nur unmittelbar positiv auf die Produktionsfirma dieser Sendung („Raab TV“) aus, sondern in weiterer Folge auch auf die aktienhaltenden Mutterunternehmen im Hintergrund („Brainpool TV GmbH“, „Viva Media“, „Viacom“ etc.). Des Weiteren profitieren der ausstrahlende Fernsehsender ProSieben, der Musikkonzern Edel Music<sup>109</sup> (Musikverkauf von Max Mutzke), der Musikkonzern Polydor<sup>110</sup> (Musikverkauf von Mustafa Sandal), die Europäische Rundfunkunion als Ausrichterin des „Eurovision Songcontest“ sowie dessen Sponsoren, die Werbeunternehmen sowie die Musikindustrie etc. Zusammenfassend repräsentieren die in der Fernsehshow „Istanbul Total“ auftretenden Gäste somit auf unterschiedlichen Ebenen Shareholderinteressen. Durch die Unterhaltungssendung „Istanbul Total“ entstehen nachhaltige finanzielle Multiplikatoreffekte, in welche sowohl die Protagonist/inn/en als auch die eingeladenen Gäste eingebunden sind.

„Istanbul Total“ wird im Kern von vier Personen getragen, denen in der Show spezifische mediale Funktionen zugeordnet sind: Stefan Raab ist der Moderator, Elton sein Gehilfe, Max der Gesangsstar und die Deutsch-Türkin Gülcan schafft die Verbindung zur türkischen Sprache und Kultur. Ihre Showauftritte sind für die Beteiligten sehr marketingwirksam – und dies bringt nicht nur ihnen persönlich, sondern auch weiteren Medienproduktionen etwas: So hat Elton etwa selbst eine Show namens „Elton TV“, die auf dem deutschen Musiksender Viva regelmäßig ausgestrahlt wird. Gülcan wiederum moderiert bei Viva die Music-News. Weil Max populär ist, wird auf Viva täglich sein Musikvideo gezeigt. Vivas Einschaltquote erhöht sich proportional zum Erfolg von „TV Total“. Dies ist in Anbetracht der Unternehmensverschachtelungen ein mehrfacher Erfolg. Dem ausstrahlenden Privat-TV-Sender ProSieben hat Stefan Raab mit „Istanbul Total“ übrigens einen Zugang zur European Broadcasting Union (EBU), dem Veranstalter des Songontests, ermöglicht. Zum Zeitpunkt der Sendung waren aus deutscher Sicht in der EBU nämlich nur die öffentlich-rechtlichen Sender (ARD und ZDF) vertreten.

---

<sup>109</sup> Die Edel-AG ist in der Musikindustrie eines der letzten und größten noch unabhängigen Musikunternehmen Europas (sog. „Independent Label“).

<sup>110</sup> Polydor gehört zu Universal Music, einem der fünf größten Musikkonzerne der Welt (sog. „Major Label“).

Zusammenfassend dokumentiert sich anhand des Beispiels „Istanbul Total“, dass jede Person, jede Maßnahme, jedes Medienevent, welche/s in dieser Sendung vorkommt, eine unmittelbare multiplikative und extensive Wirkung auf sich selbst sowie auch auf die mit ihr in Verbindung stehenden Personen, Sendungen, Medienereignisse und Unternehmen darstellt.

Die Hauptmarke „TV Total“, von der „Istanbul Total“ ja nur ein Ableger war, macht, wie der Name der Sendung schon sagt, das Fernsehen selbst, also mediale Produkte zum Gegenstand (des medialen Produkts). Sparsamer, mithin ökonomischer geht es kaum. Die Bezugnahme auf bereits erfolgreich Produziertes erlaubt eine äußerst schlanke Produktion und erzeugt eine Menge wirtschaftlich erwünschter (ökonomischer) Synergien. Das rechtfertigt sogar die kurzzeitige Übersiedlung der Produktion nach Istanbul. Denn indem „TV Total“ dadurch dem verstaubten Songcontest massive PR zuteilwerden lässt, versorgt es sich selbst mit Inhalt und stärkt noch seine eigene Marktposition als innovatives und lebensnahes Unterhaltungsmedium.

Im Kontext der kommerziellen Verbindungen, die im Zusammenhang mit der Fernsehsendung „Istanbul Total“ und ihren Protagonist/inn/en stehen, lassen sich die übergeordneten Orientierungsrahmen der Show besser verstehen. Die kommerziellen Hintergründe sind durch den Sendungshabitus der ständigen Ironisierung und Distanzierung gut kaschiert. Als für die Sendung „Istanbul Total“ strukturell bedeutsame Orientierungsrahmen treten sie aber durch komparative Analyse zutage.

### **Interpretationsansatz 3: Sexismus und Chauvinismus**

In der Fernsehsendung „Istanbul Total“ wie auch in der Version des Musikvideos „Araba“ (2004) für den internationalen Markt ist die Gleichberechtigung zwischen Männern und Frauen ein positiver Horizont, der durchgängig implizit wirksam ist, ohne offen thematisiert zu werden. In verdeckter Weise manifestiert sich der Stellenwert der Gleichberechtigung am in der Sendung immer wieder aufgeworfenen, negativen Gegenhorizont des Sexismus und Chauvinismus, der hauptsächlich den Türken zugeschrieben und vorgeworfen wird. Im Hinblick auf die Monostrukturierung durch den Moderator Stefan Raab ist noch zu bemerken, dass sexistische und chauvinistische Aussagen sehr wohl von seinen Assistenten Elton und seinem Off-Kommentator (vgl. Beyaz-Show in Kap. 6) gemacht werden können (vgl. Sequenz „Max‘ Konkurrenz“). Dies erlaubt es, die Figur Stefan Raabs so weit moralisch zu immunisieren, dass sie von einem sicheren Standort aus mit dem Finger auf türkische Stereotypen zeigen kann.

Ein gutes Beispiel für die deutliche Zuschreibungstendenz des Sexismus und Chauvinismus in „Istanbul Total“ ist die komparative Analyse der „Harems-Sequenz“ (1. Sendetag von „Istanbul Total“) mit dem ZDF-Reisevideo „Reiselust Istanbul“. In beiden Fällen wird der Topkapi-Palast gezeigt, wobei dem „faszinierendste[n] Teil des Topkapi-Palastes“ (Reiselust Istanbul, 5:16), dem Harem, jeweils beinahe der gesamte Beitrag gewidmet ist. Im Gespräch mit dem türkischen Reiseführer stellt Stefan Raab unablässig anzügliche Fragen im Hinblick auf das Sexualleben des Sultans: „Was war der beste Anmachspruch vom Sultan? Weiß man das? ... Mach hin Alte, ich bin der Sultan.“ (Istanbul Total, 562-00-03, 2:55). Der Harem wird in diesem Kontext somit auf dem Niveau eines Bordells abgehandelt, wobei der Orientierungsrahmen des Sexismus dabei aber von Raab selbst vorgegeben wird.

Im ZDF-Reisevideo werden demgegenüber auf der Bild- und Textebene das Geheimnisvolle des Harems („Harem kommt aus dem Arabischen und bedeutet unzugänglich“, Reiselust Istanbul, 5:22) sowie die sozialen Errungenschaften der dort lebenden Frauen betont („Kaum bekannt ist jedoch, dass die Frauen im Harem eine ganz normale Ausbildung erhielten“, Reiselust Istanbul, 5:38).

Im Kontrast zum stereotypen, sexistischen Verhalten der Türken übt sich der Moderator Stefan Raab in – für seine Verhältnisse (d. h. verglichen mit anderen Stereotypen, die von ihm in der Show elaboriert werden) – fast protestantischer Ergebenheit gegenüber jenen Frauen, die zum Interview in sein Studio eingeladen sind. Beim Interview mit Nina Öger beispielsweise, der Geschäftsführerin von Öger Tours (des weltweit größten Reiseunternehmens für Türkeireisen), besteht auf der Bildebene eine formal gleichberechtigte Raumaufteilung (Abbildung 59).



**Abbildung 59: Interview Nina Öger (4:23)**

Im Gespräch gibt Raab Frau Öger (Abbildung 59) ausreichend Raum für ihre persönlichen Äußerungen und unterlässt Schlüpfrigkeiten wie jene in der Haremsequenz. Dies ist umso bemerkenswerter, als die Haremsequenz direkt auf das Interview mit Nina Öger folgt. Gegenüber Frau Öger bemüht sich Raab auch um die korrekte türkischer Aussprache des Wortes „Topkapı“ wobei die türkische Sprache im Rest der Sendung entweder ignoriert, absichtlich missverstanden oder verballhornt (s. u.) wird. So wie bereits gegenüber der Interviewpartnerin Gülcan (vgl. Bohnsack, 2009) erlaubt sich Raab lediglich stereotype Äußerungen über die kleinen Gaunereien der Türken beim Herstellen, Verkaufen und Handeln von Waren.

Neben den Interviews wird das Prinzip der Gleichberechtigung zwischen Mann und Frau in den Episoden „Elton & Gülcan unterwegs in Istanbul“ wiederum im Sinne einer dumpfen Gleichmacherei zwischen den Geschlechtern verballhornt. Die gemeinsamen Aktivitäten der beiden Protagonisten gehen in Istanbul sogar so weit, dass sie dieselben Frauenkleider anprobieren (561-00-03, 1:59). Durch Eltons „Körperfülle“ (Raab, 561-00-01) entsteht in den knappen Frauenkleidern eine ironische Brechung des männlichen Geschlechtsstereotyps (Abbildungen 61 und 62). Elton macht sich jedoch nicht nur lächerlich, sondern ist in Frauenkleidern zugleich auch irgendwie cool. Damit handelt es sich im Falle seiner Performance letztlich ebenso um Sexismus.



**Abbildung 60: Elton & Gülcan 1 (0:52)**



Abbildung 61: Elton & Gülcan 1 (2:12)



Abbildung 62: Elton & Gülcan 1 (3:30)

Hinsichtlich der Verkleidung Eltons gibt es deutliche Homologien zur Eingangssequenz. Dort ist er wie ein Tourist gekleidet und nimmt damit eine vermittelnde „Zwischenstellung zwischen Bühnenakteur und abgebildetem Zuschauer“ ein (Bohnsack, 2009). Im Falle der vorliegenden *ironischen* Verkleidung jedoch erfüllt sich – wie bei Stefan Raab im Sultansgewand (vgl. Eingangspassage) – der Aspekt der Distanzierung. Genauer gesagt handelt es sich um die Distanzierung vom Ästhetisch-Weiblichen bzw. Sexuell-Anzüglichen. Die Anrühigkeit von beidem gilt es in „Istanbul Total“ zu vermeiden, wie sich an zwei Beispielen erläutern lässt.

Am dritten Sendetag sind Ersin Akalilar (der Geschäftsführer von „Mavi Jeans“) sowie der türkische Popsänger Mustafa Sandal zu Gast in der Show. Akalilar präsentiert seine aktuelle Modekollektion, Sandal sein neuestes Musikvideo. Der reizvolle Auftritt der Models in kurzen Röcken wird von Elton in Damenkleidern abgeschlossen. Nach

Mustafa Sandals erotischem Musikvideo „Araba“<sup>111</sup> werden das einzige Mal in der „Istanbul-Total“-Show Frauen mit Kopftüchern gezeigt; sie sitzen im Publikum und sind für vier Sekunden im Bild (563-00-04, 6:48–6:51). In beiden Fällen wird dem Anrühigen bildlich Einhalt geboten: das erste Mal durch eine Leistung eines abgebildeten Bildproduzenten, das zweite Mal durch eine Leistung der abbildenden Bildproduzentinnen.



Abbildung 63: Mavi Modeschau (3:05)



Abbildung 64: Mavi Modeschau (3:17)

In Analogie zu Eltons bauchfreiem Auftritt kann das Zeigen der Frauen mit Kopftüchern (Abbildung 66) sowohl im Sinne der ironischen Distanzierung als auch im Sinne der Gleichwertigkeit von Frauen mit und ohne Kopftuch verstanden werden.

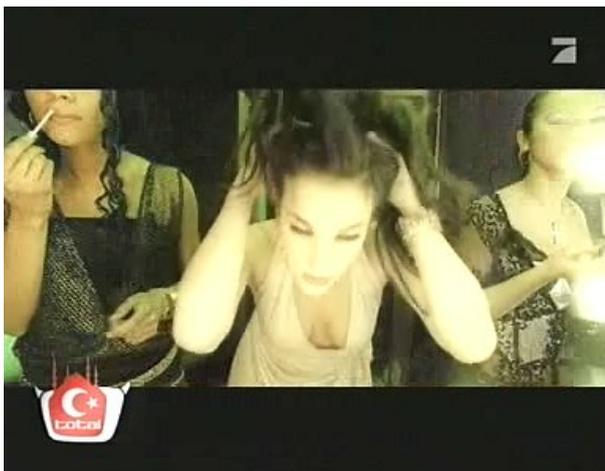


Abbildung 65: Musikvideo „Araba“ (2004) (6:18)



Abbildung 66: Frauen mit Kopftuch (6:49)

Bemerkenswert ist in jedem Fall, dass diese Einstellung gewählt wurde, obwohl Kopftücher bzw. damit in Verbindung stehende Themen wie Islam oder

<sup>111</sup> Das Musikvideo wurde in dieser erotisierten Bildversion speziell für den deutschen Markt produziert. Die türkische Bildversion des Videos enthält keine anzüglichen Elemente.

Fundamentalismus sonst nicht in der Sendung vorkommen. Nur zwei Monate vor der Ausstrahlung von „Istanbul Total“ hatten islamistische Terroristen in Madrid eines der größten Zugattentate der Geschichte verübt.

#### **Interpretationsansatz 4: Islam und Fundamentalismus**

Die Themen Islam und Fundamentalismus werden in „Istanbul Total“ umgangen. Dies wird bereits in der Eingangspassage deutlich, als Stefan Raab die Hagia Sofia als „eine der ältesten Kirchen“ bezeichnet. Weiters werden, wie im vorigen Kapitel angesprochen, in „Istanbul Total“ nur ein einziges Mal Frauen mit Kopftuch gezeigt – zur Neutralisierung der weiblichen Reize der Darstellerin im Musikvideo. Zum Kontrast: Im ZDF-Reisevideo „Reiselust Istanbul“, das aus demselben Jahr (2004) stammt, werden in den Straßen der Stadt sowohl Frauen mit Kopftüchern gezeigt als auch die Türkei als islamisches Land ausgewiesen (Reiselust Istanbul, 21:45). Im amerikanischen Reisevideo „Travelers Istanbul“ (1997) werden die ausländischen Vorurteile gegenüber Fundamentalismus offen angesprochen (Travelers Istanbul, 15:05) und von einer türkischen Reiseleiterin sogar eine Stellungnahme zum Islam abgegeben (11:00), ebenso im Dumont-Reisevideo „Istanbul OnTour“ (1997), wo sowohl von Fundamentalismus berichtet wird als auch Frauen mit Kopftuch auf Istanbuls Straßen gezeigt werden.

Dass Islam und Fundamentalismus normalerweise Themen sind, die von Stefan Raab sehr wohl behandelt werden, zeigt sich an einer „TV-Total“-Folge, die eine Woche vor „Istanbul Total“ gesendet wurde. Darin kommentiert Raab eine Zeitungsmeldung über Afghanistan. Er suggeriert dabei den terroristischen Hintergrund des dortigen „Bildungsfernsehens“ Sesamstraße. Die folgende Abbildung zeigt Raab bei der Erläuterung, wie Kinder in Afghanistan das Zählen lernen (Abbildung 67). Ansonsten gibt es noch Wortspiele mit anderen „deutschen TV-Sendungen“, die in Afghanistan angeblich „nachgemacht wurden“: „Der Mulla von Tölz“, „Guantanamo Bay-Watch“ oder die Talkshow „Bärbel Schläfer“. Auch hier sind die Bezüge zum Islam sowie zum islamischen Fundamentalismus und Terrorismus unverkennbar. „Mullas“ sind religiöse muslimische Führer, „Guantanamo Bay“ bezeichnet das Strafgefangenenlager für mutmaßliche Al-Quaida-Anhänger auf Kuba, als „Schläfer“ wurden nach den Anschlägen vom 11. September 2001 die islamistischen Terroristen bezeichnet, die angepasst und unauffällig unter uns leben, bis sie eines Tages losschlagen.

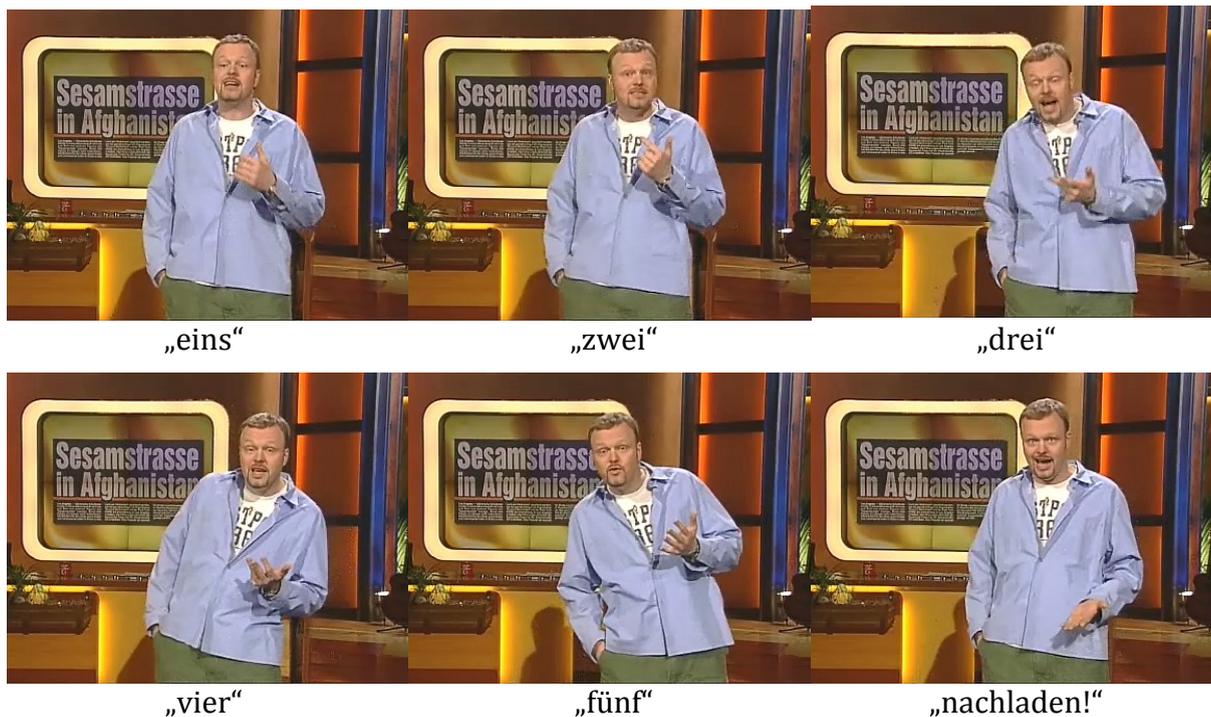


Abbildung 67: TV Total, 4. Mai 04 (0:24-0:29)

### Interpretationsansatz 5: Gewalt und Gefahr

Ebenso wie Islam und Fundamentalismus gilt es in „Istanbul Total“ die Themen Gewalt und Gefahr zu vermeiden. Erste Hinweise darauf finden sich ebenfalls bereits in der Eingangspassage, wo das „Stereotyp der Türken als kleine Gauner und Schwindler“ (Bohnsack, 2009) – in Abgrenzung zum Schwerverbrecher – entwickelt wird. Beschrieben und diskutiert werden alltägliche Delikte, bei denen niemand größeren Schaden nimmt. Beim Handeln am Bazar hat man als Kunde vielleicht einmal das Nachsehen, ein anderes Mal profitiert man davon beim Erwerb von billigen Fälschungen oder Raubkopien. Diese Form des charmanten Betrugs wird auch im Reisevideo „Travelers“ erwähnt, als die Moderatoren von einem Eisverkäufer zu wenig Retourgeld bekommen (Eisverkäufersequenz, 23:40).

Echte Gewalt und Gefahr gingen in Istanbul bestenfalls von den osmanischen Sultanen aus, deren Begehren aber historisch rasch Einhalt geboten werden konnte. So zumindest ist es den Raabs Kommentaren zum Sultan zu entnehmen: „Hat er mal gesagt, passt mal auf, Freunde, ich komm mal nach Wien. Ich bring ein paar Leute mit. Hat er aber die Rechnung ohne den Wirt gemacht. In Wien war nämlich Schluss.“ (Topkapi-Palast, 3:30–4:52)

Dass Stefan Raab selbst am ersten Sendetag so tut, also ob er ein osmanischer Sultan wäre (etwa durch seine folkloristische Verkleidung mit Turban), erscheint

paradox, ist hier aber als ironische Brechung und Distanzierung zu verstehen (vgl. Bohnsack, 2008). Das Motto scheint zu lauten: „Ich kleide mich wie die einstig unbarmherzigen Sultane und erkläre selbst herrschaftliche Ansprüche damit, distanzieren mich aber zugleich von diesen Herrschern, weil mir sie und ihr Erbe nicht ganz geheuer sind.“

Effektives Schwerverbrechen und Lebensgefahr sind Horizonte, die im Dumont-Reisevideo „OnTour“ (36:00: Auf der Straße herrscht Lebensgefahr, 39:00: Prostitution und Kriminalität im Bezirk Aksaray), sowie in ZDF-Reisevideo „Reiselust Istanbul“ (Terroranschläge von Istanbul) aufgeworfen werden. In „Istanbul Total“ werden solche Bezüge nicht hergestellt.

Dass der Eindruck von Gefahr in „Istanbul Total“ auf ironisierende Weise effektiv vermieden werden soll, dokumentiert sich bildlich in der modifizierten Version der Intro-Sequenz „Raab in Gefahr“, die üblicherweise Bestandteil jeder Stefan-Raab-Sendung ist. Wie auf den folgenden Abbildungen erkennbar, wurde in dieser Intro-Sequenz das brennende Wort „Gefahr“ vor schlammigem Hintergrund einfach durch das Wort „Istanbul“ ersetzt, das mit Stefan Raabs Turban ins Bild fliegt. Gleichzeitig wechseln die zuvor harten, rhythmischen Gitarrenriffs zu einer verspielten orientalischen Melodie über.



Abbildung 68: Raab in Gefahr (TV Total)



Abbildung 69: Raab in Istanbul (Istanbul Total)

# CURRICULUM VITAE

MMag. Stefan Hampl

[stefan.hampl@sfu.ac.at](mailto:stefan.hampl@sfu.ac.at)

+43 (676) 580 29 08

27. Oktober 1974



## Berufliche Laufbahn

- |           |   |
|-----------|---|
| seit 2012 | Finanzreferent<br>Institut für Kulturpsychologie und qualitative Sozialforschung (ikus)<br><a href="http://www.ikus.cc">www.ikus.cc</a> |
| seit 2011 | Vizerektor Lehre<br>Sigmund Freud Privatuniversität Wien<br><a href="http://www.sfu.ac.at">www.sfu.ac.at</a>                            |
| seit 2007 | Stellv. Leiter (Studiengangskoordinator) und Lehrbeauftragter<br>Department für Psychologie, Sigmund Freud Privatuniversität Wien       |
| <hr/>     |   |
| 2004–2013 | Doktoratsstudium Psychologie (Dr. phil.)<br>Fakultät für Psychologie, Universität Wien  |
| 2007–2012 | Vorsitzender<br>Institut für Kulturpsychologie und qualitative Sozialforschung (ikus)<br><a href="http://www.ikus.cc">www.ikus.cc</a>   |
| 2010–2011 | Betriebsrat<br>Sigmund Freud Privatuniversität Wien   |
| 2005–2008 | Lehrbeauftragter für Allgemeine Psychologie (Proseminar)<br>Fakultät für Psychologie, Universität Wien                                  |
| 2004–2005 | Studienassistent für qualitative Forschungsmethoden und Kulturpsychologie<br>Fakultät für Psychologie, Universität Wien                 |
| 2004      | Studienabschluss Psychologie (Mag. phil.)<br>Fakultät für Psychologie, Universität Wien   |
| 2003      | Studienabschluss Handelswissenschaften (Mag. rer. soc. oec.)<br>Wirtschaftsuniversität Wien   |

## Mitgliedschaften

- |           |  |
|-----------|--|
| seit 2011 | Mitglied der Österreichischen Privatuniversitätenkonferenz (ÖPUK)<br><a href="http://www.privatuniversitaeten.at/">http://www.privatuniversitaeten.at/</a> |
| seit 2010 | Mitglied des Centrums für qualitative Evaluations- und Sozialforschung (ces)<br><a href="http://www.ces-forschung.de">http://www.ces-forschung.de</a>      |
| seit 2007 | Gründungsmitglied des Instituts für Kulturpsychologie und qualitative Sozialforschung (ikus)<br><a href="http://www.ikus.cc">www.ikus.cc</a>               |

## Tagungsbeiträge

- Dez. 2011 *“Expanding the Concept of Planimetric Composition: The Empirical Value of Color”*  
Sign11 – international conference on signage and wayfinding  
International Institute of Information Design (IIID), Vienna
- Okt. 2010 *“Planimetric Design”*  
SpaceX:  
An Exchange Forum on Information Design for Visually Impaired People  
International Institute of Information Design (IIID), Vienna
- Sept. 2010 *„Rekonstruktive Interpretation im Medium von Bild und Video“*  
Tagung „Neue Perspektiven der dokumentarischen Methode“  
Centrum für qualitative Evaluations- und Sozialforschung (ces), Eisenach
- Juni 2009 *„Videointerpretation und -transkription“*  
Tagung „Videobasierte Methoden der Bildungsforschung - sozial-, kultur- und erziehungswissenschaftliche Nutzungsweisen“  
Stiftung Universität Hildesheim
- März 2009 *“Television programs as cultural documents: Methods of interpreting video and film”*  
Conference “East European versus West European Mentalities: Can We Hope to Understand One Another?”  
Sigmund Freud Privat Universität Wien Paris
- Sept. 2008 *“Film & Video Interpretation with the Documentary Method”*  
7th International Conference of Social Science Methodology  
Universität Federico II, Neapel
- Dez. 2007 *“Broadcasting German and Turkish Cultures: Film & Video Interpretation through the Lens of the Documentary Method”*  
ikus Research Presentation  
St. Kliment Ohridski Universität Sofia, Bulgarien
- Juni 2007 *“The Turkish City in German Entertainment TV – A Reconstructive Video Analysis”*  
International Conference on Modern Cultures and Identities of Old Cities  
Ivan Franko Nationaluniversität Lemberg, Ukraine
- Jan. 2006 *„Die mediale Konstruktion des Türkischen und des Deutschen in kulturübergreifenden Fernsehformaten am Beispiel von ‚Istanbul Total‘“*  
9. Workshop zur qualitativen Bildungs-, Beratungs- und Sozialforschung  
Otto-von-Guericke-Universität Magdeburg, Deutschland

## Vorträge

- Dez. 2011 *„Planimetrieorientierte Bild- und Videointerpretation“*  
Sigmund Freud Privat Universität Wien Paris  
Präsentation im Rahmen der Publikationsmanufaktur des Departments für Psychologie
- Nov. 2010 *“Qualitative Social Research: The Challenges of Picture and Video Interpretation”*  
Slowenisches Institut für Psychotherapie, Ljubljana  
Gastvortrag
- Juni 2009 *„Die Methode der Videointerpretation“*  
Sigmund Freud Privat Universität Wien Paris  
Donnerstagskolloquium

- April 2008 „Kulturpsychologie und Sprache“  
Institut für Germanistik, Ivan Franko Nationaluniversität Lemberg, Ukraine  
Gastvortrag auf Einladung des Lehrstuhlleiters für deutsche Philologie
- Juni 2007 „Kulturpsychologie als Wissenschaft & Berufsperspektive“  
Institut für Psychologie, Universität Graz  
Gastvortrag im Rahmen der Lehrveranstaltung Karriereplanung für PsychologInnen
- April 2007 „Kulturpsychologie und Logotherapie“  
Institut für Logotherapie und Existenzanalyse (ILE) Chur, Schweiz  
Eröffnungsvortrag zur Jahrestagung
- Juni 2006 „Die Methode der Videotranskription“  
Arbeitsbereich Qualitative Bildungsforschung, Freie Universität Berlin  
Gastvortrag bei Univ.-Prof. Dr. Ralf Bohnsack
- Mai 2006 „Materielle Kultur: Zugänge von Kulturpsychologie und qualitativer Sozialforschung“  
Universität für Angewandte Kunst Wien  
Gastvortrag im Rahmen der Lehrveranstaltung Designtheorie
- Okt. 2003 „Totem und Tabu der Führungsforschung“  
Institut für Psychologie der Universität Wien  
Fortbildungsseminar für PsychologInnen

## Ausgewählte Publikationen

- Baltruschat, A., & Hampl, S. (2013). Das Bild im Film - der Film im Bild. Zur Interpretation von Filmen nach der dokumentarischen Methode. In P. Loos, A.-M. Nohl, A. Przyborski, & B. Schäffer (Hrsg.), *Dokumentarische Methode. Grundlagen - Entwicklungen - Anwendungen*. (S. 243–267). Leverkusen, Berlin, Toronto: Verlag Barbara Budrich.
- Hampl, S. (2010). Videos interpretieren und darstellen. Die dokumentarische Methode. In F. Corsten, M. Krug, & C. Moritz (Hrsg.), *Videographie praktizieren. Herangehensweisen, Möglichkeiten und Grenzen* (S. 53–88). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Hametner, K., Hampl, S., Joerchel, A., & Wrbuschek, M. (2009). Introduction to the Special Issue on Cultural Psychology and Qualitative Social Research in Vienna. *Psychology and Society, University of Cambridge (UK)*, 1–7.
- Hampl, S. (2008). *MoviScript – Software zur Videotranskription. (Computerprogramm)*. Abgerufen am 30. April 2013 von [www.moviscript.net](http://www.moviscript.net)
- Hampl, S., Payrhuber, A., Przyborski, A., & Vitouch, P. (2006). *Marktforschung österreichische Klassenlotterie. Eine empirische Untersuchung auf der Basis von Stammkunden, unsteten Kunden, Probierkunden und Nicht-Kunden* (unveröffentl. Projektbericht). Wien.
- Hampl, S. (2003). *Management braucht Intuition. Die Notwendigkeit eines neuen Denkstils in der Führungsforschung – eine neurowissenschaftliche Begründung* (unveröffentl. Diplomarbeit). Wirtschaftsuniversität Wien.

## Sprachkenntnisse

- |                               |   |
|-------------------------------|---|
| fließend                      | Deutsch (Muttersprache), Englisch, Französisch          |
| selbständige Sprachverwendung | Bulgarisch, Italienisch, Spanisch, Russisch, Ukrainisch |