

# **Eine graduelle Handlungssemantik zur Analyse und Generierung multimedialer Gebrauchsdokumente**

Dissertation

Zur Erlangung des akademischen Grades  
Doktor-Ingenieur (Dr.-Ing.)  
vorgelegt im Fachbereich  
Elektrotechnik, Informationstechnik, Medientechnik  
an der  
Bergischen Universität Wuppertal

von

Dipl.-Math. Margit Becher

1. Gutachter: Prof. Dr. Karl-Heinrich Schmidt, Bergische Universität Wuppertal

2. Gutachter: Prof. Dr. John A. Bateman, Universität Bremen

Tag der mündlichen Prüfung: 14.01.2016

Die Dissertation kann wie folgt zitiert werden:

urn:nbn:de:hbz:468-20160421-114938-4

[<http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3A468-20160421-114938-4>]

## Danksagung

Mein ganz besonderer Dank gilt meinem Doktorvater Prof. Dr. Karl-Heinrich Schmidt, Bergische Universität Wuppertal. Ohne seine Begeisterung für das Thema, seine vielfältigen Anregungen und Hinweise, seine Geduld und sein Verständnis in Phasen, in denen die Arbeit ruhen musste, da die Lehre Vorrang hatte, würde es diese Dissertation nicht geben.

Prof. John A. Bateman, Universität Bremen, danke ich sehr für die Übernahme des Korreferats und für das meiner Arbeit entgegengebrachte Interesse.

Meinem Sohn Sebastian Becher (B.A.) danke ich herzlich für die Erstellung der Icons zur Visualisierung der Text-Bild-Relationen (Tab. 4.2).

Danken möchte ich auch dem Westdeutschen Rundfunk für die Produktion des Magazins »Die Sendung mit der Maus«. Die Sachgeschichte, die die Produktion eines Snowboards zeigt, erwies sich als ideales Beispiel zur Exemplifikation der vorgestellten Konzepte und Durchführung der »Experimente«.

Weiterhin bin ich allen zu Dank verpflichtet, die mich in den letzten Jahren in vielfältiger Weise begleitet und unterstützt haben, sei es durch Zuhören, freundliches Nachfragen, Mut machen oder Kaffee in den Vorlesungspausen.

Gehrden, im Januar 2016

Margit Becher



## Zusammenfassung

Die vorliegende Dissertation beschäftigt sich mit der Analyse und Generierung multimedialer Gebrauchsdokumente, insbesondere solchen, die filmische Komponenten enthalten.

Auf Basis des filmischen Bildes werden die paradigmatischen Relationen zwischen zwei Einstellungen und die syntagmatische Struktur eines Films analysiert. Ergänzend dazu werden auch die Beziehungen des Texttracks zum Filmbild betrachtet, um so auch eine multimediale Sicht auf ein filmisches Dokument zu realisieren.

Beispielhaft durchgeführt werden diese Analysen für ein Video, das eine alternante Struktur aufweist: in einem Handlungsstrang wird der Herstellungsprozess eines Snowboards dargestellt, im anderen Handlungsstrang werden Tests des Snowboards in verschiedenen Fertigungszuständen gezeigt.

Damit im Weiteren auch eine stärkere Berücksichtigung der in den Einstellungen ggf. repräsentierten Handlungen möglich ist, wird – ausgehend von einer Handlungssemantik für natürliche Sprachen – eine graduelle Handlungssemantik für cinematographische Dokumente entwickelt.

Um sicherzustellen, dass nur sinnvolle und verständliche Dokumente generiert werden, werden auf Basis der vorherigen Ergebnisse Randbedingungen sowohl zur sequentiellen als auch zur parallelen Anordnung von Inhaltskomponenten hergeleitet. Für die technische Umsetzung werden XML-Repräsentationen zur Darstellung der filmischen Einheiten und der verschiedenen Analysen entwickelt. Generiert werden zwei Formen einer Bildergeschichte sowie Genre-Experimente, die eine Betonung der Kohärenzrelation zwischen den alternanten Segmenten, eine Betonung der kohärent relationierten Segmente, eine Modifikation der Alternanz und einen Wegfall der Kohärenzrelation umsetzen.



## Abstract

This thesis researches the analysis and generation of multimedia functional documents, especially those including filmic components.

Considering the image track the paradigmatic relations between two shots and the syntagmatic structure of a film are analysed. To add a multimedia view of the film the relations between the text track and filmic images are analysed. These analyses are exemplified by a video, which has an alternating dramaturgy: one storyline shows the production of a snowboard, the other one shows tests of the snowboard in its partly finished states.

Furthermore, starting from a semantics of action for natural languages, a gradual semantics of action for cinematographic documents is developed to accord more attention on the actions possibly represented in the shots.

Based on the previous results constraints for both sequential and parallel positioning of content portions are derived in order to ensure that only useful and understandable documents are generated. For technical implementation XML representations of the filmic units and the various analyses are developed. There are generated two types of picture stories and also experimental genres to emphasise the coherence relation between the alternant segments, to emphasise the coherent related segments, to show a modification of the alternation and to show the effect of the non-existence of a coherence relation.



# Inhaltsverzeichnis

<b>Abbildungsverzeichnis .....</b>	<b>V</b>
<b>Tabellenverzeichnis .....</b>	<b>IX</b>
<b>Codeverzeichnis .....</b>	<b>XI</b>
<b>1 Einleitung .....</b>	<b>1</b>
<b>2 Cinematographische Dokumente.....</b>	<b>5</b>
2.1 Einleitung .....	5
2.2 Syntagmatische Organisation eines Films.....	7
2.2.1 Strukturierte cinematographische Dokumente.....	7
2.2.2 Basale Syntagmen .....	11
2.2.3 Narrative Hierarchisierung.....	21
2.2.4 Polyspatiale Alternanz .....	26
2.3 Paradigmatische Organisation eines Films.....	31
2.4 Text-Bild-Relationen .....	39
2.5 Analyse des Videos »Snowboard« .....	44
2.5.1 Vorgehensweise.....	44
2.5.2 Paradigmatische Analyse.....	46
2.5.3 Syntagmatische Analyse.....	60
2.5.4 Analyse der Text-Bild-Relationen .....	68
<b>3 Algebraische Handlungssemantik.....</b>	<b>75</b>
3.1 Einleitung .....	75
3.2 Grundlegende Begriffe und Definitionen.....	76
3.2.1 Klassen, Instanzen und ihre Relationen .....	76
3.2.2 Segmente in Film und Video.....	78
3.3 Thematische Relationen und ihre Verankerung im Filmbild .....	84
3.4 Objekt-Eindeutigkeit segment-thematischer Relationen .....	93
3.5 Relationierung von Subobjekten und Subereignissen.....	94
3.6 Handlungstypen .....	101
3.7 Repräsentation von Accomplishments im Beispielvideo .....	105
3.8 Zur cinematographischen Realisierung von Accomplishments .....	109
<b>4 Experimente mit SMIL und HTML5.....</b>	<b>121</b>
4.1 Einleitung .....	121
4.2 Randbedingungen der Dokumentgenerierung .....	124

4.2.1	Randbedingungen der thematischen Entfaltung.....	125
4.2.2	Randbedingungen der Anordnung.....	128
4.3	XML-Repräsentation der filmischen Einheiten und ihrer Relationen .....	132
4.3.1	XML-Repräsentation der filmischen Einheiten (Einstellungen) .....	133
4.3.2	XML-Repräsentation der syntagmatischen Analyse.....	135
4.3.3	XML-Repräsentation der paradigmatischen Analyse.....	136
4.3.4	XML-Repräsentation der Analyse der Text-Bild-Relationen.....	140
4.4	Verwendete Layoutsprachen.....	143
4.4.1	SMIL .....	143
4.4.2	Videointegration in HTML5.....	146
4.5	Analytische Hilfsmittel .....	148
4.5.1	Berechnung einstellungsübergreifender Ortsgruppen .....	148
4.5.2	Direktes Abspielen der einzelnen Einstellungen .....	151
4.5.3	Erzeugung der Wechseltabelle.....	152
4.5.4	Unterstützung der Analyse der paradigmatischen Relationen.....	154
4.5.5	Auswertung der Text-Bild-Relationen.....	157
4.6	Bildergeschichten .....	159
4.6.1	Einfache Bildergeschichte .....	160
4.6.2	Detaillierte Bildergeschichte.....	164
4.7	Modifikation der Alternanz unter Erhalt der Kohärenzrelation .....	167
4.7.1	Betonung der Kohärenzrelation .....	168
4.7.2	Betonung des kohärent Relationierten .....	171
4.8	Modifikation der Alternanz und der Kohärenzrelation .....	176
4.9	Wegfall der Kohärenzrelation .....	182
4.10	Zusammenfassung .....	188
<b>5</b>	<b>Literatur.....</b>	<b>189</b>
<b>6</b>	<b>Anhang.....</b>	<b>195</b>
6.1	Transkription des Videos »Snowboard«.....	195
6.2	Tabellarische Übersicht der Analyse der Text-Bild-Relationen .....	202
6.3	Dokumentation der Funktionen.....	207
6.3.1	XSLT-Funktionen.....	207
6.3.2	JavaScript-Funktionen .....	207
6.4	Übersicht aller Dateien auf der CD .....	210

# Abbildungsverzeichnis

Abb. 2.1: »Große Syntagmatik des Films« nach Metz .....	5
Abb. 2.2: Drei Sichten auf ein Dokument.....	7
Abb. 2.3: Beispiel für ein nicht-strukturiertes und ein strukturiertes Dokument .....	8
Abb. 2.4: Zuordnung von content portions zu basic logical objects .....	9
Abb. 2.5: Layoutprozess.....	9
Abb. 2.6: Beziehung Syntagma – Segment .....	11
Abb. 2.7: Ausgangssituation mit drei Syntagmen .....	13
Abb. 2.8: Zeitliche Verhältnisse zwischen Einstellungen .....	16
Abb. 2.9: Die acht Basisrelationen nach Randell (RCC8).....	16
Abb. 2.10: Darstellung einer minimalen Szene (ohne zeitliche Überlappung).....	18
Abb. 2.11: Darstellung einer Sequenz, die eine Szene enthält .....	19
Abb. 2.12: Dokumentbaum eines narrativ basalen Dokumentes .....	20
Abb. 2.13: Darstellung der »minimal story« aus Prince in drei Einstellungen .....	23
Abb. 2.14: Aggregation zweier narrativ basaler Syntagmen zu einer narrativen Folge.....	24
Abb. 2.15: Narrative Folge und ihre Darstellung als Wechseltabelle.....	27
Abb. 2.16: Wechselgraph (Übergangsgraph) einer narrativen Folge .....	27
Abb. 2.17: Dimension PROJECTION der »Grande Paradigmatique« .....	33
Abb. 2.18: Dimension TAXIS der »Grande Paradigmatique«.....	34
Abb. 2.19: Dimension »PLANE« der »Grande Paradigmatique«.....	36
Abb. 2.20: Die »Grande Paradigmatique« nach Bateman, Schmidt.....	38
Abb. 2.21: Taxonomie der logisch-semantischen Relationen nach Kong.....	40
Abb. 2.22: SMIL-Video zur Darstellung von paradigmatischen Relationen .....	47
Abb. 2.23: Einstellungen T07 – T08.....	50
Abb. 2.24: Einstellungen T12 – T13.....	52
Abb. 2.25: Paradigmatische Relationen der Einstellungen des diegetischen Ortes »Berg«.....	56
Abb. 2.26: Übersicht der paradigmatischen Relationen des Videos »Snowboard«.....	57
Abb. 2.27: Übersicht der paradigmatischen Relationen (Fortsetzung) .....	58
Abb. 2.28: Spezialisiertes Netzwerk der paradigmatischen Relationen des Videos »Snowboard«.....	59
Abb. 2.29: Syntagmatische Analyse der Einstellungen auf dem Berg.....	60
Abb. 2.30: Zustandswechsel in der Werkstatt darstellt als Petri-Netz.....	62
Abb. 2.31: Syntagmatische Analyse der Einstellungen in der Werkstatt .....	62
Abb. 2.32: Syntagmatische Analyse des Videos »Snowboard« .....	66
Abb. 2.33: Detaillierte Darstellung der syntagmatischen Analyse des Videos »Snowboard«.....	67
Abb. 2.34: Auswertung der Text-Bild-Relationen des Videos »Snowboard« .....	74
Abb. 3.1: Typhierarchie von MPEG-7 (Ausschnitt) dargestellt als UML-Klassendiagramm .....	78
Abb. 3.2: Zusammenhang zwischen VideoSegment und MovingRegion .....	79
Abb. 3.3: Zusammenhang zwischen MovingRegion, Shot und ShotRegion .....	80

Abb. 3.4: Labelling des Segmentes SW_02 = (T09, T10) .....	82
Abb. 3.5: Handlungen in SW_02 = (T09, T10).....	87
Abb. 3.6: Mögliche Fälle für die Konzeptionalisierung der Instanzen von KLEBER_AUFTRAGEN .....	88
Abb. 3.7: Semantisches Netz: AG, PAT-Relationen zur Instanz kleber_auftragen_1.....	90
Abb. 3.8: Semantisches Netz: AG, PAT, INST-Relationen zur Instanz kleber_auftragen_1 .....	91
Abb. 3.9: Subereignis-Abbildbarkeit und Subobjekt-Abbildbarkeit von »ein Glas Wein trinken« .....	95
Abb. 3.10: Beispiele für Raum-Zeit-Diagramme.....	95
Abb. 3.11: Beispiel für das Labelling bei veränderten ShotRegions.....	97
Abb. 3.12: Fehlendes Segment durch »Filmische Raffung«.....	97
Abb. 3.13: Struktur eines telischen Ereignisses .....	101
Abb. 3.14: Schema für Achievements.....	102
Abb. 3.15: (Einfaches) Schema für Activities.....	102
Abb. 3.16: (Einfaches) Schema für Accomplishments .....	102
Abb. 3.17: Detailliertes Schema für Activities und Accomplishments.....	102
Abb. 3.18: Schema des Accomplishments »Bau des Snowboards«.....	106
Abb. 3.19: Schema des Accomplishments »Test 1 des Snowboards« .....	107
Abb. 3.20: Raum-Zeit-Diagramm zu »Bau des Snowboards« .....	107
Abb. 3.21: Raum-Zeit-Diagramm des Videos »Snowboard« .....	108
Abb. 3.22: Handlungen in T09.....	115
Abb. 3.23: Subereignisse eines Accomplishments in Einstellungen.....	116
Abb. 3.24: Handlungen in (T18, T19, T20) .....	118
Abb. 4.1: Schematische Sicht eines Hypervideos .....	122
Abb. 4.2: Content Base (physikalisch) .....	122
Abb. 4.3: Content Base bestehend aus »Content Portions«.....	123
Abb. 4.4: Vorgehensweise zur Entwicklung weiterer multimedialer Genre-Experimente .....	124
Abb. 4.5: Thematische Entfaltung im Video »Snowboard« .....	127
Abb. 4.6: Syntagmatische Analyse der Einstellungen auf dem Berg.....	135
Abb. 4.7: Screenshot: Direktes Abspielen der einzelnen Einstellungen .....	151
Abb. 4.8: Übersicht: Paradigmatische Relationen – Darstellung im PDF-Dokument.....	155
Abb. 4.9: »Transformation« vom Film zur Bildergeschichte .....	159
Abb. 4.10: Screenshot der einfachen Bildergeschichte in HTML .....	160
Abb. 4.11: Screenshot der detaillierten Bildergeschichte in HTML.....	165
Abb. 4.12: Screendesign für das Genre-Experiment »Arbeitsfortschritt« .....	169
Abb. 4.13: Screenshot des Genre-Experimentes »Arbeitsfortschritt«.....	170
Abb. 4.14: Varianten des Screendesigns für das Genre »Produktion und Test« .....	171
Abb. 4.15: Varianten der Kombination von Content Portions .....	172
Abb. 4.16: Screenshot des Genre-Experimentes »Produktion und Test« .....	176
Abb. 4.17: »Drehbuch« zu Modifikation der Alternanz und der Kohärenzrelation.....	178
Abb. 4.18: Wegfall der Kohärenzrelation .....	182
Abb. 4.19: Screendesign für das Genre »Einzelne Orte«.....	183

Abb. 4.20: Screenshot des Genre-Experimentes »Einzelne Orte« ..... 186  
Abb. 4.21: Petri-Netz zum Beispielvideo »Snowboard« ..... 187



## Tabellenverzeichnis

Tab. 2.1: Übersicht: Relationen der Allen'schen Zeitlogik .....	14
Tab. 2.2: Temporale Relationen für syntagmatisch interessante zeitliche Verhältnisse .....	14
Tab. 2.3: Wechseltabelle zum Video »Snowboard« .....	45
Tab. 2.4: Einstellungsgrößen und paradigmatische Relationen der Einstellungen T21 – T31.....	54
Tab. 2.5: Wechsel zwischen Berg und Werkstatt.....	63
Tab. 2.6: Wechsel zwischen Berg und Werkstatt mit Arbeitsstationen.....	64
Tab. 3.1: Systematik der Handlungstypen von Vendler.....	103
Tab. 3.2: Übersicht: Subereignisse von »Bau des Snowboards« .....	105
Tab. 3.3: Übersicht: Subereignisses von »Test des Snowboards« .....	106
Tab. 3.4: Handlungen und Teilhandlungen in T17.....	112
Tab. 4.1: Übersicht: DTDs für die XML-Repräsentation der Einstellungen und Relationen .....	132
Tab. 4.2: Übersicht: Icons zur Visualisierung der Text-Bild-Relationen.....	165
Tab. 4.3: Zeitliche Dauer von Berg- und Werkstatt-Segmenten.....	173
Tab. 4.4: Zweidimensionales Array source im Beispielvideo »Snowboard« .....	183
Tab. 4.5: Sprechertexte in »Tests des Snowboards« .....	187



## Codeverzeichnis

Code 4.1: DTD zur Repräsentation der filmischen Einstellungen: units.dtd .....	133
Code 4.2: Ausschnitt aus snowboard.xml.....	135
Code 4.3: DTD zur Repräsentation der syntagmatischen Analyse: syntagmatic_analysis.dtd.....	135
Code 4.4: Ausschnitt aus syntagmatic_analysis.xml .....	136
Code 4.5: DTD zur Repräsentation der paradigmatischen Relationen: paradigmatic_relation.dtd .....	136
Code 4.6: Ausschnitt aus paradigmatic_relation.xml .....	137
Code 4.7: DTD zur Repräsentation der paradigmatischen Analyse: paradigmatic_analysis.dtd .....	137
Code 4.8: Ausschnitt aus paradigmatic_analysis.xml.....	138
Code 4.9: Ausschnitt aus paraRelation_gen.xml.....	139
Code 4.10: DTD zur Repräsentation der Taxonomie der Text-Bild-Relationen: textImage_relation.dtd ...	140
Code 4.11: Ausschnitt aus textImage_relation.xml .....	140
Code 4.12: DTD zur Repräsentation der Analyse der Text-Bild-Relationen: textImage_analysis.dtd.....	141
Code 4.13: Ausschnitt aus textImage_relation_gen.xml.....	142
Code 4.14: Ausschnitt aus textImage_analysis.xml.....	142
Code 4.15: Integration verschiedener Videoformate in HTML5 .....	146
Code 4.16: Template segment zur Berechnung der Segmente der verschiedenen Orte .....	149
Code 4.17: Segmente für das Video »Snowboard« berechnet mit dem Template segment.....	149
Code 4.18: Aufruf des Templates segment .....	149
Code 4.19: Template segment_loc_detail zur Berechnung der Segmente an verschiedenen Orten .....	150
Code 4.20: Segmente für das Video »Snowboard« berechnet mit dem Template segment_loc_detail .....	150
Code 4.21: Ausschnitt aus show_units.xml.....	152
Code 4.22: Ausschnitt aus transition_table.xml.....	154
Code 4.23: Ausschnitt aus paraRelation_show.xml .....	156
Code 4.24: Ausschnitt aus textImage_report.xml.....	158
Code 4.25: Ausschnitt aus picStory_simple_html.xml.....	162
Code 4.26: Benanntes Template content zur Generierung des Inhaltsverzeichnisses in HMTL .....	162
Code 4.27: Benanntes Template content zur Generierung des Inhaltsverzeichnisses in SMIL.....	164
Code 4.28: for-each-group-Schleife zur Generierung der Regionen des Inhaltsverzeichnisses .....	164
Code 4.29: Erzeugter HTML-Code für das Inhaltsverzeichnis der detaillierten Bildergeschichte.....	166
Code 4.30: Definition der Hilfsvariablen span .....	166
Code 4.31: Erzeugung des Inhaltsverzeichnisses für die detaillierte Bildergeschichte.....	167
Code 4.32: Ausschnitt aus work_progress.xml .....	170
Code 4.33: Ausschnitt aus video_synchron.xml.....	175
Code 4.34: DTD der Konfigurationsdatei zu »Modifikation der Alternanz und der Kohärenzrelation« ....	179
Code 4.35 Konfigurationsdatei zu »Modifikation der Alternanz und der Kohärenzrelation« .....	179
Code 4.36: Benanntes Template itinerar und rekursives Durchlaufen der syntagmatischen Analyse.....	180
Code 4.37: Ausschnitt aus change_alternation_smil.xml.....	182

Code 4.38: Ausschnitt aus separate_location_html.xml .....	184
Code 4.39: Ausschnitt aus separate_location_smil.xml .....	186

# 1 Einleitung

In vielen Branchen, insbesondere dem Verlagswesen und der Technischen Dokumentation, müssen – aus einer Quelle – Dokumente für verschiedene Zielgruppen und Anwendungszwecke und diese auch für verschiedene Medien produziert werden. Um diese Anforderungen kostengünstig und effizient zu erfüllen, müssen sämtliche Publikationsprozesse automatisiert werden. Dies ist für Dokumente, die mit einer Standard-Textverarbeitungssoftware erstellt wurden, oft nur mit großen Schwierigkeiten möglich. Die Ursachen hierfür liegen u. a. in der fehlenden Trennung von Inhalt und Formatierung und der oft fehlenden Möglichkeit, Teildokumente mit Metadaten so zu versehen, dass z. B. eine gezielte Filterung möglich wird.

In den letzten Jahren hat sich in diesen Branchen folgende Vorgehensweise bewährt und durchgesetzt: Für den jeweils zu erstellenden Dokumenttyp wird eine logische Struktur vorgegeben. Hierzu gibt es häufig Standardstrukturen, die fest vorgegebene Komponenten enthalten: Beispielsweise besteht ein Buch aus Kapiteln, jedes Kapitel hat eine Überschrift und mindestens einen Absatz, eine Handlungsanweisung besteht aus Handlungsschritten. Solche Standardstrukturen werden oft generisch definiert durch sogenannte Syntagmen für nicht-triviale hierarchische Teilstrukturen. Inhaltsstücke werden dann entsprechend dieser Struktur erfasst, wobei jederzeit eine Strukturprüfung des Dokuments (seiner Validität) möglich ist. Da so erstellte Dokumente noch keine Layoutinformationen enthalten, wird in diesem Kontext auch von einer Trennung zwischen Struktur, Inhalt und Layout gesprochen.

Zur Erzeugung der layoutierten Dokumente müssen Regeln definiert werden, wie die logische Struktur auf ein Layout abgebildet wird. Beispielsweise folgt – unabhängig vom Medientyp – unterhalb einer Kapitelüberschrift stets ein Absatz, und Absätze, die Handlungsanweisungen enthalten, müssen auch in ihrer logisch vorgegebenen Reihenfolge präsentiert werden; printspezifisch beginnt ein neues Kapitel oft auf einer eigenen Seite. Die schon genannten Syntagmen sind Invarianten dieser Abbildung: Alles was syntagmatisch miteinander verbunden ist, bleibt auch im Default im Layout zusammen und bildet im Default auch eine inhaltliche Einheit.

Technisch umsetzbar wurde diese Vorgehensweise durch die Entwicklung sogenannter Auszeichnungssprachen. Bereits 1986 wurde die Sprache SGML (Standard Generalized Markup Language) definiert, welche jedoch wegen ihrer Komplexität heute nur selten verwendet wird. Der Durchbruch dieser Vorgehensweise erfolgte ab 1998 mit der Entwicklung der Sprache XML (extensible Markup Language). Aus den XML-Dokumenten können die gewünschten Zieldokumente z. B. mit Hilfe von sogenannten Stylesheets der XML-Sprache XSLT (extensible Stylesheet Language Transformations) generiert werden.

Es stellt sich nun die Forschungsfrage, ob und wie sich dieser Prozess auf multimediale Dokumente, speziell solche, die filmische Komponenten enthalten, übertragen lässt.

Für die Darstellung multimedialer Dokumente wurden verschiedene Markupssprachen definiert. Insbesondere zwei sind für die hier betrachtete Problemstellung besonders geeignet: Eine ist die Synchronized Multimedia Integration Language (SMIL), eine XML-Anwendung zur Spezifikation von räumlichen und zeitlichen Beziehungen und Interaktionen zwischen Medienobjekten. Die zweite ist HTML in der Version 5, denn mit dieser Version ist es erstmals möglich, Video- und Audiodaten nativ so in ein HTML-Dokument einzubetten, dass auf Teildokumente zugegriffen werden kann.

Die Bereitstellung dieser Sprachen bietet viele technische Möglichkeiten, bei denen man sich als Autor zwei Fragen gegenübergestellt sieht:

1. Welche syntagmatischen Strukturen lassen sich in cinematographischen Dokumenten wiederfinden, speziell solchen, die inhaltlich eine klare Handlungsstruktur repräsentieren?
2. Welche Regeln müssen für die Abbildung der logischen Struktur eines cinematographischen Dokumentes auf die Layoutstruktur eines multimedialen Dokumentes definiert werden, so dass sinnvolle und verständliche Dokumente generiert werden können?

Diesen Fragen stellt sich diese Arbeit theoretisch allgemein und empirisch speziell. Die empirische Anwendung ist dabei eingegrenzt auf Gebrauchsfilme, speziell solche, die die Herstellung eines technischen Produktes zeigen. Bei diesen ist – im Unterschied zu einem künstlerischen Film – keine »besondere Interpretation« des Filmbildes notwendig.

### **Stand der Forschung**

Grundsätzlich kann man einen Film wie einen Text ansehen, als Zeichensystem, das – analog zur Sprache – über eine syntagmatische und eine paradigmatische Achse verfügt [Bateman, Schmidt 2011, S. 83f], [Monaco 2009, S. 174]. Eine Klassifikation möglicher syntagmatischer Segmente im Film wurde bereits Mitte der 1960er Jahre von Christian Metz mit seiner Veröffentlichung der »*Grande Syntagmatique du film narratif*« vorgestellt. Der Ansatz von Metz wurde von etlichen Autoren kritisiert. Eine Reformulierung der chronologischen Syntagmen, die auch den Anforderungen neuerer grammatikorientierter Dokumentbeschreibungssprachen genügt, wurde von [Schmidt, Strauch 2002], [Schmidt 2004], [Schmidt 2008] geleistet.

Da nach [Bateman 2007] in der Metz'schen Syntagmatik keine hinreichende Trennung der syntagmatischen und paradigmatischen Achsen der semiotischen Organisation erfolgt, entwickelte er einen Ansatz zur Klassifikation der verschiedenen Relationstypen zwischen filmischen Einstellungen in Form eines systemischen Netzwerks. Dieses Netzwerk bezeichnet er – auch als Reminiszenz an Metz – als »*Grande Paradigmatique*«.

Ein weiterer Mangel der bisherigen Syntagmatik besteht darin, dass sie zur Klassifikation lediglich Raum-Zeit-Beziehungen nutzt und sich nicht hinreichend an den in Einstellungen ggf. repräsentierten Handlungen orientiert. Dagegen fokussiert ein Zuschauer – zur Rezeption eines Films – seine Aufmerksamkeit sehr stark auf die gezeigten Handlungen, so dass eine sinnvolle und verständliche Layoutierung dem Rechnung tragen muss.

Der Fortschritt, der in den letzten Jahrzehnten bei der Verarbeitung von Texten im Publikationswesen erzielt wurde, betrifft auch die Klassifikation von Inhalten eines (Volltext-)Dokumentes. So sind speziell zur Beschreibung der Semantik von beschriebenen Handlungen in der Linguistik mehrere Ansätze bereitgestellt worden. Diese werden hier zur Beantwortung der beiden obigen Fragen auf multimediale Dokumente übertragen.

### **Aufbau der Arbeit**

Im Anschluss an diese Einleitung wird in **Kapitel 2** der Begriff des strukturierten cinematographischen Dokumentes definiert. Nach der Vorstellung der Redefinition der narrativ basalen Syntagmen und der polyspatialen Alternanz folgt eine Einführung in das systemische Netzwerk der »*Grande Paradigmatique*« nach Bateman.

Um auch eine multimediale Sicht auf ein filmisches Dokument zu realisieren, wird die Beziehung des Texttracks zum Filmbild betrachtet. Da kaum Forschungsergebnisse zu Relationen zwischen Text und Filmbild vorliegen, wird auf Ergebnisse zu Text-Bild-Relationen (Bild im Sinne von statischem Bild) zurückgegriffen und eine geeignete Taxonomie vorgestellt.

Für einen Film wird eine paradigmatische und eine syntagmatische Analyse sowie eine Klassifikation der vorliegenden Text-Bild-Relationen durchgeführt, um so die gesamte Vorgehensweise auch empirisch zu überprüfen.

Ausgewählt wurde dazu ein Video, das in einem Handlungsstrang den Herstellungsprozess eines Snowboards zeigt. Ein zweiter Handlungsstrang im Video demonstriert, wie ein Akteur das Snowboard in verschiedenen (teilweise fertiggestellten) Zuständen testet. Zwischen den beiden Handlungssträngen wird regelmäßig hin und her gewechselt, so dass eine typische alternante Struktur vorliegt, die sich so in Volltexten i. Allg. nicht findet.

Zur Durchführung der empirischen Arbeiten wird in **Kapitel 3** – ausgehend von einer Handlungssemantik für natürliche Sprachen – eine graduelle Handlungssemantik für cinematographische Dokumente entwickelt. Gradualität erlaubt die Verteilung einer filmischen Handlungsrepräsentation auf mehrere Einstellungen, um diese einem inhaltlichen Zusammenhang zu unterwerfen. Dies wird hier behandelt für sogenannte *Accomplishments*, die (im Unterschied zu sogenannten *Activities*) eine Einstellungsmenge einer erkennbaren Ordnung unterwerfen können.

In **Kapitel 4** steht die Generierung neuer multimedialer Dokumente im Focus. Dazu werden als Erstes XML-Repräsentationen der filmischen Einheiten und der verschiedenen in Kapitel 2 vorgestellten Relationstypen entwickelt. Es folgt eine kurze Einführung in die Sprache SMIL und die Darstellung der Möglichkeiten und Grenzen der Videointegration in HTML5. Für die Generierung der Dokumente liegen Inhaltsstücke vom Typ Video, Bild und Text vor. Unter Nutzung der Ergebnisse von Kapitel 2 und 3 werden Randbedingungen sowohl für die sequentielle als auch die parallele Anordnung der Inhaltsstücke in einem Dokument entwickelt. Dabei stellt sich auch hier – für ein relativ einfaches Dokument mit kanonischer Lesart – heraus, dass mit geringen Modifikationen der Inhaltsstücke und ihrer Anordnung eine gänzlich andere Wirkung erzielt werden kann.

## 2 Cinematographische Dokumente

### 2.1 Einleitung

Seit Mitte der 1960er Jahre veröffentlichte der Filmsemiotiker Christian Metz Untersuchungen zum Problem der Grammatik des Spielfilms [Metz 1964], [Metz 1965], [Metz 1966]. Hierbei interessierte er sich hauptsächlich für »*narrative Elemente – Syntagmen – [...], die innerhalb einer Aufnahme wie auch zwischen ihnen bestehen können [...]*« [Monaco 2009, S. 237]. Nach [Bateman 2007, S. 20] war das Ziel von Metz eine systematische Beschreibung der Montageformen filmischer Segmente:

*»Its main aim was to provide a generic classification that sets out within a single systematic description all the possible ways in which filmic segments can be meaningfully combined.«*

Metz hat »*in der filmischen Gliederung mehrere sequenzbildende Prinzipien erkannt, die sich durch unterschiedliche Zeitverwendung definieren*« [Hickethier 2007, S. 142]. Sequenz wird nach [Hickethier 2007, S. 141] als »*Einheit im Geschehen verstanden, die sich deutlich von anderen abhebt (Strukturpausen) und sich durch einen erkennbaren Handlungszusammenhang bestimmen lässt*.« Anhand der Betrachtung des Films »Adieu Philippine« hat Metz eine Typologie der Sequenzformen entwickelt. Das Ergebnis seiner Arbeit bezeichnete er als »*Große Syntagmatik des Films*« (s. Abb. 2.1, nach [Metz 1972, S. 198]).

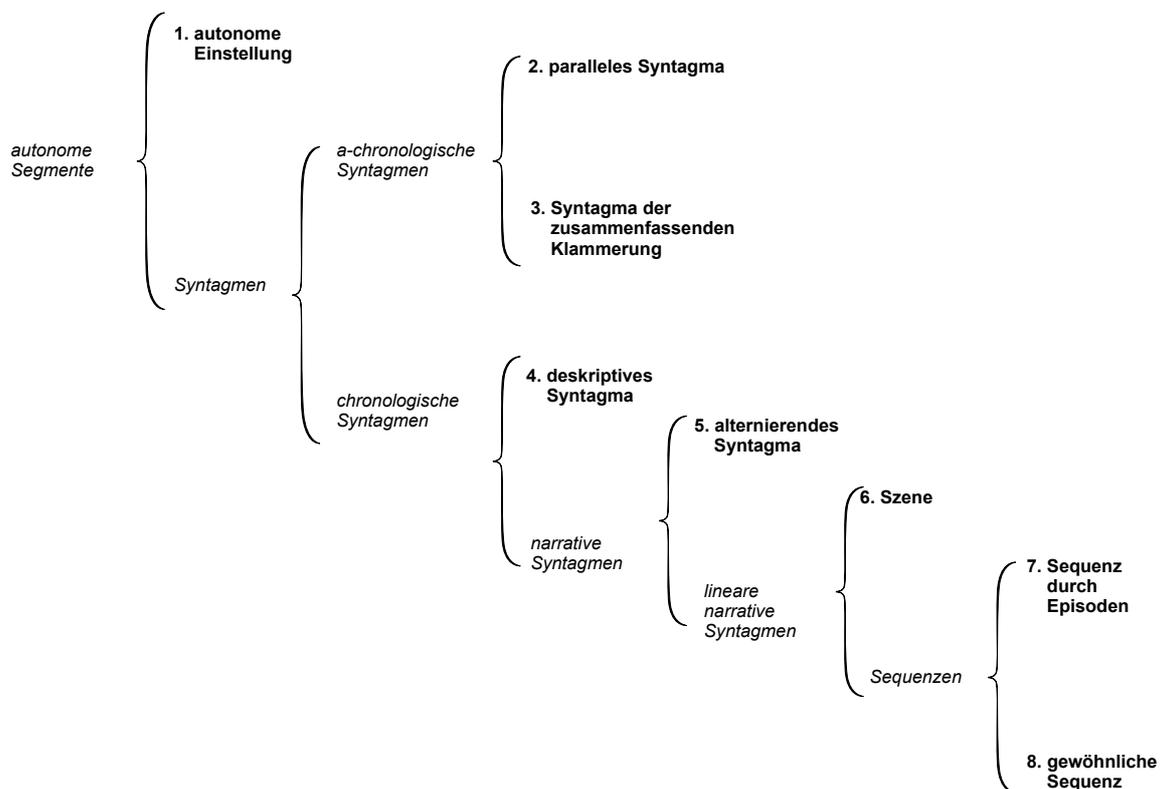


Abb. 2.1: »Große Syntagmatik des Films« nach Metz

Nach [Schmidt, Strauch 2002, S. 66] ist die Syntagmatik »*keine formale Grammatik, enthält aber eine Taxonomie, die auf den ersten Blick recht ansprechend ist.*« Etliche Autoren, z. B. [Möller-Naß 1988], [Colin 1995], [Monaco 2009], kritisierten den Ansatz von Metz und zeigten dessen Begrenztheit auf. Eine Diskussion der drei häufigsten Kritikpunkte lässt sich in [Bateman, Schmidt 2011, S. 117 – 120] nachlesen.

Eine Weiterentwicklung der Syntagmatik, insbesondere eine Reformulierung der chronologischen Syntagmen, wurde in [Schmidt, Strauch 2002], [Schmidt 2004], [Schmidt 2008], [Bateman, Schmidt 2011] geleistet, weil »*sie für strukturelle Untersuchungen nach wie vor einen guten Einstieg bietet und für eine grammatische Sicht für die elektronische Verarbeitung von Dokumenten wesentlich ist*« [Schmidt, Strauch 2002, S. 66]. Bei dieser Revision wurden insbesondere die folgenden Aspekte berücksichtigt:

- Innerhalb des Films vorhandene audiovisuelle Anhaltspunkte (*clues*), d. h. räumliche oder objektbasierte Anker, werden explizit betrachtet, da diese Anker von den Beobachtern verwendet werden, um aufeinanderfolgende Einstellungen zu verbinden [Bateman 2007, S. 28].
- Die Formalisierung genügt auch den Anforderungen neuerer grammatikorientierter Dokumentbeschreibungssprachen [Schmidt, Strauch 2002, S. 65].

Im ersten Teil dieses Kapitels wird kurz der dokumententheoretische Hintergrund erläutert. Zur Anwendung auf filmische Dokumente wird der Begriff des strukturierten cinematographischen Dokumentes eingeführt. Anschließend wird die Redefinition der Syntagmen »Planszene«, »Szene«, »Sequenz«, der »narrativen Folge« und ihres Spezialfalls der »polyspatialen Alternanz« vorgestellt.

Im Anschluss daran wird das systemische Netzwerk der »*Grande Paradigmatique*« nach [Bateman, Schmidt 2011] erläutert, da zur Produktion und auch zur Analyse eines Films sowohl die Betrachtung der syntagmatischen als auch der paradigmatischen Achse der semiotischen Organisation notwendig ist. Das Netzwerk bietet eine Klassifikationsmöglichkeit für die vielfältigen, simultan vorliegenden Relationen zwischen filmischen Einstellungen, die Ausgangspunkt der paradigmatischen Beschreibung sind.

Eine paradigmatische und syntagmatische Analyse eines Filmes basiert nur auf der Betrachtung des filmischen Bildes. Da aber auch die Relationen zwischen Bild und Text zum Verständnis des Films beitragen, wird eine Taxonomie für Text-Bild-Relationen, die ursprünglich für Text und statisches Bild entwickelt wurde, erläutert.

Für ein Beispielvideo wird eine paradigmatische, syntagmatische sowie eine Analyse der Text-Bild-Relationen durchgeführt.

## 2.2 Syntagmatische Organisation eines Films

### 2.2.1 Strukturierte cinematographische Dokumente

Bei der Modellierung strukturierter Dokumente [ISO/IEC 8613:1994] gibt es drei verschiedene, komplementäre Sichten auf ein Dokument (s. Abb. 2.2, nach [Bateman, Schmidt 2011, S. 52]):

- die logische Sicht (*logical view*),
- die Layoutsicht (*layout view*),
- die Inhaltssicht (*content view*).

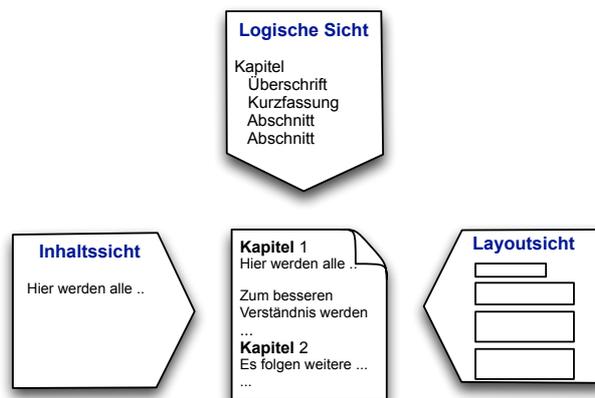


Abb. 2.2: Drei Sichten auf ein Dokument

Die logische Sicht gibt die Struktur eines Dokumentes wieder. Struktur ist die Aufteilung und die Abfolge der Inhaltsstücke (*content portions*). Beispielsweise kann ein Buch Kapitel enthalten, die jeweils aus einer Überschrift, einer Kurzfassung sowie einzelnen Abschnitten bestehen können.

Die Layoutsicht gibt die Repräsentation des Dokumentes zu einem bestimmten Zeitpunkt auf einem Ausgabemedium wieder. Hier wird also festgelegt, wie das Dokument als Printdokument auf Papier, als Webseite oder als Bewegtbild auf dem Bildschirm für einen Beobachter erscheint.<sup>1</sup>

Inhalt ist die zu vermittelnde Information, das, was den Leser bzw. Beobachter interessiert. Oft wird der Inhalt in Textform bereitgestellt. Ergänzt wird dieser durch Bilder und immer häufiger auch durch multimediale Elemente, wie Töne oder Videos, die in Binärform vorliegen.

Sowohl die Layoutsicht als auch die logische Sicht modellieren ein **Dokument als Baum**. Gemäß [ISO/IEC 8613:1994] existieren in einem Dokumentbaum drei verschiedene Typen von Objekten:

- (1) der Wurzelknoten eines Dokuments: der *document (logical oder layout) root*;
- (2) ein oder mehrere Endknoten: die *basic (logical oder layout) objects*;

<sup>1</sup> Das eigentliche Dokument wird durch einen Rendering-Prozess erzeugt.

- (3) ein oder mehrere Knoten zwischen Wurzelknoten und Endknoten: die *composite (logical oder layout) objects*.

Der Wurzelknoten ist das oberste Objekt der Hierarchie. Er ist ein *composite object*, seine unmittelbaren Nachfolger sind weitere *composite objects* oder *basic objects*, die in beliebiger Anzahl und Reihenfolge vorkommen können.

Nach [Bateman, Schmidt 2011, S. 150] wird definiert:

**Definition 2.1:** Ein Dokument heißt **strukturiert**, wenn der Wurzelknoten der logischen Struktur mindestens zwei Nachfahren (*descendants*) besitzt. Die Nachfahren müssen nicht notwendigerweise direkte Nachfahren sein.

Das in Abb. 2.3 links gezeigte Dokument ist nicht-strukturiert, da nur ein *basic logical object* Kind des *document logical root* ist. Der rechte Teil von Abb. 2.3 zeigt ein strukturiertes Dokument.

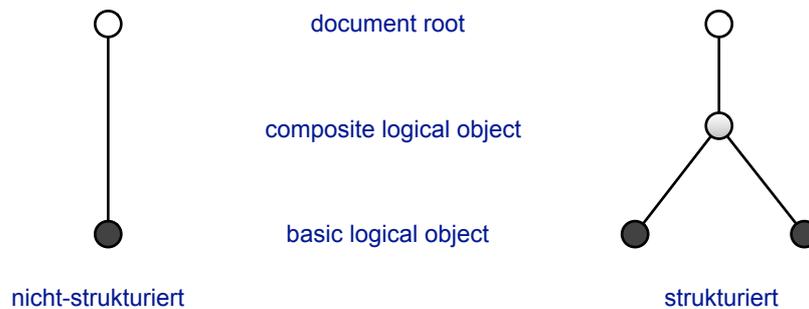


Abb. 2.3: Beispiel für ein nicht-strukturiertes und ein strukturiertes Dokument

Es ist sinnvoll, im Folgenden nur strukturierte Dokumente zu betrachten, also Dokumente, deren logische Struktur eine gewisse Mindest-Komplexität aufweist. Eine entsprechende Mindest-Komplexitätsbedingung findet man auch in der Textlinguistik. So gilt hier als Kernbedeutung der alltagssprachlichen(!) Verwendung des Wortes »Text«<sup>1</sup>: »"Text" ist eine (schriftlich) fixierte sprachliche Einheit, die in der Regel mehr als einen Satz umfasst« [Brinker 2010, S. 12]. Für den Film gilt dies historisch und für viele technische und künstlerische Anwendungen nur mit Einschränkungen. So stellte [Ast 2002, S. 11f] fest: »Die Hälfte der bis 1906 gedrehten, noch existierenden Filme besteht nur aus einer Szene mit einer Einstellung. Ab 1903 ging ein deutlicher Trend bei der Filmproduktion zu multiple scene films.«

In der »klassischen« Anwendung strukturierter Dokumente, der Textverarbeitung, verweisen die *basic logical objects* auf »Text«. Zur Filmanalyse wird ein Filmdokument so zerlegt, dass alle in ihm repräsentierten Zusammenhänge auch logisch sichtbar werden. Dazu werden alle Einstellungen eines filmischen Dokuments zunächst den *basic logical objects* als »Inhaltsstück« (*content portion*) zugeordnet (s. Abb. 2.4).

<sup>1</sup> Dies ist nicht das einzige Textualitätskriterium, s. 2.2.3.

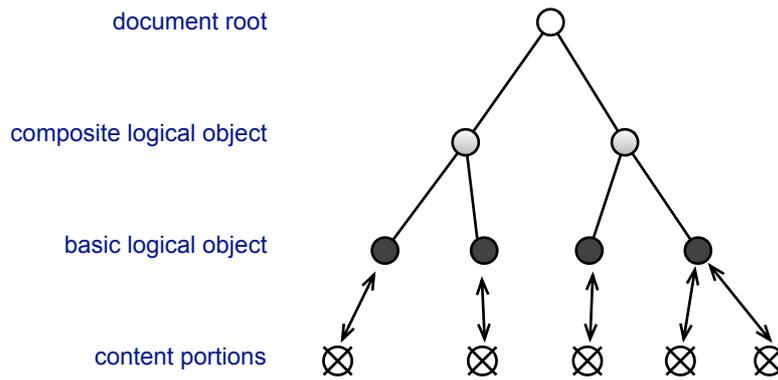


Abb. 2.4: Zuordnung von content portions zu basic logical objects

Technisch gesehen ist eine Einstellung ein Datensatz, »der einer vorgegebenen Menge von audiovisuellen Normen von Bewegbilddaten genügt« [Schmidt 2004, S. 259]. Für die filmanalytischen Anwendungen müssen die Einstellungen nach [Bateman, Schmidt 2011, S. 148, eig. Übersetzung] zudem folgende Bedingung erfüllen:

*»Jede Einstellung kann als verlässliche Raum-Zeit-Messung real wahrnehmbarer Szenen, die von einer Kamera aufgezeichnet werden, aufgefasst werden.«<sup>1</sup>*

Der Begriff »real« in obiger Bedingung umfasst sowohl reale Szenen vor der Kamera (*pro-filmic*) als auch künstliche Welten und Animationen [Bateman, Schmidt 2011, S. 145]. Diese erweiterte Auffassung von »real« ist sinnvoll, da heute eine Vielzahl von Filmen nur mit Hilfe einer digitalen (Nach-)Bearbeitung am Computer geschaffen werden.

**Der Layoutprozess**

Für die cinematographische Analyse sind die logische Sicht und die Layoutsicht und insbesondere auch ihre Beziehung zueinander von Interesse. [Bateman, Schmidt 2011, S. 197f]:

*»It is then in the relationship between the logical perspective and the layout perspective that we situate many of the basic mechanisms of filmic interpretation.«*

Jeder **Layoutprozess** erzeugt aus einem logischen Baum einen Layoutbaum (s. Abb. 2.5).

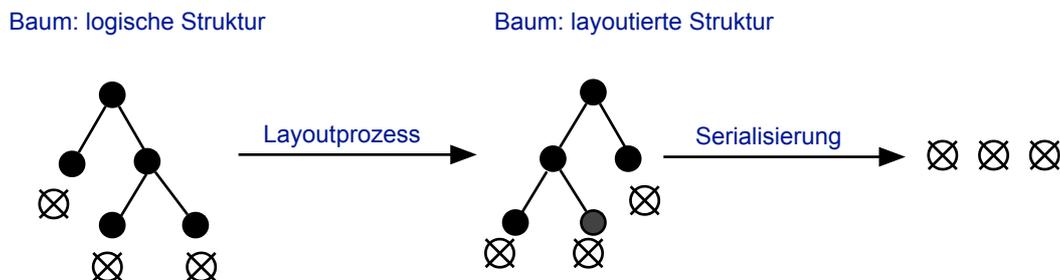


Abb. 2.5: Layoutprozess

<sup>1</sup> Im Original: »[...] the audiovisual content portions of film can be considered to be reliable space-time measurements of perceptually real pro-filmic events.« Als *pro-filmic* wird die Szene, die die Kamera aufzeichnet, bezeichnet.

Betrachtet werden kann ein Dokument – unabhängig vom Ausgabemedium – oft erst nach einer Serialisierung.<sup>1</sup> Der Layoutprozess legt dazu insbesondere fest, welche Inhaltsstücke auf einem Ausgabemedium benachbart sind. Bezogen auf filmische Dokumente legt der Layoutprozess fest, welche Einstellungen zeitlich aufeinanderfolgen.<sup>2</sup> Somit legt der Layoutprozess den »wirklichen Film« fest, den ein Zuschauer betrachten kann, wenn er aufgeführt wird.

Ein Layoutprozess kann mehrere Ausgabeströme vorsehen [Schmidt 2004, S. 258]. Bei einem Printdokument ist dies z. B. eine lebende Kopfzeile mit der Kapitelüberschrift, die – zusätzlich zum eigentlichen Inhalt – auf jeder Seite erscheint. In den filmanalytischen Anwendungen wird jedoch in dieser Arbeit davon ausgegangen, dass der Layoutprozess lediglich einen Ausgabe-  
strom erzeugt, so dass zu jedem Zeitpunkt nur jeweils maximal ein Filmbild betrachtet werden kann.

Es lässt sich nun definieren [Bateman, Schmidt 2011, S. 58]:

**Definition 2.2:** Ein Dokument  $D$  heißt **cinematographisch**<sup>3</sup>, wenn

- (1) jedem *basic logical object* mindestens eine audiovisuelle *content portion* zugeordnet werden kann,
- (2) ein Layoutprozess  $L$  existiert, der für mindestens eine audiovisuelle *content portion* eines jeden *basic logical object* des Dokumentes ein *basic layout object* generiert,
- (3) der Layoutprozess  $L$  die Menge aller layoutierten Einstellungen des Dokumentes einer strengen (räumlichen und/oder zeitlichen) Ordnung unterwirft.

Umgekehrt kann man auch sagen, dass jedes Filmsegment das Ergebnis eines Layoutprozesses ist, so dass für die layoutierten Teile eines cinematographischen Dokumentes definiert werden kann [Bateman, Schmidt 2011, S. 199]:

**Definition 2.3:** Ein **Segment** in einem cinematographischen Dokument ist eine nichtleere Teilfolge in der layoutierten zeitlichen Folge aller Einstellungen für einen gegebenen Layoutprozess.

### **Bemerkung**

Mit »Segment« bezeichnete Metz grundsätzlich jeden Analyseabschnitt, welcher eine zu klassifizierende Einheit darstellt. Ein Segment ist »also ein Abschnitt, dessen spezifische Klassifikation noch nicht erfolgt ist« [Schmidt, Strauch 2002, S. 67]. Dies steht nicht im Widerspruch zu Definition 2.3, da diese keine Klassifikationsbedingungen an die Einstellungsfolge stellt.

---

<sup>1</sup> Zur Betrachtung ist i. Allg. ein geeignetes Programm, z. B. Webbrowser oder Videoplayer, notwendig.

<sup>2</sup> d. h. aus den  $n!$  möglichen Anordnungen der Einstellungen wird eine ausgesucht.

<sup>3</sup> Hier wird bewusst der allgemeinere Begriff »cinematographisches Dokument« statt »Film« gewählt, was [Schmidt 2008, S. 220] wie folgt begründet: »Bei geeigneter Wahl der Zulässigkeitsbedingungen für cinematographische Inhaltsstücke kann auch ein Videoprotokoll medizinischer Operationen ein cinematographisches Dokument sein, das alltagssprachlich nicht gleich als "Film" angesehen werden würde.«

Ein hinreichend komplexer Teilbaum der generischen logischen Struktur wird nach [Schmidt 2004, S. 261] als Syntagma klassifiziert. Genauer gilt (s. auch Abb. 2.6):

**Definition 2.4:** Ein **Syntagma** klassifiziert Teilbäume der logischen Struktur eines Dokumentes, die mindestens ein *composite logical object* enthalten und durch einen Layoutprozess auf mindestens ein Segment abgebildet werden.

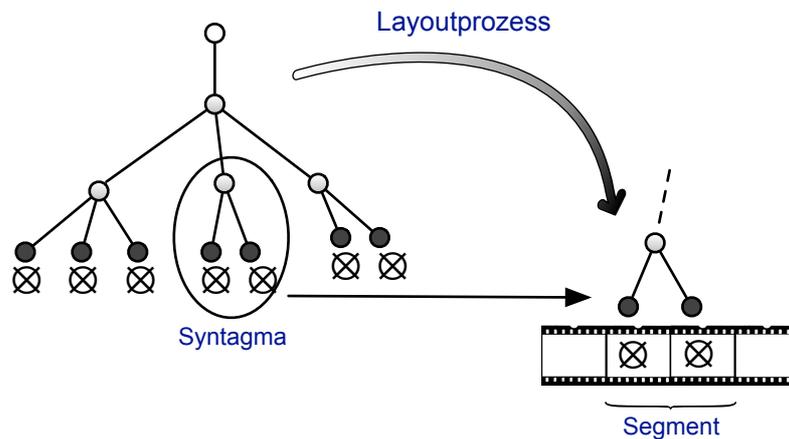


Abb. 2.6: Beziehung Syntagma – Segment

Unter dem Gesichtspunkt der Transformation eines einzelnen Dokumentes in verschiedene Layouts ist ein Syntagma im Default also eine Invariante der Transformation der logischen Struktur des Dokumentes.

Nach [Schmidt 2004, S. 261] kann man abkürzend von den »Segmenten eines Syntagmas« sprechen. Ebenso kann man abkürzend von den »Einstellungen eines Syntagmas« sprechen, wenn diese Einstellungen zu einem der Segmente eines Syntagmas gehören.

## 2.2.2 Basale Syntagmen

[Bordwell 1985, S. 29ff] weist darauf hin, dass der Zuschauer<sup>1</sup> eines Films nicht passiv ist, sondern eine aktive Rolle ausfüllt, wie der Leser eines Textes, der diesen aktiv rezipieren muss. [Hickethier 2007, S. 106] formuliert dies wie folgt: »Nicht der Film vermittelt Bedeutung, sondern der Zuschauer erkennt aufgrund bestimmter Bedingungen in ihm Bedeutungen.« Für diesen Kontext bedeutet dies, dass die Interpretation der audiovisuellen Messungen immer von der Beobachtermenge und deren Weltwissen abhängig ist. Aus diesem Grund werden die entsprechenden Definitionen um den Zusatz »für alle Beobachter einer Beobachtermenge« ergänzt [Bateman, Schmidt 2011, S. 149]. Nicht vorausgesetzt wird – wie schon bei [Bordwell 1985, S. 30] – das Vorhandensein einer idealen Beobachtermenge.

<sup>1</sup> Bordwell führt für den Zuschauer die Bezeichnung »spectator« oder »viewer« ein.

Für die weiteren Überlegungen gelten folgende **Bezeichnungen**:

$D$  sei ein cinematographisches Dokument bestehend aus Einstellungen  $T_i, 1 \leq i \leq n$ ,

$M_D = \{T_1, \dots, T_n\}$  sei die Menge der Einstellungen des Dokumentes  $D$ , für einen Teilbaum  $U$  der logischen Struktur sei  $M_U$  die zugehörige Menge der Einstellungen aus  $U$ ;

ferner seien  $S_j, 1 \leq j \leq k$ , Segmente von  $D$ ,  $Seg(D)$  die Menge aller Segmente von  $D$ .

Zudem gilt folgende **Voraussetzung**:

- (V1) Jedem Segment mit wenigstens einer diegetischen Einstellung kann eine diegetische Raumzeit *DieSpaceTime* zugeordnet werden. Diese ist ein »universeller Bildraum für Beobachtermodelle der Urbilder filmischer Messungen« [Schmidt 2008, S. 221].

Die diegetische Raumzeit ergibt sich als direktes Produkt eines diegetischen Raumes *DieSpace* und einer diegetischen Zeit *DieTime*:  $DieSpaceTime = DieSpace \times DieTime$ . Der diegetische Raum ist nach [Schmidt 2008, S. 221] eine Teilmenge eines (für die Anschauung euklidischen, aber im Folgenden nur als sogenannter Messraum festgelegten)  $\mathbb{R}^3$ . Die diegetische Zeit jeder Einstellung lässt sich in der Menge aller abgeschlossenen, nicht entarteten Intervalle  $I(M)$  eines geeigneten zeitlichen Messraums  $M$  repräsentieren (vgl. [Schmidt 2004, S. 262]).<sup>1</sup> Formal kann nach [Schmidt 2008, S. 221] die Zuordnung von Segmenten zum diegetischen Raum bzw. zur diegetischen Zeit durch zwei (ggf. partielle) Abbildungen *dieS* und *dieT* definiert werden:

$$dieS: B \times Seg(D) \rightarrow DieSpace$$

$$(B, S) \mapsto dieS(B, S)$$

$$dieT: B \times Seg(D) \rightarrow DieTime$$

$$(B, S) \mapsto dieT(B, S)$$

Die Zuordnung der diegetischen Raumzeit ergibt sich dann aus der Kombination dieser Abbildungen:

$$dieST: B \times Seg(D) \rightarrow DieSpace \times DieTime$$

$$(B, S) \mapsto dieS(B, S) \times dieT(B, S)$$

Da Einstellungen als einpunktige Segmente betrachtet werden können, gelten obige Abbildungen auch für (einzelne) Einstellungen.

Im Folgenden wird in Anlehnung an [Bateman, Schmidt 2011] die diegetische Raumzeit einer Einstellung  $T_i$  mit  $SpaceTime(T_i)$  bezeichnet.

<sup>1</sup> Zur allgemeinen Definition eines Messraums vgl. [Davis 1990, S. 147]. In diesem Anwendungsfall gilt: Die Spielzeit eines Segmentes wird dementsprechend durch das Intervall repräsentiert, das durch den Anfangszeitpunkt der ersten Einstellung und den Endzeitpunkt der letzten Einstellung während eines Imaging des Segmentes festgelegt wird. Dies ist eine Festlegung, die für die Theoriebildung geeignet ist. In Normen, wie z. B. Media Fragments URI (<http://www.w3.org/TR/media-frags/> [letzter Aufruf: 04.03.2015]), werden häufig halboffene Intervalle gewählt.

### Ausgangssituation

Es sei nun die in Abb. 2.7 dargestellte Situation gegeben: Im logischen Baum eines cinematographischen Dokumentes  $D$  werden drei Syntagmen  $S1$ ,  $S2$  und  $S3$  identifiziert. Weitere Detaillierungen der Situation folgen.

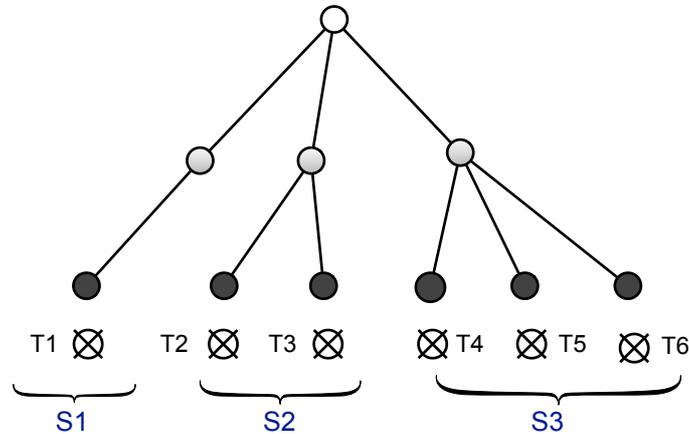


Abb. 2.7: Ausgangssituation mit drei Syntagmen

Die Einstellung  $T1$  kann keinem der Syntagmen  $S2$  bzw.  $S3$  zugeordnet werden. Hier liegt also der Fall vor, dass ein *composite logical object* lediglich ein *basic logical object* enthält. [Metz 1972, S. 172] beschreibt eine solche Situation wie folgt: Eine »ganze Szene wird in einer Einstellung behandelt, d. h. die Einheit der "Handlung" verleiht der Einstellung ihre Autonomie.« Er bezeichnet dieses Syntagma als Plansequenz.<sup>1</sup> In neuerer Literatur über Filmmontage findet sich folgende Definition: »A long and usually complex shot involving much camera movement during which a whole scene is shot in one take without cuts« [Reisz, Millar 2010, S. 338]. [Schmidt, Strauch 2002, S. 71] begründen ausführlich, dass das Wort »Plansequenz« leider ein großer Missgriff ist und schlagen »Planszene« oder »szenische Einstellung« vor. Sie definieren Planszene lediglich als Residualkategorie: »Ihr ist als Ergebnis des vorgegebenen Layoutprozesses eine einzelne Einstellung zugeordnet, die in einer vollständigen Klassifikation des Dokumentes keinem anderen Syntagma zugeordnet wird« [Schmidt 2004, S. 261]. Eine weitere Kennzeichnung wird nicht geliefert. Diese wird in Kapitel 3 dieser Arbeit erfolgen.

Grundsätzlich bildet eine Planszene den zeitlichen Ablauf homomorph ab [Schmidt 2004, S. 262]. Das heißt, die zeitliche Reihenfolge von Ereignissen in einer *content portion* wird durch den Layoutprozess nicht verändert. Die Präsentation kann ggf. in Zeitraffung oder Zeitlupe erfolgen.

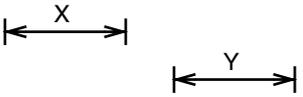
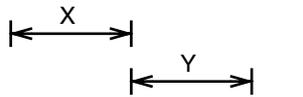
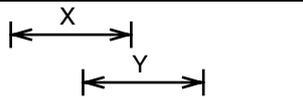
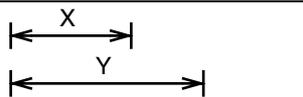
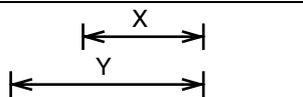
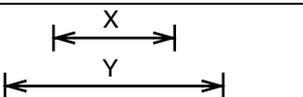
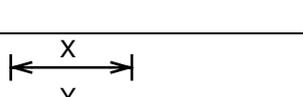
Betrachtet werden dazu nun nicht-triviale Segmente, also Segmente, zu denen mehr als eine Einstellung gehört. In Abb. 2.7 ist dies bei  $S2$  und  $S3$  der Fall.

<sup>1</sup> *plan (fr)*: Einstellung

Da nach Voraussetzung (V1) jeder Einstellung eine diegetische Zeit zugeordnet werden kann, kann – bei Vorliegen von geeigneten temporalen Relationen – eine Relationierung dieser Zeiten durchgeführt werden.

### Allen'sche Zeitlogik

Geeignete temporale Relationen stellt die Allen'sche Zeitlogik [Allen 1983], [Allen, Hayes 1985] bereit. Sie umfasst dreizehn Basisrelationen, die die verschiedenen Möglichkeiten, wie Zeitintervalle zueinander in Beziehung stehen können, beschreiben. Eine Übersicht zeigt Tab. 2.1. In Spalte 1 ist in Klammern jeweils die inverse Relation angegeben.

<i>X before Y</i> ( <i>Y after X</i> )	Das Intervall <i>X</i> endet bevor das Intervall <i>Y</i> startet.	
<i>X meets Y</i> ( <i>Y metBy X</i> )	Das Intervall <i>X</i> trifft das Intervall <i>Y</i> , d. h. der Endzeitpunkt von <i>X</i> ist der Startzeitpunkt von <i>Y</i> .	
<i>X overlaps Y</i> ( <i>Y overlapped-by X</i> )	Das Intervall <i>X</i> startet vor dem Intervall <i>Y</i> und endet während diesem.	
<i>X starts Y</i> ( <i>Y started-by X</i> )	Das Intervall <i>X</i> startet gleichzeitig mit dem Intervall <i>Y</i> und endet vor diesem.	
<i>X finishes Y</i> ( <i>Y finished-by X</i> )	Das Intervall <i>X</i> endet gleichzeitig mit dem Intervall <i>Y</i> , das Intervall <i>X</i> beginnt nach dem Intervall <i>Y</i> .	
<i>X during Y</i> ( <i>Y contains X</i> )	Das Intervall <i>X</i> findet während des Intervalls <i>Y</i> statt, d. h. das Intervall <i>X</i> startet nach dem Intervall <i>Y</i> und endet vor dem Ende von Intervall <i>Y</i> .	
<i>X equal Y</i>	Das Intervall <i>X</i> startet und endet gleichzeitig mit dem Intervall <i>Y</i> .	

Tab. 2.1: Übersicht: Relationen der Allen'schen Zeitlogik

Nach [Schmidt 2004, S. 262] sind für die weitere Analyse die in Tab. 2.2 aufgeführten zeitlichen Verhältnisse und die zugehörigen temporalen Relationen interessant.

Zeitliches Verhältnis	Temporale Relation
(1) diegetisches Aufeinanderfolgen »ohne Lücke«	<i>meets(X, Y)</i>
(2) diegetisches Aufeinanderfolgen »mit Lücke«	<i>before(X, Y)</i>
(3) diegetische Gleichzeitigkeit/Kopräsenz	<i>equal(X, Y)</i>
(4) diegetische Überlappung	<i>overlaps(X, Y)</i> oder <i>starts(X, Y)</i> oder <i>during(X, Y)</i> oder <i>finishes(X, Y)</i>
(5) kein zeitliches Verhältnis	---

Tab. 2.2: Temporale Relationen für syntagmatisch interessante zeitliche Verhältnisse

Betrachtet wird zuerst der Fall, dass zwischen Einstellungen eine der Relationen (1) bis (4) aus Tab. 2.2 vorliegt. Mit Hilfe dieser Relationen kann für Dokumentteile eine chronologische Aggregation durchgeführt werden und für eine Beobachtermenge  $\mathcal{B}$  eines cinematographischen Dokumentes  $D$  wird nach [Schmidt 2004, S. 262] sowie [Bateman, Schmidt 2011, S. 205] definiert:

**Definition 2.5:** Ein Teilbaum  $U$  der logischen Struktur eines cinematographischen Dokumentes  $D$  heißt für eine Beobachtermenge  $\mathcal{B}$  **chronologisch**, wenn für die zugehörige Menge von Einstellungen  $M_U$  eine diegetische Relation  $t_{\mathcal{B}}$  existiert mit  $t_{\mathcal{B}}(T, T')$  für alle  $T, T' \in M_U$ , wobei  $t_{\mathcal{B}} \in \langle \textit{before, meets, overlaps, starts, during, finishes} \rangle$ . Dabei besteht  $\langle \dots \rangle$  aus allen Oder-Verknüpfungen der obigen Relationen und ihrer Inversen.

Ein cinematographisches Dokument  $D$  ist **chronologisch**, wenn alle seine Teilbäume chronologisch sind.

Die Definition verlangt, dass die Beobachter eine der aufgeführten zeitlichen Relationen zwischen je zwei Einstellungen erkennen. Welche der Relationen dies ist, ist für die Klassifikation eines Teilbaums bzw. Dokumentes als »chronologisch« nicht relevant.

Einen Spezialfall der chronologischen Dokumente stellen die Dokumente dar, bei denen »*man sich über den diegetischen Verlauf in der Beobachtermenge  $\mathcal{B}$  einig ist*« [Schmidt 2004, S. 262]. Nach [Bateman, Schmidt 2011, S. 205] wird definiert:

**Definition 2.6:** Ein Teilbaum  $U$  der logischen Struktur eines cinematographischen Dokumentes  $D$  ist **monochron** für eine Beobachtermenge  $\mathcal{B}$ , wenn er chronologisch für diese Beobachtermenge ist, und jeder Beobachter der Beobachtermenge die gleiche temporale Relation als durchweg gültig ansieht.

Ein Segment  $S$  heißt **monochron**, wenn ein Layoutprozess existiert, der einen monochronen Teilbaum der logischen Struktur auf dieses Segment abbildet.

Liegt für ein cinematographisches Dokument der Fall (5) aus Tab. 2.2 vor, handelt es sich um ein nicht-chronologisches Dokument. Ein solches Dokument kann ggf. in mehrere chronologische Teildokumente zerlegt werden (s. Abb. 2.8, oben), die dann jeweils getrennt analysiert werden [Schmidt 2004, S. 262]. Die nicht-zeitliche Beziehung zwischen den Teildokumenten – im Extremfall liegt zwischen allen Einstellungen kein zeitliches Verhältnis vor (s. Abb. 2.8, unten) – fällt nicht in den Bereich einer chronologischen Syntagmatik.

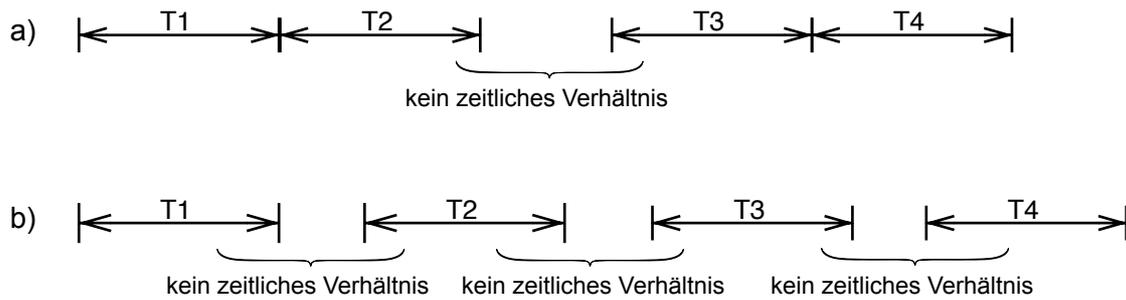


Abb. 2.8: Zeitliche Verhältnisse zwischen Einstellungen

### Spatiale Relationen

Analog zur Beschreibung von temporalen Relationen existieren auch mehrere Ansätze zur Beschreibung von spatialen (räumlichen) Relationen. Häufig verwendet wird der *Region Connection Calculus* von [Randell et al. 1992]. Acht Basisrelationen (bekannt als RCC8) beschreiben alle Möglichkeiten wie zwei Regionen zueinander in Beziehung stehen können. Dies sind *DisConnected* (DC), *Externally Connected* (EC), *Partially Overlapping* (PO), *Tangential Proper Part* (TPP), *Tangential Proper Part inverse* (TPPi), *Non-Tangential Proper Part* (NTPP), *Non-Tangential Proper Part inverse* (NTPPi), *Equal* (EQ). Abb. 2.9 nach [Li, Ying 2003] zeigt eine Darstellung der acht Basisrelationen.

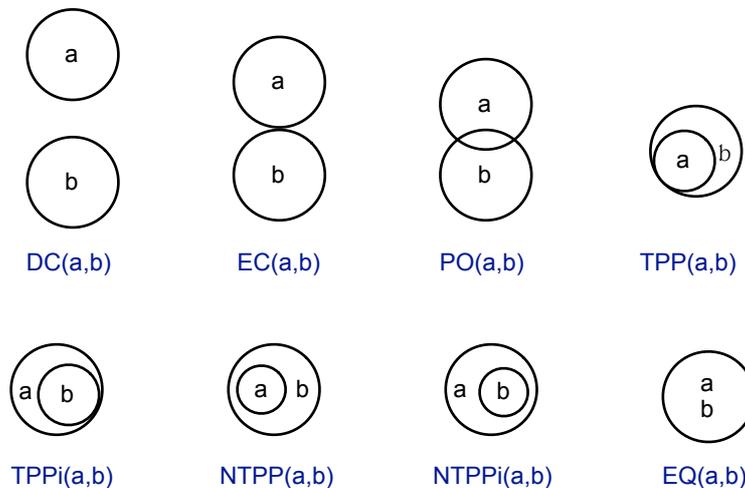


Abb. 2.9: Die acht Basisrelationen nach Randell (RCC8)

Die folgenden Definitionen der syntagmatischen Klassen »Szene«, »Sequenz« und der »narrativen Folge« verlangen jeweils diegetische Zeitverhältnisse, die von allen Beobachtern einer Beobachtermenge konzeptualisiert werden. Es wird also in allen Fällen davon ausgegangen, dass monochrome Dokumente bzw. mindestens monochrome Teilbäume vorliegen.

### Die Szene

Der Begriff »Szene« ist in der Theaterwelt wohl bekannt und bezeichnet dort einen Teil eines Aktes. [Kawin 1992, S. 243] führt aus: »*The scene is an occasion for significant behavior and dramatic interaction. As a theatrical unit taking place in continuous time and space, it has a long history of definition.*« Beim Film wird der Begriff im Drehbuch und als Terminus bei der Produktion verwendet: »[...] *a scene is a complete unit of action capable of being covered in a single shot*« [Kawin 1992, S. 243]. In diesem Sinne formuliert auch [Schmidt 2004, S. 263] folgende Ersetzungsprobe für eine Szene: »*Eine Szene ist für einen menschlichen Beobachter erkennbar, wenn er eine Ersetzungsprobe vornehmen kann: Dies heißt, dass er in einem Gedankenexperiment eine "Szene" als eine einzige Einstellung in einem räumlichen Zusammenhang inszenieren kann.*«

Der Begriff »Szene« wird nun auf cinematographische Dokumente übertragen [Bateman, Schmidt 2011, S. 207]:

**Definition 2.7:** Ein Teilbaum  $U$  der logischen Struktur eines cinematographischen Dokumentes  $D$ , der mindestens zwei Einstellungen enthält, heißt **Szene** für eine Beobachtermenge  $\mathcal{B}$ , wenn

- (1) die diegetischen Räume aller Einstellungen von allen  $B \in \mathcal{B}$  als zusammenhängend konzeptualisiert werden können (Monospacialität),
- (2) die diegetischen Zeiten der Einstellungen von allen  $B \in \mathcal{B}$  als zusammenhängend konzeptualisiert werden (keine zeitlichen Lücken),
- (3) ein Layoutprozess existiert, so dass die durch ihn erzeugte Reihenfolge der Einstellungen und ihre diegetische Reihenfolge von allen Beobachtern als homomorph angenommen werden kann,
- (4) in  $D$  keine weitere Einstellung existiert, die die Bedingungen (1) – (3) erfüllt und die einem *basic logical object* außerhalb dieses Teilbaums zugeordnet ist.

### Bemerkung

Bedingung (1) bedeutet, dass zwischen den diegetischen Räumen aller Einstellungen eine der Randell'schen Basisrelationen aus RCC8, aber **nicht** die Relation DC oder EC vorliegt. [Schmidt 2004, S. 263] betont: »*Der räumliche Zusammenhang muss anhand der abgebildeten Gegenstände erkennbar sein; ein nur vom Beobachter rekonstruierter Zusammenhang, der nicht Bezug auf die tatsächlich abgebildeten Gegenstände nimmt, reicht nicht aus.*« Das bedeutet, der räumliche Zusammenhang muss durch sogenannte »visuelle Anker« erkennbar sein. Technisch gesehen müs-

sen die Bilder (nicht notwendig) aufeinanderfolgender Einstellungen über gemeinsame Pixelmengen so verfügen, dass keine Einstellung existiert, die nicht mit wenigstens einer anderen überlappt.

Die Bedingung (2) ist z. B. erfüllt, wenn zwischen den Einstellungen die Zeitrelation *meets* von den Beobachtern konzeptionalisiert werden kann.

Die Homomorphie-Bedingung in (3) bedeutet, dass die Reihenfolge der Ereignisse in den layoutierten Einstellungen der Reihenfolge der Ereignisse in der diegetischen Welt entsprechen kann.

Die Bedingung (4) der Definition ist ein Maximalitätskriterium: Szenen sind maximal, da das Hinzufügen einer weiteren Einstellung keine Szene mehr liefert.

Die Definition einer Szene stellt lediglich Bedingungen an Raum, Zeit und den Layoutprozess. Dazu merkt [Schmidt 2004, S. 264] an: »Die Definition verzichtet auf jede handlungstheoretische Fundierung; sie ist daher weniger restriktiv als Definitionen, die Handlungseinheiten einbeziehen.«

Abb. 2.10 zeigt nun eine Detaillierung der Situation in Abb. 2.7. Den Einstellungen  $T_2$  und  $T_3$  sind diegetische Zeiten zugeordnet. Zwischen den diegetischen Zeitintervallen der Einstellungen liege die Relation *meets* vor. Der gezeigte Teilbaum stellt daher eine minimale Szene ohne zeitliche Überlappung dar.

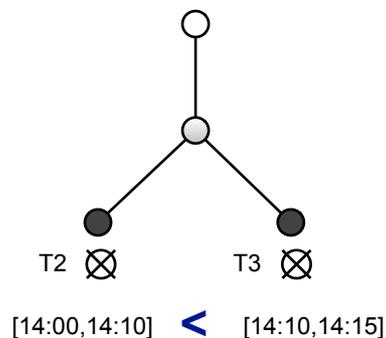


Abb. 2.10: Darstellung einer minimalen Szene (ohne zeitliche Überlappung)

### Die Sequenz

Im Unterschied zu einer Szene wird bei einer Sequenz der zeitliche Zusammenhang nicht verlangt, genauer gesagt, es muss sogar zeitliche Lücken geben. Es gilt folgende Definition [Bateman, Schmidt 2011, S. 210]:

**Definition 2.8:** Ein Teilbaum  $U$  der logischen Struktur eines cinematographischen Dokumentes  $D$ , der aus mindestens zwei Einstellungen besteht, ist für eine Beobachtermenge  $\mathcal{B}$  eine **Sequenz**, wenn

- (1) die diegetischen Räume aller Einstellungen von allen Beobachtern  $B \in \mathcal{B}$  als zusammenhängend konzeptionalisiert werden können,

- (2) die diegetischen Zeiten der Einstellungen von allen Beobachtern  $B \in \mathcal{B}$  als nicht zusammenhängend konzeptionalisiert werden,
- (3) ein Layoutprozess existiert, so dass die durch ihn erzeugte Reihenfolge der Einstellungen und ihre diegetische Reihenfolge von allen Beobachtern als homomorph angenommen werden kann,
- (4) keine weitere Einstellung existiert, die die Bedingungen (1) bis (3) erfüllt.

### Bemerkung

Im Unterschied zu einer Szene muss eine Sequenz nach (2) mindestens eine zeitliche Lücke beinhalten, d. h. zwischen zwei Einstellungen muss die Zeitrelation *before* von den Beobachtern konzeptionalisiert werden.

Eine Sequenz kann – zwischen den Zeitsprüngen – Szenen als Unterelemente enthalten. Abb. 2.11 zeigt links eine Detaillierung der Situation aus Abb. 2.7. Den Einstellungen  $T4$ ,  $T5$  und  $T6$  sind diegetische Zeiten zugeordnet. Zwischen Einstellung  $T4$  und  $T5$  existiert eine zeitliche Lücke. Der gezeigte Teilbaum kann daher als Sequenz klassifiziert werden. Da zwischen  $T5$  und  $T6$  keine zeitliche Lücke existiert, enthält die Sequenz eine Szene als Unterelement. Daher wird ein *composite logical object* in den Baum eingefügt, der die Einstellungen der Szene als Kindelemente enthält (s. Abb. 2.11, rechts).

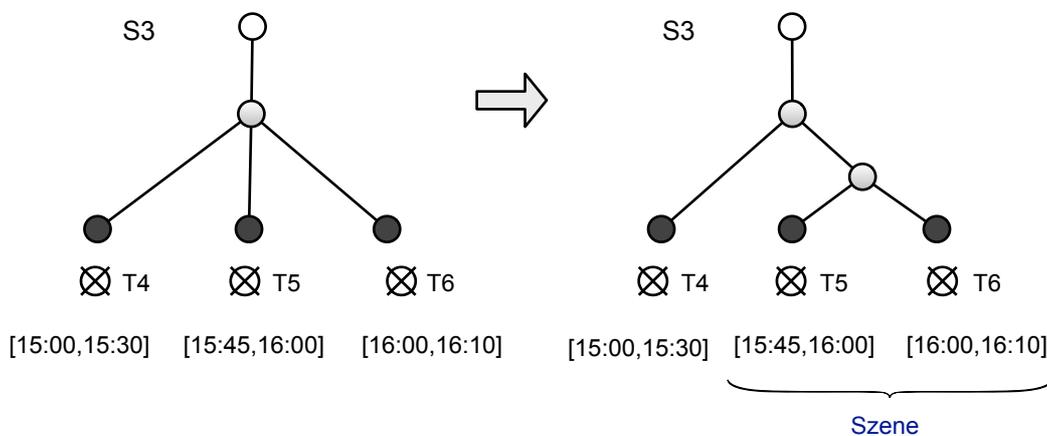


Abb. 2.11: Darstellung einer Sequenz, die eine Szene enthält

Für narrative Filme reichen die Syntagmen Planszene, Szene und Sequenz aus, um auch komplexere Strukturen zu erzeugen. Dies wird schon in [Schmidt, Strauch 2002, S. 74ff] ausführlich begründet und motiviert folgende Definition [Bateman, Schmidt 2011, S. 212]:

**Definition 2.9:** Ein Syntagma ist ein **narrativ basales Syntagma** (*basic narrative syntagma*), wenn es entweder eine Planszene, eine Szene oder eine Sequenz ist.

Ein cinematographisches Dokument, das nur aus narrativ basalen Syntagmen besteht, ist **narrativ basal**.

**Bemerkung**

Der Dokumentbaum eines narrativ basalen Dokumentes (s. Abb. 2.12) hat eine maximale Tiefe von vier [Bateman, Schmidt 2011, S. 212]:

- die Wurzel,
- die Ebene der *composite elements*, die den basalen Syntagmen des Films entsprechen,
- mögliche Szenen, die in Sequenzen enthalten sind, und
- *basic logical objects*, die die Einstellungen als *content portions* enthalten.

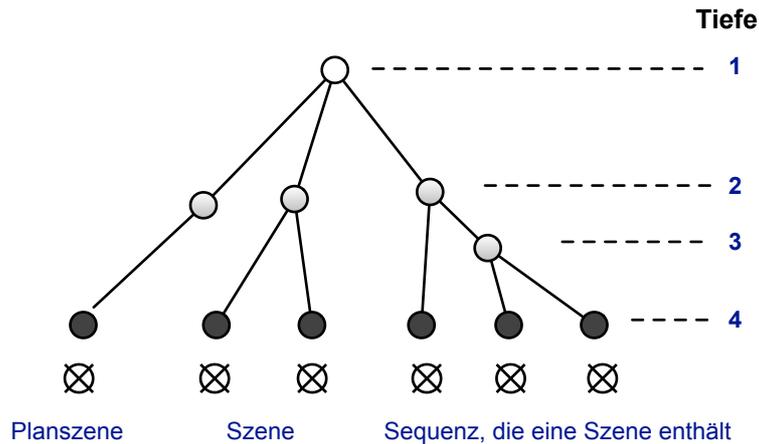


Abb. 2.12: Dokumentbaum eines narrativ basalen Dokumentes

Es bleibt zu fragen, ob und in welchen Anwendungsbereichen narrativ basale Dokumente existieren. [Schmidt 2004, S. 265] nennt als Beispiel für narrativ basale Dokumente Videoüberwachungsprotokolle einer Raumeinheit, »wo ohne Montage Ereignisse (ggf. stückweise) mitgeschnitten werden.« In diesem Fall liefert nur die Raumzeit Zusammenhangsbedingungen. Im entsprechenden Dokumentbaum werden sich daher unterhalb des *document logical root* maximal zwei Ebenen von *composite logical objects* befinden.

Narrativ interessantere Dokumente erfordern Zusammenhangsbedingungen, die nicht nur auf Beobachteranalysen der Raumzeit basieren. Der Dokumentbaum wird in solchen Fällen weitere Hierarchieebenen enthalten können.

### 2.2.3 Narrative Hierarchisierung

In natürlich-sprachlichen Repräsentationen ist eine beliebige Folge von Sätzen noch kein Text. [Brinker 2010, S. 17] ergänzt die in Abschnitt 2.2 notierte Definition von »Text« wie folgt:

*»Der Terminus "Text" bezeichnet eine begrenzte Folge von sprachlichen Zeichen, die in sich kohärent ist und die als Ganzes eine erkennbare kommunikative Funktion signalisiert.«*

Unter Kohärenz<sup>1</sup> versteht man den »semantisch-kognitiven Sinnzusammenhang eines Textes« [Bußmann 2002, S. 351]. Nach [Pörings, Schmitz 2003, S. 197] gibt es folgende Mittel zur Herstellung von Kohärenz:

*»Kohärenz wird auf zweierlei Art erreicht: erstens durch referentielle Kohärenz, d. h. die wiederholte Referenz auf dieselben Objekte innerhalb des Textes und zweitens durch die Verknüpfung von Textteilen durch so genannte Kohärenzrelationen wie "Ursache – Wirkung" und "Kontrast", d. h. durch relationale Kohärenz.«*

Kohärenzrelationen können explizit durch Konnektoren, z. B. Konjunktionen, angezeigt werden oder implizit bleiben. Kohärenz beruht jedoch weniger auf den verwendeten sprachlichen Mitteln, sondern ist ein Ergebnis der Interpretationsanstrengungen des Lesers, wie [Pörings, Schmitz 2003, S. 201] ausführen:

*»Jeder Leser oder Hörer kann einen Text nicht wirklich verstehen, wenn er oder sie nicht die Kohärenzrelationen zwischen den einzelnen Sätzen oder Teilsätzen eines Textes bei der Interpretation herstellt.«*

Eine spezielle Textsorte ist die Erzählung (auch Narration<sup>2</sup>). [Martínez, Scheffel 2007, S. 188] definieren eine Erzählung als »schriftliche oder mündliche Darstellung einer Handlung.« [Genette 2010, S. 11] unterscheidet drei Bedeutungen des Begriffs »Erzählung«. Die erste und am weitesten verbreitete Bedeutung bezeichnet Erzählung als »die narrative Aussage, den mündlichen oder schriftlichen Diskurs (discours), der von einem Ereignis oder einer Reihe von Ereignissen berichtet.« Die zweite Bedeutung bezieht sich auf den narrativen Inhalt: Erzählung als »die Abfolge der realen oder fiktiven Ereignisse, die den Gegenstand dieser Rede ausmachen, und ihre unterschiedlichen Beziehungen zueinander – solche des Zusammenhangs, des Gegensatzes, der Wiederholung usw.« In der dritten (wahrscheinlich der ältesten) Bedeutung bezeichnet Erzählung »den Akt der Narration selber.«

Eine allgemeine Voraussetzung für das Vorhandensein einer Erzählung ist die Existenz einer minimalen narrativen Struktur (Erzählstruktur) zwischen Teilen eines Dokuments. So definiert

<sup>1</sup> Die Definition von Kohärenz wird nicht einheitlich verwendet. Einige Autoren unterscheiden zwischen Kohärenz und Kohäsion. Vgl. dazu die Diskussion bei [Brinker 2010, S. 17].

<sup>2</sup> von *narrare* (lat.): erzählen. Die »Narratologie« (Erzähltheorie, Erzählforschung) ist die Wissenschaft von der Erzählung. Ihr Ziel ist eine systematische Beschreibung der Prinzipien eines erzählenden Dokumentes.

[Prince 1977, S. 31], dass eine minimale Erzählstruktur (*minimal story*) vorliegt, wenn eine Ausgangssituation durch einen die Situation verändernden Vorgang in eine Endsituation übergeht und die Endsituation »invers« zur Ausgangssituation ist.<sup>1</sup> [Prince 1977, S. 19] gibt folgendes Beispiel für eine *minimal story*:

»A man was happy, then he met a woman, then, as a result, he was very unhappy.«

Während nach den Definitionen von Genette, Martínez und Scheffel eine Erzählung in mündlicher oder schriftlicher Form erfolgt, betont [Chatman 1980, S. 121] die Medienunabhängigkeit einer Erzählung:

»One of the most important observations to come out of narratology is that narrative itself is a deep structure quite independent of its medium.«

Nach Chatman kann also auf unterschiedliche Weise, durch gesprochene oder geschriebene Worte, Zeichnungen, Schauspiel, Tanz oder Film, eine Erzählung realisiert werden.

[Bordwell, Thompson 2010, S. 79] haben sich intensiv mit der Narration im Film befasst. Sie definieren:

»We can consider a narrative to be a chain of events linked by cause and effect and occurring in time and space.«

Eine Erzählung beginnt mit einer auslösenden Aktion, es folgt eine Reihe von Ereignissen, die im Zeitverlauf an einem oder mehreren Orten stattfinden und in einem kausalen Zusammenhang stehen.

Die oben zitierte *minimal story* aus [Prince 1977] könnte filmisch<sup>2</sup> mit drei Einstellungen realisiert werden: Einstellung 1 zeigt den glücklichen Mann, Einstellung 2 Mann und Frau und Einstellung 3 den unglücklichen Mann (s. Abb. 2.13). Ohne weiteren erklärenden Text wird der Zuschauer die Geschichte im intendierten Sinn nur verstehen, wenn die Einstellungen auch in der aufgeführten Reihenfolge vorgeführt werden. Das heißt, in diesem Fall wird eine Reihenfolge im Layout impliziert.

---

<sup>1</sup> Im Original: »A minimal story consists of three conjoined events. The first and third events are stative, the second is active. Furthermore, the third event is the inverse of the first. Finally, the three events are conjoined by three conjunctive features in such a way that (a) the first event precedes the second in time and the second precedes the third, and (b) the second event causes the third.«

<sup>2</sup> Hierbei wird von einem Stummfilm oder einem Film ohne Erzähler im Off ausgegangen.

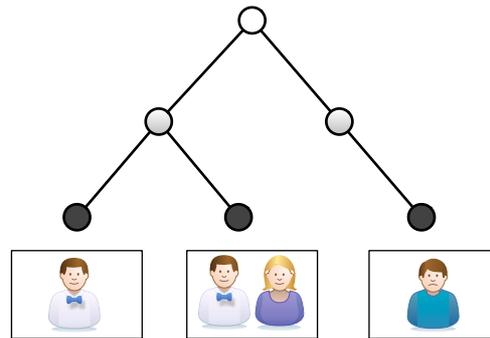


Abb. 2.13: Darstellung der »minimal story« aus Prince in drei Einstellungen

Speziell im Film ist die kausal-chronologische Abfolge von Ereignissen ein zentraler Punkt für eine Erzählung und ihre Repräsentation. Dies betont auch [Bordwell 2010, S. 492] in seiner Definition von Linearität in einer Erzählung:

*»Linearity: In a narrative the clear motivation of a series of cause and effects that progress without significant digression, delays or irrelevant actions.«*

Jedoch kann die Reihenfolge der Inhaltsstücke im Film verändert werden und dennoch für die Zuschauer eine kohärente Struktur darstellen, z. B. durch ein Abweichen von der chronologischen Reihenfolge in Form einer Rückblende (*flashback*) oder einer Vorausblende (*flashforward*).

Eine weitere Möglichkeit ist die sogenannte **Parallelmontage** (*parallel editing*), bei der die Einstellungen von Handlungssträngen im Wechsel montiert werden. Dies erweckt zum einen den Eindruck einer Gleichzeitigkeit von Handlungen, zum anderen den einer Beziehung zwischen den gezeigten Handlungen. Spannung wird erzeugt, wenn die Handlungen auf einen gemeinsamen Schnittpunkt zulaufen [Hickethier 2007, S. 135].

[Doane 2002, S. 194] beschreibt die Wirkung der Parallelmontage und betont die zusätzlichen Anforderungen, die durch die Raum- und Zeitsprünge an die Beobachter gestellt werden:

*»The yoking together of noncontiguous spaces through parallel editing forced a certain denaturalization of the filmic discourse. It required the spectator to accept enormous leaps in space and to allow the disfiguration of continuous time, its expansion or contraction.«*

Im Weiteren werden komplexere narrative Strukturen, wie z. B. die Parallelmontage, aus syntagmatischer Sicht behandelt. Interessant ist insbesondere die Frage, in welcher Form für die Zuschauer eine kohärente Struktur entsteht und wie diese formalisiert werden kann. Auch im Folgenden werden chronologische Dokumente bzw. Teildokumente betrachtet.

Komplexere narrative Dokumente können durch die Kombination von narrativ basalen Syntagmen entstehen und zu weiteren Hierarchisierungen führen [Bateman, Schmidt 2011, S. 212]. Im logischen Dokumentbaum werden die narrativ basalen Syntagmen zunächst unterhalb eines

*composite logical objects* gruppiert. Durch deren Verknüpfung mittels dieses *composite logical objects* entsteht eine weitere Montageebene, welche als zweite Montageebene gekennzeichnet wird. Abb. 2.14 zeigt die Situation mit zwei basalen Syntagmen  $S1$  und  $S2$ . Eine solche Situation wird im Folgenden, wenn bestimmte Voraussetzungen erfüllt sind, als narrative Folge klassifiziert.

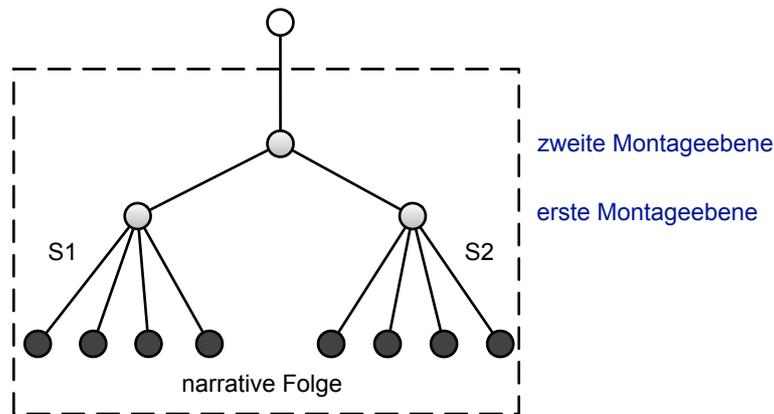


Abb. 2.14: Aggregation zweier narrativ basaler Syntagmen zu einer narrativen Folge

Es gelten folgende Anforderungen [Bateman, Schmidt 2011, S. 213f]:

- (V2) Zwischen  $S1$  und  $S2$  besteht kein in den Messdaten verankerter räumlicher Zusammenhang. Jedoch wird zwischen  $S1$  und  $S2$  von den Beobachtern ein Zusammenhang erkannt, so dass sie beobachterrelational unter einem *composite logical object* gruppiert werden können.
- (V3) Zwischen den Syntagmen der ersten Montageebene liege eine minimale zeitliche Ordnung vor, damit eine Beobachtermenge  $\mathcal{B}$  das Dokument als syntagmatisch chronologisch erfassen kann.

Die in (V2) dargestellte Anforderung wird nun für beliebig viele verschiedene Raumsituationen beschrieben. Dazu wird als Erstes der Begriff »räumlich  $r$ -partit« eingeführt, vgl. [Schmidt 2004, S. 268], [Bateman, Schmidt 2011, S. 216]:

**Definition 2.10:** Ein Teilbaum  $U$  der logischen Struktur eines cinematographischen Dokumentes  $D$  ist syntagmatisch **räumlich  $r$ -partit** für eine Beobachtermenge, wenn

- (1) der Teilbaum mindestens zwei Einstellungen als *content portions* enthält,
- (2) für die Menge  $M_S$  der Einstellungen des zugehörigen Segmentes  $S$  eine Partitionierung existiert mit

$$M_S = \bigcup_{1 \leq i \leq r} M_i \text{ und } M_i \cap M_j \neq \emptyset \text{ für } i \neq j, 1 \leq i, j \leq r,$$

so dass jede dieser partitionierenden Mengen  $M_i$  als narrativ basales Syntagma klassifiziert werden kann.

Ein filmisches Segment heißt **räumlich  $r$ -partit**, wenn es unter Anwendung eines Layoutprozesses das Bild eines räumlich  $r$ -partiten Teilbaums der logischen Struktur ist.

Der in Abb. 2.14 dargestellte Teilbaum, dessen Wurzel der *composite logical object* der zweiten Montageebene ist, ist räumlich bi-partit, er lässt sich in zwei Partitionen aufteilen.

Die Anforderung (V3) wird nun genauer betrachtet. Jedes Segment eines narrativ basalen Syntagmas kann (per Definition) chronologisch geordnet werden. (V3) verlangt nun, dass zwischen den Syntagmen eine minimale zeitliche Ordnung vorliegt. Es stellt sich die Frage, welche zeitliche Beziehung zwischen den Segmenten vorliegt. Für diese Beziehung sind zwei Extremfälle denkbar [Schmidt 2004, S. 266], [Bateman, Schmidt 2011, S. 214]:

*Fall 1:* In allen Segmenten existiert jeweils nur eine Einstellung, die ein zeitliches Verhältnis zu einer Einstellung außerhalb des betrachteten Segmentes hat.

*Fall 2:* Alle Einstellungen aus allen Segmenten lassen sich (ggf. perlschnurartig) in eine zeitliche Ordnung bringen.

In diesem Kontext wird davon ausgegangen, dass alle Varianten zwischen den beiden Extremfällen erlaubt sind, solange das Dokument monochron ist [Bateman, Schmidt 2011, S. 214].

Aufgrund dieser Vorüberlegungen lässt sich eine narrative Folge wie folgt definieren, vgl. [Bateman, Schmidt 2011, S. 216]:

**Definition 2.11:** Für die syntagmatische Klassifikation eines Teilbaums  $U$  der logischen Struktur eines cinematographischen Dokumentes  $D$  als **narrative Folge** ist es notwendig, dass

- (1) der Teilbaum räumlich  $r$ -partit mit  $r \geq 2$  ist,
- (2) die Beobachter in jeder Partitionsmenge  $M_i$ ,  $1 \leq i \leq r$ , eine Einstellung finden, die in einer temporalen Relation zu mindestens einer Einstellung in einer anderen (verschiedenen) Menge  $M_j$ ,  $j \neq i$ ,  $1 \leq j \leq r$ , steht.

### **Bemerkung**

Zusammenfassend lässt sich sagen, dass die wesentlichen Eigenschaften einer narrativen Folge der Wechsel des Ortes und die temporale Relation zwischen den Einstellungen verschiedener Partitions Mengen sind. [Schmidt 2004, S. 269] bezeichnet dies kurz auch als »*Wechsel des Ortes bei weiterlaufender Zeit*«.

Wird Bedingung (2) der Definition 2.11 nicht erfüllt, d. h. liegen (mindestens) zwei Mengen zeitlich nicht verbundener Einstellungen vor, so dass der ganze Teilbaum nicht mehr chronologisch ist, kann es sich nach der »*Großen Syntagmatik von Metz*« um ein paralleles, also ein achronologisches Syntagma handeln [Bateman, Schmidt 2011, S. 217].

### Beispiele für narrative Folgen

Als Beispiele für narrative Folgen nennen [Schmidt 2004, S. 269] und [Bateman, Schmidt 2011, S. 217] folgende Situationen:

- (1) Ein **Itinerar** (*itinerary*<sup>1</sup>), d. h. eine Situation, bei der im Zeitverlauf eine Figur verschiedene Orte besucht. Dies kann z. B. ein Bericht oder eine Dokumentation über eine Reise mit verschiedenen Reisestationen sein oder die Darstellung verschiedener Etappen eines Handlungsplans, der an verschiedenen Orten durchgeführt wird. Der logische Teilbaum eines Itinerars ist  $r$ -partit, wobei  $r$  gleich der Anzahl der verschiedenen Orte ist.
- (2) Eine **polyspatiale Alternanz**, d. h. eine Situation, bei der zwischen zwei (oder mehr) Orten hin und her gewechselt wird. Beispiele sind Dialogsituationen, wie z. B. ein Telefongespräch, bei dem die Gesprächspartner abwechselnd gezeigt werden, oder etwa eine Verfolgungsjagd, bei der Verfolger und Verfolgter abwechselnd gezeigt werden. Der logische Teilbaum ist in diesem Fall analog zu dem in Abb. 2.14 dargestellten. Ein Layoutprozess angewandt auf diesen Baum erzeugt eine Präsentation, bei der die Einstellungen in ihrer zeitlichen Ordnung gezeigt werden, die Sequenzen werden dabei »ineinander verzahnt«.

#### 2.2.4 Polyspatiale Alternanz

[Schmidt 2004, S. 270] schlägt vor, zur Darstellung einer polyspatialen Alternanz eine Tabelle zu verwenden. Die Zahl der Spalten der Tabelle entspricht der Zahl der verschiedenen Raumumgebungen. Die (durchnummerierten) Einstellungen werden in die Zeilen eingetragen, so dass bei einer Lesereihenfolge von oben nach unten die Serialisierung und die Wechsel des Ortes (»Raumsprünge«) erkennbar sind.

Abb. 2.15 rechts zeigt die Darstellung einer narrativen Folge als Wechseltabelle. Im Beispiel existieren zwei verschiedene Raumumgebungen. Es gilt:

$$\begin{aligned} SpaceTime_1 &\subseteq SpaceTime \text{ und} \\ SpaceTime_2 &\subseteq SpaceTime \text{ und} \\ SpaceTime_1 \cap SpaceTime_2 &= \emptyset. \end{aligned}$$

---

<sup>1</sup> Reiseplan, Reiseroute

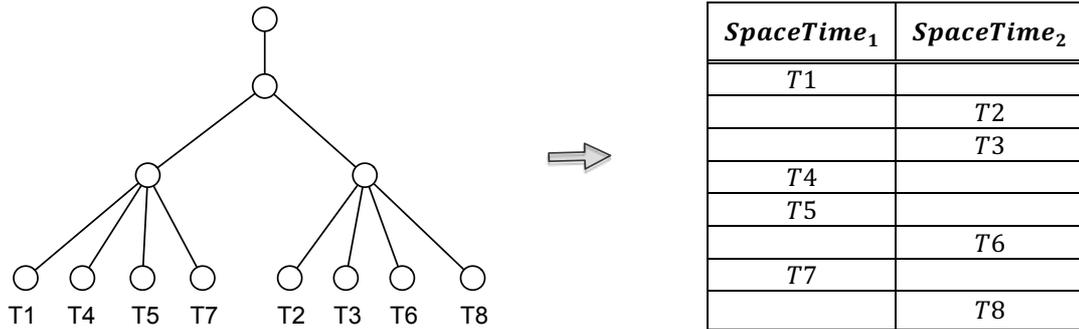


Abb. 2.15: Narrative Folge und ihre Darstellung als Wechseltabelle

Die Tabelle in Abb. 2.15 kann auch als Wechselgraph (Übergangsgraph, *transition graph*) dargestellt werden [Bateman, Schmidt 2011, S. 218]. Die Einstellungen werden als Knoten dargestellt, die Kanten verbinden im Layout zeitlich benachbarte Einstellungen, die zu verschiedenen Partitionen gehören (s. Abb. 2.16). Der Wechselgraph enthält somit alle räumlichen Sprünge im Layout. Damit lediglich **ein** zusammenhängender Graph entsteht, werden auch die benachbarten Einstellungen der gleichen Partition verbunden. In Abb. 2.16 werden sie durch doppelte Linien dargestellt.

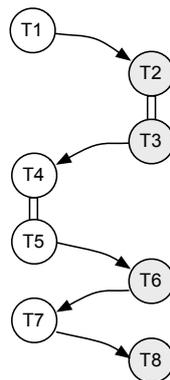


Abb. 2.16: Wechselgraph (Übergangsgraph) einer narrativen Folge

Für einen  $r$ -partiten Wechselgraph werden folgende **Bezeichnungen** eingeführt (vgl. [Schmidt 2004, S. 270f]):

Für ein Layout  $L$  seien  $r$  Partitionsmengen  $M_i, 1 \leq i \leq r$ , gegeben. Der Wechselgraph wird mit  $Alt_L(V, E)$  bezeichnet, wobei die Knoten  $V = V(Alt_L)$  aus der Menge aller Einstellungen bestehen und die Kanten  $E = E(Alt_L)$  aus den im Layout benachbarten Einstellungspaaren  $(T_\mu, T_\nu)$  mit  $T_\mu \in M_i$  und  $T_\nu \in M_j, 1 \leq \mu, \nu \leq n, \mu \neq \nu, 1 \leq i, j \leq r, i \neq j$ . In  $Alt_L(V, E)$  sind damit alle »Raumsprünge« repräsentiert.

## Kohärenz

Der Graph in Abb. 2.16 ist bi-partit in dem Sinne, dass hier zwei Raumzeitumgebungen miteinander verbunden sind. Es stellt sich die Frage, wie Beobachter einen inhaltlichen Zusammenhang, eine Kohärenz, zwischen den Einstellungen verschiedener Partitionen konzeptualisieren können.

Nach [Schmidt 2004, S. 270] ist jegliches Kriterium, das einen räumlichen Zusammenhang voraussetzt, für diese Wechsel als Zusammenhangskriterium ausgeschlossen (siehe auch (V2)). Weiterhin formuliert er:

*»Da wir aber für den filmischen Zusammenhang nur einen Ausgabestrom unterstellen, sind diese mit einem Layoutprozess festgelegten "Raumsprünge" zwischen benachbarten Einstellungen aus unterschiedlichen Raumzeitumgebungen für das Verständnis des alternierenden Falls entscheidend.«*

[Hickethier 2007, S. 113] formuliert folgendes (allgemeine) Prinzip der Kohärenz:

*»Auch wenn ständig gewechselt wird, um die Aufmerksamkeit des Betrachters zu halten, muss zwischen allen Teilen ein Zusammenhang hergestellt werden, der die Teile als Bestandteile einer Einheit erscheinen lässt. So wie die Bedeutungen sich aus dem Kontext bestimmen, in dem die einzelnen Bedeutung tragenden Elemente auftreten, so müssen sich die Teile einer audiovisuellen Einheit zu dieser Einheit erkennbar zusammenfügen. Kohärenz zeichnet eine Bilderfolge als filmischen "Text" aus.«*

Nach [Hickethier 2007, S. 113f] kann Kohärenz auf unterschiedliche Weise erzielt werden:

- durch einstellungsinterne Verweise, die Einstellungen miteinander in Beziehung setzen,
- durch einstellungsübergreifende Elemente (wie z. B. gleiche Bewegungsverläufe, fort-dauernde Musik)
- und nicht zuletzt durch erkennbaren Sinnzusammenhang, eine vom Zuschauer im Gezeigten erkannte Geschichte, eine Erzählform, ein Thema etc.

Einstellungsinterne Verweise sorgen für eine referentielle Kohärenz. Diese Verweise können durch Objekte realisiert sein, die in den Einstellungen beider Partitionen vorkommen, z. B. haben bei einem Telefongespräch beide Gesprächspartner ein Telefon. Ein einstellungsübergreifendes Element in Form eines gleichen Bewegungsablaufes ist bei einem Telefongespräch oder in einer Verfolgungssituation leicht erkennbar. In den genannten Fällen wird es für einen geeigneten Beobachter möglich sein, einen Sinnzusammenhang, eine Kohärenz, zwischen den Einstellungen zu erkennen.

Formal kann die Kohärenz zwischen Einstellungen durch eine Relation beschrieben werden. Fasst man die Aussagen von [Hickethier 2007] und [Schmidt 2004] zusammen, muss diese Rela-

tion eine Beziehung zwischen beobachteten oder schätzbaren Objekten bzw. Handlungen beschreiben. Sie darf jedoch keine Konzeption eines tatsächlichen räumlichen Zusammenhangs etablieren, da dann keine Polyspatialität vorliegt und ggf. basale Syntagmen zu Klassifikation heranzuziehen sind. Dies kann wie folgt notiert werden, vgl. [Schmidt 2004, S. 271], [Bateman, Schmidt 2011, S. 219f]:

**Definition 2.12:** Im bi-partiten Fall ist eine (polyspatiale) **Alternanz** für einen Beobachter konzeptionalisierbar, wenn zu seinem Beobachterwissen eine 2-stellige Relation  $R$  gehört, die folgende Eigenschaften erfüllt:

- (1)  $R \subseteq SpaceTime_1 \times SpaceTime_2$ , d. h.  $R$  ist über zwei Mengen von Entities definiert, eine aus  $SpaceTime_1$ , die andere aus  $SpaceTime_2$ .
- (2) Für ein gegebenes Layout  $L$  lässt sich  $R$  über alle Wechsel, die durch den entsprechenden Wechselgraph  $Alt_L(V, E)$  gegeben sind, etablieren.
- (3)  $R$  ist symmetrisch. Alternativ kann eine weitere Relation  $R'$  für den Beobachter angenommen werden, die die Bedingungen (1) und (2) erfüllt und für die die Disjunktion  $R \vee R'$  eine symmetrische Relation bildet.
- (4)  $SpaceTime_1$  und  $SpaceTime_2$  werden vom Beobachter nicht als räumlich zusammenhängend konzeptionalisiert.

Zu beachten ist, dass man nur von einer Alternanz für eine gegebene Layout-Struktur und daher in Bezug auf den entsprechenden Wechselgraphen sprechen kann. Alternanz ist eine Klassifikation, die nur auf Segmente anwendbar, also auch nur in einem layoutierten cinematographischen Dokument beobachtbar und nur indirekt abhängig von der logischen Struktur des Dokumentes ist. Die logische Struktur muss jedoch so sein, dass eine Alternanz möglich ist [Bateman, Schmidt 2011, S. 221].

Mit diesen Vorüberlegungen ist es nun möglich, alternante Segmente des Typs »Wechsel des Ortes bei fortschreitender Zeit« (»*change of location with [...] running time*«) zu definieren, vgl. [Schmidt 2004, S. 273], [Bateman, Schmidt 2011, S. 221]:

**Definition 2.13:** Ein Segment  $S$  eines cinematographischen Dokumentes  $D$  ist  **$r$ -alternant** bezüglich eines gegebenen Layoutprozesses  $L$  und einer Beobachtermenge  $\mathcal{B}$ , wenn

- (1)  $S$  monochron für die Beobachtermenge  $\mathcal{B}$  ist,
- (2)  $S$  räumlich  $r$ -partit mit den Partitions Mengen  $M_1, \dots, M_r$  ist,
- (3) für jedes Paar von Partitions Mengen  $M_i$  und  $M_j$ ,  $1 \leq i, j \leq r, i \neq j$ , für die eine Transition zwischen einem Element der ersten und einem Element der zweiten Partitions Menge existiert, d. h.  $(T, T') \in E(Alt_L)$  mit  $T \in M_i$  und  $T' \in M_j$ , eine symmetrische Relation  $R \subseteq SpaceTime(T) \times SpaceTime(T')$  von allen Beobachtern einer Beobachtermenge etabliert wird und die für alle Transitionen der Partition gültig ist,

- (4) für alle Transitionspaare zwischen den Elementen einer Partition alle Beobachter der Beobachtermenge die Raumzeiturbilder der Elemente dieser Paare als disjunkt erfassen, d. h. für alle  $(T, T') \in E(Alt_L)$  mit  $T \in M_i$  und  $T' \in M_j, 1 \leq i, j \leq r, i \neq j$ , und für alle  $B \in \mathcal{B}$  gilt  $SpaceTime(T) \times SpaceTime(T') = \emptyset$ ,
- (5) es in der Darstellung entsprechend dem gegebenen Layoutprozess  $L$  zwischen den verschiedenen Elementen von je zwei Partitions Mengen mindestens drei Transitionen gibt, also für jede Partitionsmenge  $M_i$  zwischen dieser und einer anderen Partitionsmenge  $M_j, i \neq j$ , mindestens dreimal in  $Alt_L$  gewechselt wird:

$$|\{(T, T') | (T, T') \in E(Alt_L) \text{ mit } T \in M_i \text{ und } T' \in M_j\}| \geq 3$$

### Bemerkung

Die in der Definition geforderte Relation formalisiert die Kohärenz für die Beobachter der alternanten Segmente. Beispiele für kohärenzbildende Relationen sind:

- *Telefoniert*( $a, b$ ):  $a$  und  $b$  telefonieren miteinander,
- *Verfolgt*( $a, b$ ):  $a$  verfolgt  $b$  mit der Relation *Wird-Verfolgt*( $a, b$ ):  $a$  wird von  $b$  verfolgt,

wobei jeweils  $a$  aus  $SpaceTime_1$  und  $b$  aus  $SpaceTime_2$ .

In Bedingung (5) der Definition werden mindestens drei Transitionen gefordert. Dies ist plausibel, denn nur wenn eine solche Mindestzahl von Wechseln vorliegt, kann ein Beobachter überhaupt eine Kohärenzrelation über mehrere Wechsel bestätigen [Bateman, Schmidt 2011, S. 219].

Ein 2-alternantes Segment wird als bi-partite und ein 3-alternantes Segment als tri-partite Alternanz bezeichnet [Bateman, Schmidt 2011, S. 222].

Die basalen Syntagmen wurden als spezielle Teilbäume der logischen Struktur eines Dokumentes ausgezeichnet. Zur ihrer Definition (s. o.) genügte die Forderung der Existenz eines Layoutprozesses. Eine Alternanz ist hingegen nur im layoutierten Dokument beobachtbar, daher muss die Definition einer Alternanz auch zuerst für Segmente erfolgen, erst dann kann sie für die logische Struktur definiert werden [Bateman, Schmidt 2011, S. 226].

**Definition 2.14:** Ein Teilbaum  $U$  der logischen Struktur eines cinematographischen Dokumentes  $D$  ist **syntagmatisch  $r$ -alternierbar** für eine Beobachtermenge, wenn ein Layoutprozess existiert, der ein entsprechendes  $r$ -alternantes Segment erzeugt.

### 2.3 Paradigmatische Organisation eines Films

Nach [Bateman 2007] wird sowohl im Framework der »*Grande Syntagmatique*« von Christian Metz als auch bei darauf basierenden Anwendungen (z. B. [Roth 1983]) keine hinreichende Unterscheidung der syntagmatischen und paradigmatischen Achsen der semiotischen Organisation getroffen. Zur Produktion eines Films und zu seiner Analyse ist jedoch eine Betrachtung beider Achsen notwendig, wie auch [Monaco 2009, S. 174] anmerkt:

*»Diese zwei Bedeutungsachsen – die paradigmatische und die syntagmatische – sind wertvolle Hilfsmittel zum Verständnis dessen, was ein Film bedeutet. Tatsächlich hängt der Film als Kunst völlig von diesen zwei Auswahlmöglichkeiten ab. Nachdem sich der Filmemacher entschieden hat, was er filmen will, sind die zwei zwingenden Fragen, wie er dies filmen soll (welche Wahlen er trifft: die paradigmatische) und wie er diese Aufnahmen präsentieren soll (wie er sie montiert: die syntagmatische).«*

Ausgangspunkt einer paradigmatischen Beschreibung sind die Relationen zwischen den Einstellungen eines Films. [Bateman, Schmidt 2011, S. 166] formulieren:

*»Our paradigmatic description will, in particular, classify the kinds of relationships that hold between elements within particular montage configurations.«*

Die Beziehungen zwischen den einzelnen Einstellungen sind entscheidend für das Verständnis eines Films, denn für jedes neue Element, das ein Beobachter perzipiert, muss er entscheiden, in welcher Beziehung dies zu dem steht, was zuvor gezeigt wurde [Bateman, Schmidt 2011, S. 171]. Ganz allgemein ausgedrückt, stellt sich also der Beobachter die Frage: »Ist das neue Element verschieden oder ähnlich zu dem, was ich bisher gesehen oder gehört habe?« Dies wird während der Betrachtung des gesamten Films durchgeführt und so entstehen Gruppen mit mehr oder weniger eng verbundenen Elementen. Es zeigt sich, dass zu deren Klassifikation relativ wenige Grundtypen benötigt werden, die [Bateman, Schmidt 2011, S. 171] als »*broad syntagmatic dependency types* (BST)« bezeichnen. Zu ihrer formalen Beschreibung verwenden [Bateman, Schmidt 2011, S. 172f] folgende Notation: Großbuchstaben, wie *W*, *X*, *Y*, *Z*, repräsentieren Folgen, ein Element in einer Folge wird mit  $X_i$ ,  $Y_i$ , ... bezeichnet. Zusätzlich (quasi als weitere Dimension) wird der Grad der Abhängigkeit angegeben: Griechische Buchstaben kennzeichnen eine Abhängigkeit (ein Element ist vom anderen abhängig, *hypotaxis*), Zahlen eine Unabhängigkeit (beide Elemente spielen eine gleiche Rolle, *parataxis*).<sup>1</sup>

<sup>1</sup> vgl auch die Beschreibung der Dimension TAXIS des Netzwerks der »*Grande Paradigmatique*«.



werk wird – auch als Reminiszenz an Christian Metz – als »*Grande Paradigmatique*« bezeichnet. [Bateman, Schmidt 2011, S. 168]:

»The resulting system network we term a **grande paradigmatic**, both to bring out its obvious debt to Metz and to make explicit its exclusive focus on the paradigmatic options available within its semiotic mode.«

Das Netzwerk sieht vor, dass jede Beziehung zwischen zwei Einstellungen immer mit den drei paradigmatisch unabhängigen Dimensionen PROJECTION, TAXIS und PLANE klassifiziert werden kann, wobei jede Dimension weitere Auswahlmöglichkeiten, die weitere Detaillierungen darstellen, enthält. Simultane Auswahlmöglichkeiten werden im Netzwerk durch geschweifte Klammern angezeigt. Betrachtet man das Netzwerk von links nach rechts, erreicht man die Detaillierungen innerhalb der Dimensionen.

Die drei Dimensionen des Netzwerks PROJECTION, TAXIS und PLANE werden mit jeweils einem Schaubild erläutert; die Schaubilder insgesamt ergeben die graphische Repräsentation des gesamten Modells (s. Abb. 2.20).

### PROJECTION

Abb. 2.17 zeigt das Teilnetzwerk der Dimension PROJECTION.

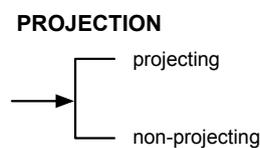


Abb. 2.17: Dimension PROJECTION der »Grande Paradigmatique«

In der systemisch-funktionalen Grammatik [Halliday, Matthiessen 2014] gehört die Relation *projection* zu den logisch-semantischen Relationen. Sie beschreibt die Beziehung zwischen einer »*mental clause*« und einer »*idea clause*«. Beispiele sind »*He says 'that's enough'*« (Typ: *quoting locution, projection of wording*) oder »*He think that it's too hot*« (Typ: *reporting idea, projection of meaning*) [Halliday, Matthiessen 2014, S. 433]. [Martinec, Salway 2005] betrachten Relationen zwischen Text und Bildern. In dieser Situation liegt eine Relation vom Typ *projection* vor, wenn Inhalte, die durch Text oder Bild repräsentiert werden, auch im jeweils anderen Modus dargestellt werden [Martinec, Salway 2005, S. 354]. Insbesondere im Comic kommt die Relation *projection* häufig vor. Hier wird die mentale Welt eines Darstellers durch Sprech- bzw. Denkblasen visualisiert.

Im Film liegt die Relation *projection* zwischen zwei Einstellungen vor, wenn beim Wechsel von einer Einstellung zur folgenden ein Wechsel vom Fokalisierer (*focalizer*) zur mentalen Welt dieses Fokalisierers stattfindet. Dies können Erinnerungen, Ahnungen oder Träume oder Perzepti-

onen sein [Bateman 2007, S. 45]. Für den Fall der perzeptiven Projektionen hat [Branigan 1984, S. 103] den Begriff *Point-of-View-Shot (POV-Shot)* eingeführt:

»The POV shot is a shot in which the camera assumes the position of a subject in order to show us what the object sees.«

Bei einem *POV-Shot* nimmt die Kamera die Perspektive einer wahrnehmenden Figur ein, um dem Zuschauer zu zeigen, was diese Figur sieht. [Grimm 1996, S. 52] beschreibt dies kurz und treffend mit den Worten: »Ein Blick ist zu sehen, dann das, worauf er/sie blickt.« Häufig zeigen die Einstellungen zusammenhängende Räume und ununterbrochene Zeiten.

## TAXIS

Abb. 2.18 zeigt das Teilnetzwerk der Dimension TAXIS.

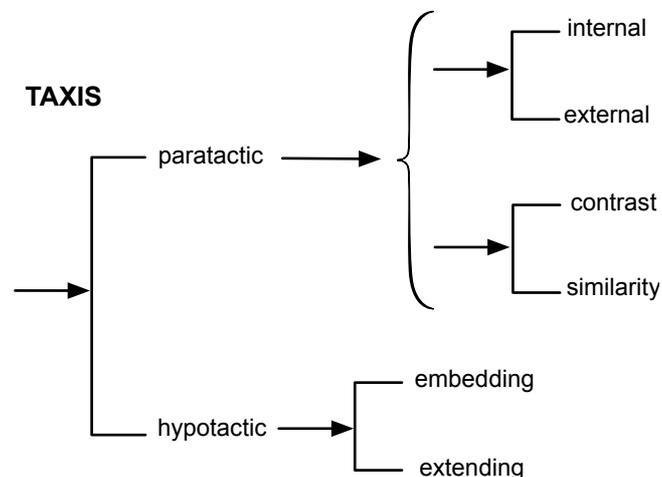


Abb. 2.18: Dimension TAXIS der »Grande Paradigmatique«

Der Begriff *taxis* stammt ebenfalls aus der systemisch-funktionalen Grammatik nach Halliday [Halliday, Matthiessen 2014, S. 451f]. *Taxis* bezeichnet den Grad der gegenseitigen Abhängigkeit (Interdependenz, *interdependency*). Es werden zwei Grade unterschieden: *hypotaxis*<sup>1</sup> (*unequal status*) und *parataxis*<sup>2</sup> (*equal status*). *Hypotaxis* (dt. Hypotaxe) ist die Relation zwischen einem abhängigen Element (Satellit) und seiner Dominante, dem Element von dem es abhängig ist (Kern, Nukleus). Ein Beispiel ist die Unterordnung von Nebensätzen unter Hauptsätze. *Parataxis* (dt. Parataxe) ist die Relation zwischen zwei ähnlichen Elementen vom gleichen Status, eines beginnt, das andere setzt fort. Dies ist z. B. der Fall bei einer Aneinanderreihung von selbstständigen Hauptsätzen.

Eine Relation zwischen zwei Einstellungen eines Films ist hypotaktisch (*hypotactic*), wenn eine raumzeitliche Entwicklung stattfindet, bei der der Ort und/oder die Zeit des folgenden Elementes direkt mit dem vorangegangenen verbunden ist. Hier liegt also der *broad syntagmatic depen-*

<sup>1</sup> *hypotaxis* (griech.): *hypo* (griech.): unter, *taxis* (griech.): Ordnung

<sup>2</sup> *parataxis* (griech.): Beiordnung, Koordination

*dependency type 1* (BST1) vor. Im Netzwerk werden bei einer hypotaktischen Relation zwei Optionen unterschieden: *extending*, wenn eine Erweiterung des Raum- oder Zeitgebietes stattfindet, *embedding*, wenn dies nicht der Fall ist.

Parataktische Relationen werden im Film durch eine Folge von Segmenten, die in keiner Abhängigkeitsbeziehung zueinander stehen, realisiert. Im Teilnetzwerk *paratactic* findet sich eine Kreuzklassifikation mit zwei Dimensionen: *internal / external* und *similarity / contrast*.

Die Klassifikation *internal / external* geht auf [van Leeuwen 1991] zurück. Sie reflektiert – in allgemeinerer Form – die aus der Narratologie bekannte Unterscheidung zwischen *histoire* und *discourse time*. Eine Relation vom Typ *internal* konstruiert eine Beziehung »*in the telling of the story*«, d. h. hier findet eine Gegenüberstellung (*comparison*) von Themen (*topics*) der Erzählung statt. Eine Relation vom Typ *external* repräsentiert eine Beziehung »*between "the world of events" depicted in the story*«. Es liegt also eine Gegenüberstellung von Ereignissen vor. In beiden Fällen ist somit der Zweck der Einstellungsfolge die Darstellung einer Gegenüberstellung, allerdings muss im Falle einer Relation vom Typ *internal* der Handlungsverlauf nicht fortgesetzt werden, während dies bei einer Relation vom Typ *external* der Fall ist, z. B. bei einer Verfolgungsjagd.

Parallel dazu werden Relationen als *similarity* oder *contrast* klassifiziert. Die Relation *similarity* besteht zwischen einer Folge von Segmenten, die in keiner abhängigen Beziehung zueinander stehen. Die Elemente können verschiedene denotative Interpretationen besitzen, entscheidend ist, dass die Relation, die zwischen den Elementen besteht, konstant ist.<sup>1</sup> Syntagmatisch gesehen liegt hier der *broad syntagmatic dependency type 2* (BST2) vor.

Bei einer Relation vom Typ *contrast* liegt zwischen aufeinanderfolgenden Elementen eine konstante, sich wiederholende Relation vor, die in sich bereits kontrastierend ist. Syntagmatisch gesehen müssen hierzu (mindestens) Paare von Elementen vorliegen. Eine Realisation vom Typ *contrast* entspricht dem *broad syntagmatic dependency type* »*multitracking*«, also BST4.

[Bateman, Schmidt 2011, S. 177] weisen darauf hin, dass die Relationstypen '*hypotactic: embedding*' und '*paratactic: contrast*' in der verbalen Sprache – im Unterschied zum Film – nicht (jedenfalls nicht als »*major structuring devices in their own right*«) vorkommen.

---

<sup>1</sup> Dies entspricht dem internen Fall des achronologischen Syntagms der zusammenfassenden Klammerung (*achronological bracket syntagma*) nach Metz.

**PLANE**

Abb. 2.19 zeigt das Teilnetzwerk der Dimension PLANE.

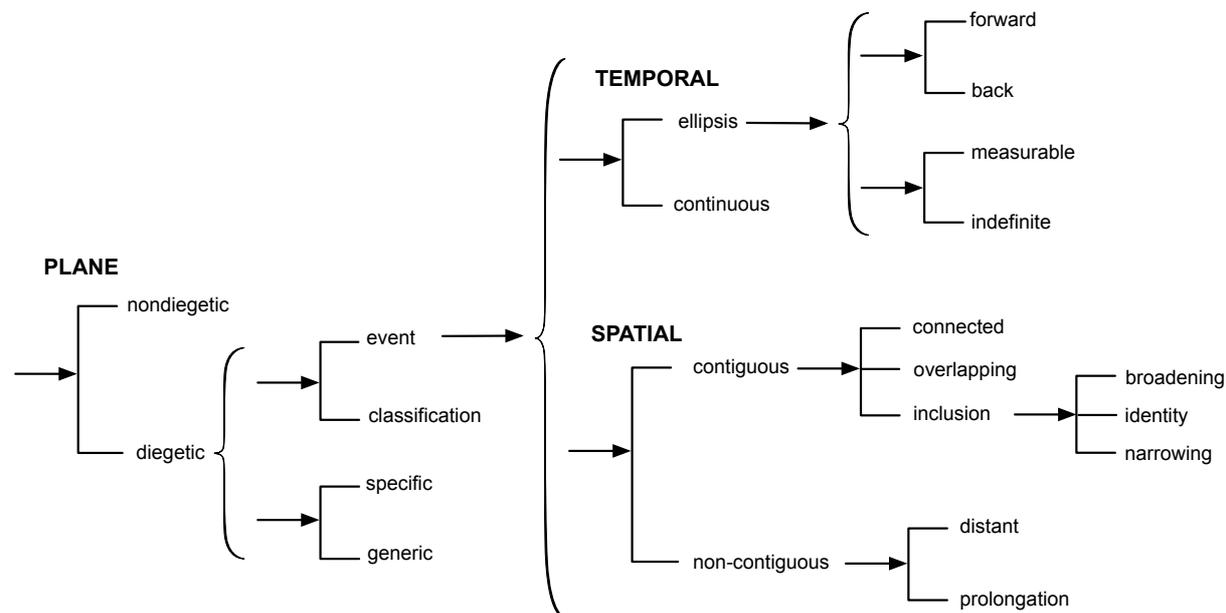


Abb. 2.19: Dimension »PLANE« der »Grande Paradigmatique«

Die oberste Option im Teilnetzwerk PLANE ist die Unterscheidung zwischen *diegetic* / *nondiegetic* nach [Colin 1995]. Die *diegesis* ist für den cinematographischen Fall der gesamte hör- und sehbare Bereich, der für die Figuren auf der Leinwand als wahrnehmbar angenommen werden kann. Dementsprechend ist etwas nicht-diegetisch, wenn es nur für den Beobachter des Dokuments bestimmt ist und nicht für die Figuren im Film. Dies ist z. B. bei Filmmusik oder textuellen Einblendungen der Fall.

Eine *diegetic* Relation lässt sich weiter spezifizieren durch die Dimensionen *event* / *classification* und *specific* / *generic*. Die Unterscheidung *event* / *classification* geht zurück auf [Kress, van Leeuwen 1996], die die verschiedenen Typen eines »visuellen Prozesses« in Bildern (*static images*) diskutieren. Nach ihrer Meinung können alle Bilder in zwei Kategorien eingeteilt werden: narrative und konzeptuelle. In narrativen Bildern werden Aktionen dargestellt, dabei kann es sich um »*unfolding actions and events, processes of change, transitory spatial arrangements*« handeln [Kress, van Leeuwen 1996, S. 56]. Konzeptuelle Bilder zeigen etwas Statisches und Zeitloses. Eine ihrer Unterkategorien sind die klassifizierenden Bilder, die die verschiedenen Relationen zwischen Personen, Orten oder Dingen darstellen oder sie in einer Taxonomie präsentieren. Diese beiden Typen »narrativ« und »klassifikatorisch« wurden in das Netzwerk als *event* bzw. *classification* eingefügt.

Das Teilnetzwerk innerhalb von *event* enthält die Subklassifikationen TEMPORAL und SPATIAL, basierend auf Ideen von [Burch 1973]. Eine temporale Beziehung ist entweder *continuous*, d. h. ununterbrochen, fortdauernd, kontinuierlich, oder vom Typ *ellipsis*, d. h. sie enthält eine zeitliche

Lücke. Innerhalb von *ellipsis* gibt es die parallelen Dimensionen *forward / back*, d. h. es gibt einen Zeitsprung in die Zukunft bzw. in die Vergangenheit, und *measurable / indefinite*, d. h. eine genaue Angabe über die Zeitdifferenz ist möglich oder schätzbar bzw. ist die Zeitdifferenz unbestimmt.

Bei einer spatialen Relation sind Raumgebiete entweder *contiguous* (zusammenhängend, angrenzend) oder *non-contiguous* (nicht-zusammenhängend). Bei der *contiguous*-Relation (basierend auf [Randell et al. 1992], s. auch 2.2.2) wird zwischen *connected* (verbunden), *overlapping* (überlappend) und *inclusion* (enthalten sein) unterschieden. Eine Relation vom Typ *inclusion* enthält die Detaillierungen *broadening*, *identity* und *narrowing*. Mit den Relationstypen *broadening* bzw. *narrowing* können Wechsel zwischen Einstellungsgrößen charakterisiert werden: *broadening*, wenn das »Kamera-Auge« sich vom Objekt entfernt, z. B. bei einem Wechsel von der Halbtotalen zur Totalen, und *narrowing* im umgekehrten Fall. Nicht-zusammenhängende Räume (*non-contiguous*) sind entweder weit voneinander entfernt (*distant*) oder der eine ist – obwohl vom anderen entfernt – eine Fortsetzung bzw. Verlängerung des anderen (*prolongation*) wie z. B. bei der Darstellung einer Reise. Eine Verbindung der Räume wird vom Beobachter erkannt, wenn er sieht, wie sich ein Darsteller von einem Ort zum anderen bewegt [Bateman, Schmidt 2011, S. 251].

Die Klassifikation *specific / generic* geht auf [Schmidt, Strauch 2002], [Schmidt 2008] zurück. Dort wird gezeigt, wie die beiden Modi der Symboltheorie von Zeichensystemen nach [Goodman 1969], die »Denotation« und die »Exemplifikation«, zur Klassifikation von Syntagmen angewandt werden können. Denotativ<sup>1</sup> verwendete Zeichen »passen als Etiketten auf – alles was sie darstellen, benennen oder beschreiben« [Goodman 1984, S. 127]. Der Spezialfall der piktoralen Denotation umfasst eine »Abbildung oder Repräsentation, und zwar durch eine Zeichnung, ein Gemälde, eine Skulptur, eine Fotografie, einen Film und so weiter« [Goodman 1987, S. 89]. Das heißt, ein Bild, das einen Gegenstand darstellt, denotiert ihn. Umgekehrt bedeutet Exemplifikation<sup>2</sup> »Bezugnahme durch einen Einzelfall einer Probe auf ein Etikett, das ihn denotiert« [Goodman, Elgin 1989, S. 35]. Das heißt, bei einer Exemplifikation handelt es sich um eine Bezugnahme zwischen Symbol und Eigenschaft. Ein Beispiel für eine Exemplifikation ist eine Stoffprobe oder ein Farbmuster in einem Musterbuch, die die Textur bzw. die Farbe eines Stoffes exemplifiziert. Im Netzwerk ist die Denotation als *specific* und die Exemplifikation als *generic* eingefügt.

<sup>1</sup> *denotare* (lat.): bezeichnen

<sup>2</sup> *exemplum* (lat.): Erläuterung, Beweis durch Beispiele

Abb. 2.20 zeigt die Gesamtsicht auf das Netzwerk der »Grande Paradigmatique« [Bateman, Schmidt 2011, S. 178 bzw. S. 330].

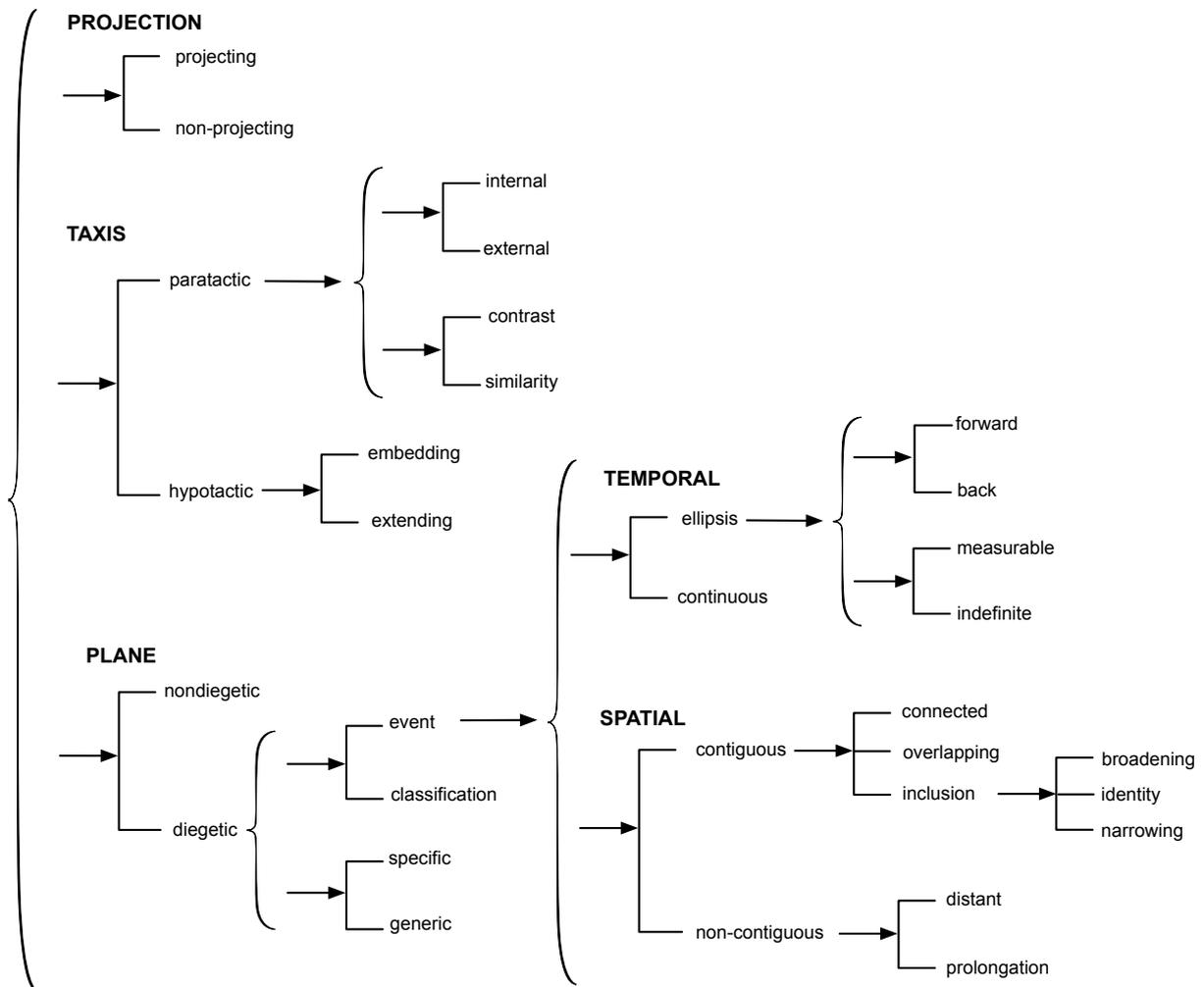


Abb. 2.20: Die »Grande Paradigmatique« nach Bateman, Schmidt

## 2.4 Text-Bild-Relationen

Die bisher vorgestellten analytischen Mittel beziehen sich primär auf die bildliche Darstellung im Film. Nicht betrachtet werden hierbei Relationen zwischen Bild und Text. Wie jedoch [Taboada, Habel 2013, S. 66] anmerken, sind auch diese Relationen Kohärenzrelationen und tragen damit zur Kohärenz eines Dokumentes und somit auch zu dessen Verständnis bei:

*»The rhetorical relations between figures and text can be understood as coherence links, contributing to the perceived coherence of a document.«*

Da kaum Forschungen zu Relationen zwischen Text und Filmbild vorliegen, wird hier auch auf sonstige Forschungsergebnisse zu Text-Bild-Relationen zurückgegriffen, die nicht Filmbilder in den Fokus stellen. Dies ist hier auch deshalb mit Gewinn möglich, da in dieser Arbeit kein künstlerischer Film analysiert wird, sondern ein Video, bei dem es sich eher um einen »Gebrauchstext« handelt, für den eine (angenähert) äquivalente Information sich auch durch eine geeignete Bilderfolge (s. 4.6) repräsentieren ließe, zu der der transkribierte Text ergänzend präsentiert wird.

In der Literatur finden sich zwei Taxonomien zu logisch-semantischen Relationen zwischen Text und Bild, die beide auf Ergebnissen der systemisch-funktionalen Linguistik und Semiotik ([Halliday, Matthiessen 2014], [Kress, van Leeuwen 1996]) basieren. Dies sind die Arbeiten von [Martinec, Salway 2005] und [Kong 2006]. Als Basis für die weitere Analyse wird die Taxonomie von [Kong 2006] verwendet, da sie etwas detaillierter und begrifflich präziser als die von [Martinec, Salway 2005] ist. Zudem wurde sie auch explizit zur Unterstützung empirischer Analysen entwickelt [Bateman 2014, S. 200]:

*»His principal aim with the account is also to support empirical analysis – that is, he offers the framework as a multi-layer annotation scheme which can be applied to collections of naturally occurring data in order to search for patterns and regularities; [...].«*

Abb. 2.21 zeigt die Taxonomie der logisch-semantischen Relationen nach [Kong 2006, S. 213]. Nach Kong werden die logisch-semantischen Relationen in drei Kategorien eingeteilt: *expansion*, *projection*<sup>1</sup> und *decoration*. Im Folgenden werden die Relationstypen in der Reihenfolge dieser Kategorien erläutert.

<sup>1</sup> *Expansion* und *projection* sind die beiden Haupttypen der logisch-semantischen Relationen von Halliday.

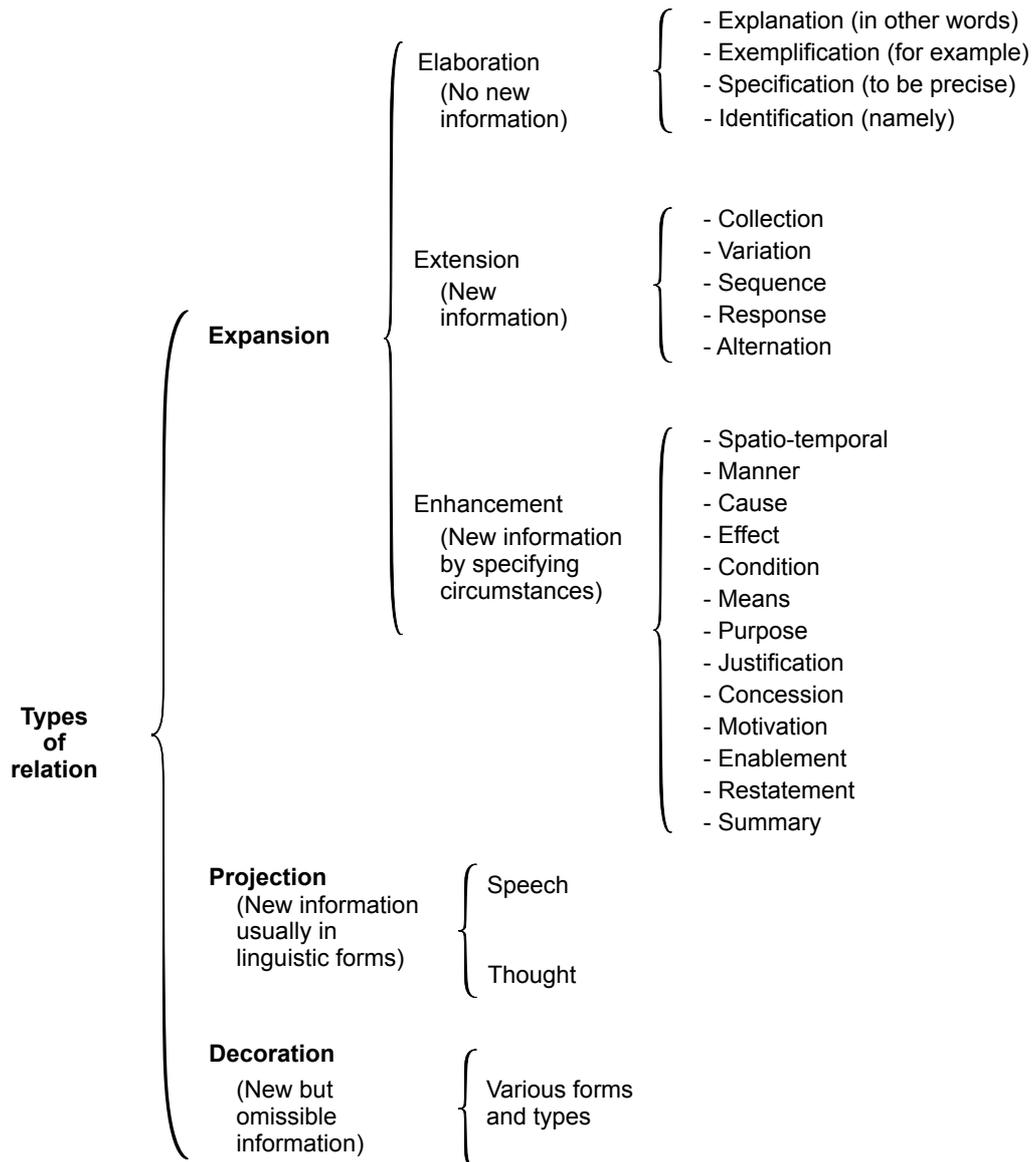


Abb. 2.21: Taxonomie der logisch-semantischen Relationen nach Kong

Der Relationstyp *expansion* liegt vor, wenn ein Modus den anderen Modus erweitert. Dieser Relationstyp hat folgende Subtypen:

- *Elaboration*

In diesem Fall liefert ein Modus eine detaillierte, nicht unbedingt neue Beschreibung des anderen Modus. [Halliday, Matthiessen 2014, S. 461] präzisiert dies wie folgt: »*The secondary clause does not introduce a new element in the picture but rather provides a further characterization of one that is already there, restating it, clarifying it, or adding a descriptive attribute or comment.*«

Der Relationstyp *elaboration* hat vier Subtypen: *explanation*, *exemplification*, *specification* und *identification*.

Der Relationstyp *explanation* liegt vor, wenn beide Modi die gleiche Information präsentieren, jedoch auf unterschiedliche Art oder Präsentationsmethode (»mit anderen Wor-

ten«). In einem Film liegt dieser Relationstyp folglich immer dann vor, wenn der Text, genau die im Filmbild gezeigten Handlungen beschreibt, jedoch keine zusätzlichen Informationen oder direkte Verweise auf das Bild enthält.

Im Falle des Relationstyps *exemplification* liefert der erweiternde Modus ein spezifischeres Beispiel.

Der Relationstyp *specification* ist äquivalent zum Ausdruck »um präzise zu sein«. In der Filmanalyse kann eine Relation als *specification* klassifiziert werden, wenn der Text eine sehr detaillierte Beschreibung des Geschehens liefert, z. B. durch genaue numerische Angaben.

Der Relationstyp *identification* liegt vor, wenn ein Wort oder ein Text ein spezielles Bild identifiziert oder der umgekehrte Fall vorliegt. Dieser Relationstyp ist sehr häufig in wissenschaftlicher Literatur, Bedienungsanleitungen oder Reiseführern anzutreffen.

- *Extension*

Beim Relationstyp *extension* erweitert ein Modus die Information des anderen Modus. [Kong 2006, S. 212] formuliert hierzu folgenden Vergleich: »*Like an extension of a building, a unit with new meaning can be added to the original unit.*« Es existieren die Subtypen *collection*, *variation*, *sequence*, *response* und *alternation*.

Beim Relationstyp *collection*<sup>1</sup> enthält der erweiterte Modus ein zusätzliches Element, ohne dass eine kausale oder temporale Beziehung zwischen den Modi vorliegt. In der Sprache entspricht dies einer »und«-Verknüpfung.

Der Relationstyp *variation* liegt vor, wenn eine Einheit als vollständiger oder teilweiser Ersatz der anderen Einheit betrachtet werden kann. Dies ist im Sinne von »Vergleich und Gegenüberstellung« gemeint.<sup>2</sup>

Der Relationstyp *sequence* liegt vor, wenn beide Modi, Bild und Text, Ideen in chronologischer Reihenfolge übermitteln, ohne dass eine Einheit der anderen untergeordnet ist. Ein Beispiel hierzu sind Bildergeschichten oder Comicstrips, bei denen ein Panel auch nur Text oder nur ein Bild enthalten kann.

Sind zwei Einheiten durch den Relationstyp *alternation* miteinander verbunden, kann die eine die andere vollständig ersetzen, ohne dass Bedeutung verloren geht.

Der Relationstyp *response* tritt auf, wenn Einheiten als Frage-Antwort-Kombination miteinander verbunden sind; eine »*response*«-Einheit enthält eigenständige Informationen.

- *Enhancement*

Ein Modus reichert die Information des anderen Modus an, indem zusätzliche, ausführli-

<sup>1</sup> [Halliday 2013, S. 471] nennt diesen Relationstyp *addition*.

<sup>2</sup> Im Original [Kong 2006, S. 212]: »*Which is similar to the function of 'comparing and contrasting'.*«

chere Informationen referenziert werden<sup>1</sup>, z. B. zu einem Ort (*spatial*), einer Zeit (*temporal*), einem Zweck (*purpose*), einer Bedingung (*condition*), einem Ziel (*goal*) oder andere zirkumstanzielle Elemente.

Der Relationstyp *projection* wurde bereits in 2.3 erläutert. Da dieser Typ im zu untersuchenden Video nicht auftritt, wird an dieser Stelle nicht weiter darauf eingegangen.

Der Relationstyp *decoration* ist nach [Kong 2006, S. 212] »*more diffuse and less constrained by convention and more subject to interpretation.*« Elemente, die als »in der Beziehung *decoration* stehend« klassifiziert werden, haben wenig oder keinen Informationswert. Ihre Aufgabe besteht darin, die Aufmerksamkeit der Leser bzw. Zuschauer zu wecken, oder sie werden aus ästhetischen Gründen hinzugefügt [Kong 2006, S. 218]. Weiterhin führt Kong aus, dass bei einer Kombination von Bild und Text meistens das Bild eine dekorative Funktion für den Text hat und der umgekehrte Fall eher unwahrscheinlich ist.

### **Kritik**

Fraglich ist, ob die Abgrenzung der Relationstypen voneinander wirklich trennscharf ist. Ein Beispiel: Eine Relation wird als *expansion – elaboration – specification* klassifiziert, wenn der Text eine sehr detaillierte Beschreibung des Geschehens liefert, z. B. durch genaue numerische Angaben. Hier kann man jedoch diskutieren, ob diese numerischen Angaben eine Zusatzinformation darstellen und die Relation demzufolge als *expansion – extension – collection* einzuordnen wäre. Es ist also nicht auszuschließen, dass verschiedene Analysten zu verschiedenen Ergebnissen kommen.<sup>2</sup>

### **Anwendung von Text-Bild-Relationen auf filmischen Text**

Bei der Analyse von Text-Bild-Relationen beim Medium Film/Video sollten noch folgende Aspekte beachtet werden:

Nach [Taboada, Habel 2013, S. 81f] können in einem multimodalen Dokument mit statischem Text und Bild mehrere Relationen zwischen Text und Bild vorliegen. Auch wenn der Text linear gelesen wird, wird der Lesevorgang häufig durch Blickwechsel zu den Bildern unterbrochen, oft werden Beschreibungen und Bilder mehrmals gelesen bzw. betrachtet. Von daher kann davon ausgegangen werden, dass je nach Beobachtungssituation auch andere Relationen angenommen werden können. Im Unterschied zu einer Lesesituation, die »Zurückblättern« erlaubt, kann ein Zuschauer einen Film in vielen Beobachtungssituationen nur »linear« erfassen<sup>3</sup>, wobei der gesprochene Text und das filmische Bild häufig parallel präsentiert werden. Bei der Modellierung

---

<sup>1</sup> Im Original [Kong 2006, S. 212]: »*to expand a main idea unit by specifying circumstances.*«

<sup>2</sup> Vergleiche hierzu auch die Ausführungen von [Mann, Thompson 1987, S. 28f] zur Problematik der »*ambiguity*« bei einer RST-Analyse.

<sup>3</sup> Es wird davon ausgegangen, dass der Zuschauer keine Möglichkeit hat, einzelne Teile eines Films wiederholt zu betrachten.

der Beobachter sollte man hier minimal davon ausgehen, dass der Zuschauer überhaupt eine Relation konzeptualisieren kann. Bei einer Analyse hat der Ausführende dagegen die Möglichkeit, Filmabschnitte wiederholt anzusehen und den transkribierten Text zu lesen, wobei er typischerweise die Situation des »normalen« Zuschauers nicht aufrechterhalten kann.

Zudem muss man fragen, ob in einem Film, in dem verschiedene Medien zur gleichen Zeit präsentiert werden, diese auf einen Zuschauer in gleicher Weise wirken. Dazu führt [Scheffel 2009, S. 19] aus:

*»Anders als im Fall des Dramas, an dessen Ursprung der Figurendialog steht, ist im Fall des Films eben nicht der Ton, d. h. das gesprochene oder in Schrift notierte Wort oder auch die Musik, sondern das qua Bewegung per definitionem irgendeine Art von Geschehen darstellende Bild als Ursprung und Leitmedium anzusehen.«*

[Scheffel 2009, S. 19] folgert, dass sich die Bedeutung der verschiedenen Medien hierarchisieren lässt:

*»In der Regel gilt, dass der bildepische Faktor dem gesprochenen Wort übergeordnet ist.«*

Da dies nicht mehr als eine plausible Annahme ist, wird hier mit empirischer Vorsicht die Analyse der Text-Bild-Relationen des Beispielvideos in 2.5.4 lediglich »in einer Richtung« durchgeführt, d. h. es wird untersucht, welche Beziehung bzw. Funktion ein Textabschnitt zum parallel gezeigten filmischen Bild hat. Zudem wird davon ausgegangen, dass sich das empirisch behandelte Material an einen Standardzuschauer richtet.

## 2.5 Analyse des Videos »Snowboard«

Die empirische Anwendung der gesamten, in dieser Arbeit präsentierten Vorgehensweise wird eingegrenzt auf Gebrauchsfilm, speziell solche, die die Herstellung eines technischen Produktes zeigen. Beispielführend durchgeföhrt werden alle Analysen und auch die »Experimente« in Kapitel 4 anhand eines Videos, das den Herstellungsprozess eines Snowboards zeigt.<sup>1</sup> Das Video gehört zu einer Reihe von Sachgeschichten der »Sendung mit der Maus« des Westdeutschen Rundfunks (WDR). Es wurde am 16.12.2007 erstmals ausgestrahlt und hat eine Länge von 7:39 Minuten. Zielgruppe dieser Sachgeschichten sind Kinder im Vorschul- und Grundschulalter. Viele der Sachgeschichten wurden nach demselben dramaturgischen Muster produziert, wie [Stötzel 1990, S. 135f] feststellt:

*»Klassisches Element vieler Sachgeschichten sind Zwischenschnitte, die den fortschreitenden Produktionsprozess am Produkt im Standbild oder in neutraler Umgebung festhalten [...]. Zudem markieren diese Zwischenschnitte Übergänge zu einzelnen, wengleich nicht immer unterschiedlichen Produktionsphasen.«*

In Beispielvideo »Snowboard« bilden die Zwischenschnitte einen zweiten Handlungsstrang, in dem gezeigt wird, wie ein Akteur, ein junges Mädchen namens »Johanna«, das Snowboard im jeweils vorliegenden Zustand testet. Zwischen den beiden Handlungssträngen »Bau des Snowboards« und »Test durch Johanna« wird regelmäßig hin und her gewechselt. Das gesamte Video wird von einem männlichen Sprecher aus dem OFF kommentiert.

### 2.5.1 Vorgehensweise

[Bateman, Schmidt 2011, S. 252] betonen, dass die paradigmatische Analyse eine führende Rolle (»guiding role«) spielt. Würde mit der syntagmatischen Analyse begonnen, ergäbe sich eine sehr große Anzahl von Kombinationen von Einstellungen, die auf Erfüllung der Bedingungen der verschiedenen Syntagmen getestet werden müssten. Wird jedoch zuerst die paradigmatische Organisation der Einstellungen untersucht, entsteht dadurch eine Priorisierung. Insbesondere die Ergebnisse zur räumlichen und zeitlichen Beziehung zweier Einstellungen liefern wichtige Informationen, welches der Syntagmen vorliegen könnte. Die Anzahl der zu untersuchenden Kombinationen wird somit deutlich reduziert. Bei der folgenden syntagmatischen Analyse kann es passieren, dass einzelne Werte der paradigmatischen Relationen korrigiert werden müssen und dadurch evtl. auch die syntagmatische Analyse. Es entsteht so ein zyklischer Prozess, der beendet ist, wenn der komplette Film paradigmatisch und syntagmatisch betrachtet wurde [Bateman, Schmidt 2011, S. 245]. Wie häufig dieser Zyklus durchlaufen werden muss, hängt auch von der Komplexität und der »narrativen Tiefe« des Films ab.

---

<sup>1</sup> [http://www.youtube.com/watch?v=xDu\\_-CO6eCE](http://www.youtube.com/watch?v=xDu_-CO6eCE) [letzter Aufruf: 04.03.2015]

Der Empfehlung von Bateman folgend wird nun zuerst – Einstellung für Einstellung – die paradigmatische Analyse und anschließend die syntagmatische Analyse durchgeführt.

Diese Analysen werden zudem unter der Maßgabe der maschinellen Umsetzbarkeit durchgeführt. So finden sich die hier eingeführten Nummerierungen und Bezeichner in den entsprechenden XML-Dokumenten wieder. Umgekehrt werden in Kapitel 4 auch »maschinelle« Hilfsmittel« zur Unterstützung der Analysen entwickelt. Tab. 2.3 ist ein erstes Hilfsmittel: die Wechseltabelle, wie in 2.2.4 vorgestellt.<sup>1</sup> In dieser ist erkennbar, dass zwei diegetische Hauptorte, genannt »Berg« und »Werkstatt«, existieren und für den Bereich »Werkstatt« eine weitere Unterteilung in 14 Teilbereiche, »S1« bis »S14«, vorliegt.

**Wechseltabelle: Snowboard**

Berg	Werkstatt													
	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14
T01														
T02														
T03														
T04														
T05														
T06														
T07														
T08														
T09														
T10														
T11														
T12														
T13														
T14														
T15														
T16														
T17														
T18														
T19														
T20														
T21														
T22														
T23														
T24														
T25														
T26														
T27														
T28														
T29														
T30														
T31														
T32														
T33														
T34														
T35														
T36														
T37														
T38														
T39														
T40														
T41														
T42														

Tab. 2.3: Wechseltabelle zum Video »Snowboard«

<sup>1</sup> Die Wechseltabelle wurde auf Basis der Transkriptionsanalyse und der Darstellung der einzelnen Einstellungen in XML (s. 4.2.1) durch ein XSLT-Stylesheet (s. 4.4.3) generiert.

### 2.5.2 Paradigmatische Analyse

Die paradigmatische Analyse eines Films besteht in der Klassifikation der Übergänge (*transitions*) zwischen den Einstellungen. Genauer gesagt, wird für die Übergänge zwischen den Einstellungen untersucht, welche der Optionen im Netzwerk der »Grande Paradigmatique« (s. Abb. 2.20) für den jeweils betrachteten Übergang gültig sind. Hierbei werden nicht nur Übergänge zwischen aufeinanderfolgenden Einstellungen betrachtet, sondern, insbesondere bei Ortswechseln, auch Übergänge zu/von weiter zurückliegenden Einstellungen, die den gleichen Ort zeigen.

In den Beispielen in [Bateman 2007] und [Bateman, Schmidt 2011] wird die Analyse primär auf Basis der bildlichen Darstellung durchgeführt. Das hier untersuchte Video ist vollständig mit Sprechertext unterlegt. Dieser erläutert die gezeigten Handlungen und liefert zudem Informationen, z. B. über Ort und Zeit<sup>1</sup>, die allein aus der bildlichen Darstellung nicht ersichtlich sind. Diese additiven Informationen werden bei der Analyse ggf. berücksichtigt.

#### Notation

Eine »Kurzfassung« der paradigmatischen Relationen wird, ähnlich zu [Bateman, Schmidt 2011] angegeben. Es wird jeweils der maximal mögliche Weg durch das Netzwerk der »Grande Paradigmatique« notiert. Die »Werte« der einzelnen Dimensionen werden durch Semikolon getrennt und umgeben von eckigen Klammern. Die Reihenfolge entspricht der Reihenfolge im Netzwerk in der Leserichtung von oben nach unten: [Projection; Taxis; Plane]. Zur besseren Lesbarkeit wird nach dem Semikolon ein Zeilenumbruch eingefügt. Detaillierungen innerhalb der Dimensionen werden durch Bindestriche getrennt, simultane Auswählmöglichkeiten untereinander angeordnet.<sup>2</sup> Die paradigmatischen Relationen werden in der Reihenfolge ihres Auftretens durchnummeriert ([Px]); so sind Verweise auf bereits definierte Relationen leicht möglich.

#### Hilfsmittel

Zur Klassifikation der Beziehungen von genau zwei Einstellungen zueinander ist es hilfreich, wenn die zu analysierenden Einstellungen isoliert betrachtet werden können. Zu diesem Zweck wurden folgende Hilfsmittel entwickelt (s. 4.4.4):

- (1) ein XSLT-Stylesheet, das ein XML-Dokument als »Vorlage« erzeugt. In dieser Vorlage sind alle zu analysierenden Übergänge eingetragen, es müssen lediglich die Relationsnummern aus der Datei `paradigmatic_relation.xml` ergänzt werden<sup>3</sup>,
- (2) ein PDF-Dokument, in dem die Screenshots zweier Einstellungen jeweils in einer Zeile dargestellt werden, sowie SMIL- und HTML5-Dokumente, bei denen jeweils zwei Einstellungen in Folge abgespielt werden.

---

<sup>1</sup> Beispiele: T08: »Was passiert in der grünen Kiste?« T25: »und nach 15 bis 20 Minuten ...«

<sup>2</sup> Falls dies layout-technisch nicht möglich ist, wird dies durch einen »|« visualisiert.

<sup>3</sup> Bei einer ersten Analyse (wie in diesem Fall) wird diese Datei parallel erstellt.

Eine Kontrolle der Analyse ist leicht möglich, wenn die XSLT-Transformation erneut durchgeführt wird, da nun die bereits erfassten Ergebnisse ebenfalls angezeigt werden. Abb. 2.22 zeigt die Darstellung eines SMIL-Dokuments, Abb. 4.8 einen Screenshot des PDF-Dokuments.

**Paradigmatische Relationen - Video: Snowboard**

**T32 Ort: Werkstatt /S10 (5:54 - 5:55.8)**



**T33 Ort: Werkstatt /S10 (5:56 - 5:59.5)**



Relation	P1
PPROJECTION	hypotactic - extending
TAXIS	hypotactic - extending
PLANE	diegetic - event ellipsis - forward   measurable contiguous - inclusion - narrowing specific

Abb. 2.22: SMIL-Video zur Darstellung von paradigmatischen Relationen

## Analyse

Allgemeine Aussagen sind für die Dimensionen PROJECTION und PLANE möglich.

Das Snowboard-Video ist ein Gebrauchstext. Die Darsteller spielen im Video eine untergeordnete Rolle. Die Akteure in der Werkstatt führen die Handlungsschritte zur Produktion des Snowboards vor. Johanna zeigt, ob ein Fahren mit dem teilweise fertiggestellten Snowboard möglich ist. In keiner Einstellung werden mentale Zustände oder mentale Ereignisse eines Darstellers gezeigt. Bezüglich der Dimension PROJECTION sind daher alle Relationen als non-projecting zu klassifizieren.

Auf der Bildebene gibt es keine nicht-diegetischen Einfügungen, zudem werden im gesamten Video Handlungen dargestellt.<sup>1</sup> Bezüglich der Dimension PLANE spricht daher nichts dagegen, alle Einstellungen als diegetic - event zu betrachten.

Da es sich bei allen Bildern für die intendierte Beobachtermenge um eine Denotation der gezeigten Welt handelt, können alle Relationen als specific klassifiziert werden. Das Video zeigt den

<sup>1</sup> Im Unterschied dazu ist der Ton (bestehend aus dem Sprechertext) sowie der Musikuntermalung (bei den Szenen mit Johanna auf dem Berg) nicht-diegetisch. Eine Ausnahme bilden die Einstellungen in der Werkstatt, in denen Maschinengeräusche zu hören sind.

chronologischen Ablauf der Produktion und der Tests des Zwischenproduktes. Dies bedeutet: Wenn zwischen zwei Einstellungen eine zeitliche Lücke existiert, ist diese jeweils als *forward* zu klassifizieren.

Die zu untersuchenden Einstellungsübergänge werden nun in der Reihenfolge analysiert, in der sie auch in der erstellten Vorlage angeordnet sind: Als Erstes werden die paradigmatischen Relationen aller aufeinanderfolgenden (benachbarten) Einstellungen analysiert, dann folgt die Analyse der nicht-benachbarten Einstellungen des gleichen Hauptortes und abschließend die der nicht-benachbarten Einstellungen, die in der Werkstatt spielen und bei denen ein Wechsel der Arbeitsstation vorliegt.

### Analyse der Relationen zwischen aufeinanderfolgenden Einstellungen

#### T01 – T02

Beide Einstellungen zeigen den gleichen Ort, den Berg, auf dem sich Johanna befindet. Die Kameraeinstellung wechselt von der Totalen zur Nahaufnahme<sup>1</sup>, daher ist die räumliche Beziehung *contiguous – inclusion – narrowing*. Da nicht zu sehen ist, wie Johanna das Snowboard erhält, muss zwischen den Einstellungen eine zeitliche Lücke existieren. Die zeitliche Lücke wird als *measurable* klassifiziert, da für einen Beobachter eine kleine zeitliche Dauer schätzbar ist. Da T02 als Fortsetzung von T01 angesehen werden kann, ist die Relation als *hypotactic – extending* zu klassifizieren. Es ergibt sich:

[P1] [non-projecting;  
hypotactic – extending;  
diegetic – event – ellipsis – forward  
– measurable  
– contiguous – inclusion – narrowing  
– specific]

#### T02 – T03

Beide Einstellungen zeigen Johanna auf dem Berg. Es findet ein Wechsel von der Nahaufnahme zur Totalen statt, daher liegt für einen Beobachter die räumliche Beziehung *contiguous – inclusion – broadening* vor. Zwischen den Einstellungen ist keine zeitliche Lücke zu erkennen. Auch dieser Übergang ist *hypotactic – extending*. Es liegt als Relation [P2] vor:

[P2] [non-projecting;  
hypotactic – extending;  
diegetic – event – continuous  
– contiguous – inclusion – broadening  
– specific]

#### T03 – T04

T04 ist die erste Einstellung, die die Werkstatt (Arbeitsstation 1) zeigt. Zwischen T03 und T04 existiert keine visuelle Überlappung. Die räumliche Beziehung ist daher als *non-contiguous –*

<sup>1</sup> Für die Einstellungsgrößen wird die Nomenklatur aus [Hickethier 2007, S. 55ff] verwendet.

distant zu klassifizieren. Eine zeitliche Lücke unbekannter Dauer wird diegetisch nahegelegt, daher ist die zeitliche Relation mit *indefinite* zu klassifizieren. In der Dimension *TAXIS* ist der Übergang als *paratactic* zu klassifizieren, da die Einstellungen in keiner Abhängigkeitsbeziehung zueinander stehen, und als *external*, da in beiden Einstellungen unabhängige Handlungen repräsentiert werden. Ob die parallele Sub-Dimension als *similarity* oder *contrast* zu klassifizieren ist, kann an dieser Stelle ein Beobachter noch nicht entscheiden. Hierzu muss er erst weitere Einstellungen betrachten. Es wird somit die Relation [P3] definiert, die an einer Position noch ein »?« für »undetermined« enthält:

```
[P3] [non-projecting;
      paratactic - external
        - ? ;
      diegetic - event - ellipsis - forward
                - indefinite
                - non-contiguous - distant
      - specific]
```

#### **T04 – T05**

In *T04* bewegt sich der Akteur vom Vorratsstapel zu einer Fräse, es folgt ein Schnitt und in *T05* wird der Akteur in Nahaufnahme gezeigt. Die diegetischen Räume der Einstellungen sind verbunden mit einer visuellen Überlappung (Kameraschwenk). Es findet ein Einstellungswechsel von »HN« zu »N« statt. Dieser hinterlässt beim Beobachter einen größeren Eindruck als der Kameraschwenk, so dass die räumliche Beziehung als *contiguous - inclusion - narrowing* klassifiziert werden kann. Die Einstellungen sind zeitlich zusammenhängend. Es liegt als Relation [P4] vor:

```
[P4] [non-projecting;
      hypotactic - extending;
      diegetic - event - continuous
                - contiguous - inclusion - narrowing
      - specific]
```

#### **T05 – T06**

Beide Einstellungen zeigen einen Akteur an derselben Maschine (Arbeitsstation 1). Da ein Einstellungswechsel von »N« zu »G« stattfindet, ist die räumliche Beziehung als *contiguous - inclusion - narrowing* zu klassifizieren. Es keine zeitliche Lücke zu erkennen, so dass auch hier die Relation [P4] vorliegt.

#### **T06 – T07**

In *T06* befindet sich der Akteur am »rechten« Ende der Maschine, in *T07* bewegt er sich zum »linken« Ende, es findet ein Einstellungswechsel von »G« zu »HT« statt. Hier kann man jedoch den Eindruck haben, dass die räumliche Veränderung stärker auf die Beobachter wirkt als der Einstellungswechsel. Daher wird hier die räumliche Relation als *contiguous - overlapping* klassifiziert. Eine zeitliche Lücke ist offenbar nicht vorhanden. Hier liegt also als Relation [P5] vor:

[P5] [non-projecting;  
hypotactic – extending;  
diegetic – event – continuous  
– contiguous – overlapping  
– specific]

### T07 – T08

Abb. 2.23 zeigt die Einstellungen T07 – T08. In Einstellung T07 wird eine Fräse gezeigt, die die Dicke der Bretter verändert. Die Kameraeinstellung ist »HT«. In T08 wird dieser Vorgang erneut gezeigt, jetzt allerdings in Großaufnahme. Zudem ist das Innere dieser Maschine zu sehen, vermutlich wurde vorab die Abdeckung entfernt.

T07	T08
 <p>Und danach in so 'ne grüne Kiste. Vorne rein und hinten wieder raus. Und immer drei Bretter gehören zusammen. Und der Strich bedeutet nicht, dass da abgeschnitten werden soll, sondern: da ist später hinten.</p>	 <p>Was passiert in der grünen Kiste? Da laufen die Bretter unter einer Fräse durch. Das ist das runde Ding, was sich da dreht. Und was macht diese Fräse? Hier kann man's sehen: die macht das Brett dünner an manchen Stellen, ...</p>

Abb. 2.23: Einstellungen T07 – T08

Eine ähnliche Situation diskutieren [Schmidt, Strauch 2002] bei der Analyse von Segment 7 des Films »Adieu Pilippine«. Dort ist in der ersten Einstellung zu sehen, wie ein Wagen rückwärts anfährt und mit hoher Geschwindigkeit zurücksetzt. Eine scheinbare Wiederholung des Ereignisses wird nochmals in einer unbedeutend näheren Einstellung wiedergegeben.<sup>1</sup> [Schmidt, Strauch 2002, S. 75] unterscheiden bei der Klassifikation mehrere Fälle, wovon zwei im Beispielvideo »Snowboard« vorliegen könnten.

- *Fall 1: »In der zweiten Einstellung wird tatsächlich mit einer anderen Einstellungsgröße dasselbe Ereignis noch einmal dargestellt. In diesem Fall führt die Erzählung eine Irritation herbei, die die Aufmerksamkeit auf das Erzählen selbst lenkt. [...]«*
- *Fall 2: »In der zweiten Einstellung wird ein zeitlich nachfolgendes (Rückwärtsfahren) an demselben Ort dargestellt. Das Ganze ist dann eine Sequenz, weil das Vorfahren des Autos zwischen der ersten und zweiten Einstellung elliptisch ausgelassen wurde.«*

Im betrachteten Video kann der Vorgang nicht mit dem gleichen Brett wiederholt werden, da dieses ja verändert wurde. Eine Irritation des Zuschauers ist nicht der Fall, insbesondere auch deshalb nicht, da die Wiederholung durch den Sprechertext kommentiert wird. Fall 1 ist daher auszuschließen. Eher wahrscheinlich ist, dass folgende Situation vorliegt: Für ein Snowboard müssen drei der Bretter zusammengefügt werden, so dass es kein Problem ist, die Einstellung

<sup>1</sup> In [Schmidt, Strauch 2002] gibt es noch eine dritte Einstellung: Das Auto fährt wieder vorwärts.

T08 mit dem zweiten oder dritten Brett aus einer anderen Perspektive zu filmen. Diese Situation ist somit mit Fall 2 vergleichbar. Allerdings scheint zwischen T07 und T08 keine (oder nur eine minimale) zeitliche Lücke zu liegen (die Bretter werden automatisch eingezogen). Die Beziehung zwischen den Einstellungen wird daher mit der Relation [P4] klassifiziert.

### **T08 – T09**

Die Einstellung T09 zeigt die weitere Bearbeitung des Snowboards an einer anderen Arbeitsstation (Station 2). In der Dimension TAXIS könnte die Option *paratactic* gültig sein, da ein anderer Ort und eine andere Person in dieser Einstellung zu sehen sind. Durch den Sprechertext, die Umgebung (z. B. die Farben) und nicht zuletzt durch das verbindende Objekt »Snowboard« soll nahelegt werden, dass es sich hier um eine Fortsetzung des Erstellungsprozesses handelt. Die Relation sollte daher als *hypotactic – extending* klassifiziert werden. Die spatiale Relation ist als *non-contiguous – prolongation* zu klassifizieren, da eine Bewegung (des Snowboards) von einem Ort zu einem anderen stattfindet.<sup>1</sup> Zwischen den Einstellungen existiert eine zeitliche Lücke, deren Dauer ein Beobachter nicht schätzen kann, da er zum einen die genaue Raumsituation der Werkstatt nicht kennt, also nicht abschätzen kann, wie lange ein Akteur von einer Station zur nächsten braucht, zum anderen kennt er den exakten zeitlichen Ablauf nicht, weiß also nicht, ob und wie lange eventuelle Pausen zwischen den Arbeitsschritten dauern. Die temporale Relation wird daher als *ellipsis – indefinite* klassifiziert. Insgesamt liegt somit als »Itinerar-Relation« [P6] vor:

[P6] [non-projecting;  
hypotactic – extending;  
diegetic – event – ellipsis – forward  
– indefinite  
– non-contiguous – prolongation  
– specific]

### **T09 – T10**

Aus unterschiedlichen Perspektiven zeigen T09 und T10 einen Arbeitsschritt, das Zusammenkleben von drei Brettern an einer Arbeitsstation (Station 2). In T09 steht die Kamera auf der rechten Seite des Akteurs, die Einstellung endet mit einer Großaufnahme der Klebepistole. T10 zeigt den Akteur in der Einstellung »HN« von schräg vorne. Die räumliche Relation ist *contiguous – overlapping*. Es gibt keine zeitliche Lücke. Es liegt daher Relation [P5] vor.

### **T10 – T11 – T12**

Es findet ein Wechsel von der Werkstatt zum Berg und zurück statt. Der diegetische Ort von Einstellung T12 ist Arbeitsstation 3 in der Werkstatt. Bei beiden Übergängen liegt die gleiche paradigmatische Relation wie beim Übergang T03 – T04 (Wechsel Berg – Werkstatt), die Relation

<sup>1</sup> »Prolongation because of the character's movement which takes us from the old location to the new.« [Bateman, Schmidt 2011, S. 251]

[P3], vor. Es wurden nun genügend Einstellungen betrachtet, so dass klar ist, dass zwischen den Einstellungen  $T03 - T04$ ,  $T10 - T11$  und  $T11 - T12$  jeweils eine Relation vorliegt, die in sich kontrastierend ist. Der noch fehlende Wert in der Relation [P3] ist also contrast. Insgesamt gilt somit für die »Wechsel-Relation« [P3]:

[P3] [non-projecting;  
paratactic - external  
- contrast;  
diegetic - event - ellipsis - forward  
- indefinite  
- non-contiguous - distant  
- specific]

### **$T12 - T13$**

<b><math>T12</math></b>	<b><math>T13</math></b>
 <p>Also kommt das Ding jetzt unter diese Maschine. Da wird das Brett eingeklemmt. Und was passiert da?</p>	 <p>Da laufen dann wieder Fräsen - nicht nur eine - sondern mehr, außen am Rand entlang und geben dem Ding die Form: vorne und hinten rund und an der Seite etwas schmaler.</p>

Abb. 2.24: Einstellungen  $T12 - T13$

Die Einstellungen  $T12$  und  $T13$  (s. Abb. 2.24) zeigen – auf den ersten Blick – eine analoge Situation zu der in  $T07 - T08$ . In  $T12$  ist ein Akteur zu sehen, der ein Brett in die Fräse legt. Die Kameraeinstellung ist »N«. In Einstellung  $T13$  wird nun im Detail, mit Kameraeinstellung »D«, gezeigt, wie die Fräse arbeitet. Im Unterschied zu den Einstellungen  $T07 - T08$  ist hier jedoch keine Wiederholung, sondern eine Fortsetzung des Prozessschrittes anzunehmen. Es liegt also Relation [P4] vor.

### **$T13 - T14$**

Bei diesen Einstellungen findet ein Wechsel zur nächsten Arbeitsstation (Station 4) statt. Der Arbeitsprozess wird fortgesetzt. Zwischen den Einstellungen ist eine zeitliche Lücke anzunehmen. Für diesen Übergang gilt daher die »Itinerar-Relation« [P6].

### **$T14 - T15$**

Beide Einstellungen zeigen die gleiche Arbeitsstation (Station 4). Die Kameraeinstellung wechselt von »D« zu »HT«. Es ist keine zeitliche Lücke anzunehmen. Die Situation ist analog zu der bei  $T02 - T03$ . Es liegt Relation [P2] vor.

**T15 – T16 – T17**

Hier findet wieder ein Wechsel von der Werkstatt zum Berg und zurück statt. Der diegetische Ort der Einstellung T17 ist Arbeitsstation 5. Bei beiden Übergängen liegt die »Wechsel-Relation« [P3] vor.

**T17 – T18 – T19 – T20 – T21**

Jede dieser Einstellungen spielt an einer weiteren Arbeitsstation, so dass Wechsel von Arbeitsstation 5 bis zu Arbeitsstation 9 vorliegen. Der Arbeitsprozess wird fortgesetzt. Zwischen den Einstellungen liegen jeweils zeitliche Lücken vor. Für alle Übergänge gilt daher die »Itinerar-Relation« [P6].

**T21 – T22 – T23 – T24 – T25 – T26 – T27 – T28 – T29 – T30 – T31**

In diesen Einstellungen werden verschiedene Arbeitsschritte, die alle an einer Arbeitsstation (Station 9) stattfinden, gezeigt. Im Video dauert diese Folge von 4:08 bis 5:46 Minuten, also insgesamt 1:38 Minuten und wirkt sehr langatmig. Zwischen den Einstellungen liegen minimale zeitliche Lücken vor. Vermutlich wurde die Folge am Stück gedreht und anschließend wurden kleinere Bildfolgen (Randbereiche) herausgeschnitten, um die Sequenz zu kürzen. Die zeitliche Beziehung zwischen den Einstellungen kann daher mit *ellipsis – forward – measurable* klassifiziert werden. Die Einstellungsgröße variiert zwischen »N« und »G« (s. Tab. 2.4). Bei einem Einstellungswechsel von »N« zu »G« gilt für die spatiale Relation *contiguous – inclusion – narrowing*. In diesem Fall liegt somit Relation [P1] vor. Für die übrigen Fälle müssen noch zwei paradigmatische Relationen definiert werden, eine, [P7], für den Fall, dass Ort und Einstellungsgröße gleichbleiben, die spatiale Relation also *contiguous – inclusion – identity* ist, und eine, [P8], mit der spatiale Relation *contiguous – inclusion – broadening* für das »Wegzoomen«:

[P7] [non-projecting;  
hypotactic – extending;  
diegetic – event – ellipsis – forward  
– measurable  
– contiguous – inclusion – identity  
– specific]

[P8] [non-projecting;  
hypotactic – extending;  
diegetic – event – ellipsis – forward  
– measurable  
– contiguous – inclusion – broadening  
– specific]

Tab. 2.4 zeigt die Einstellungsgrößen und die paradigmatischen Relationen der Einstellungen an Station 9 in einer Übersicht.

Ein- stellung	T21		T22		T23		T24		T25		T26		T27		T28		T29		T30		T31
Ein- stellungs- größe	N		N		N		G		N		G		G		N		N		N		N
Relation		[P7]		[P7]		[P1]		[P8]		[P1]		[P7]		[P8]		[P7]		[P7]		[P7]	

Tab. 2.4: Einstellungsgrößen und paradigmatische Relationen der Einstellungen T21 – T31

### T31 – T32

Von T31 nach T32 findet ein Ortswechsel von Arbeitsstation 9 zu 10 statt und es existiert eine zeitliche Lücke, daher liegt die »Itinerar-Relation« [P6] vor.

### T32 – T33 – T34

Alle Einstellungen zeigen Arbeitsstation 10. Bei jeder dieser Einstellungen wird die Maschine näher herangezoomt. Die räumliche Relation ist daher jeweils contiguous – inclusion – narrowing. Es existieren kleine zeitliche Lücken, deren Dauer schätzbar scheint. Es gilt daher die Relation [P1].

### T34 – T35

T34 und T35 zeigen Arbeitsschritte an einer Arbeitsstation, der Station 10. Der Sprechertext von T35 beginnt mit den Worten: »Und nach etwa 15 bis 20 Minuten ....«. Hierdurch ist klar, dass eine diegetische zeitliche Lücke bekannter Dauer zwischen T34 und T35 existiert. Es liegt als paradigmatische Relation [P9] vor:

[P9] [non-projecting;  
hypotactic – extending;  
diegetic – event – ellipsis – forward  
– measurable  
– contiguous – overlapping  
– specific]

### T35 – T36 – T37

Hier existiert ein Wechsel von der Werkstatt zum Berg und zurück. Der diegetische Ort von T37 ist Arbeitsstation 11. Für beide Übergänge ist die gleiche paradigmatische Relation wie bei T03 – T04 (Wechsel Berg – Werkstatt), die »Wechsel-Relation« [P3], gültig.

### T37 – T38 – T39 – T40

Bei diesen Einstellungen finden Wechsel von einer Station zu einer weiteren statt (von Station 11 bis zu Station 14). Es liegen jeweils zeitliche Lücken vor. Daher können die Übergänge mit der »Itinerar-Relation« [P6] klassifiziert werden.

### T40 – T41

Hier findet ein Wechsel von der Werkstatt zum Berg statt. Es liegt die gleiche paradigmatische Relation wie bei T03 – T04 vor, die »Wechsel-Relation« [P3].

**T41 – T42**

Beide Einstellungen zeigen Johanna auf dem Berg. Es gibt keine zeitliche Lücke, die Räume überlappen sich. Daher liegt hier die Relation [P5] vor.

**Analyse nicht aufeinanderfolgender Einstellungen des gleichen Hauptortes**

Bisher wurde das ganze filmische Dokument in Layoutordnung analysiert, die deutlich eine alternante Struktur hat. Es stellt sich nun die Frage, ob innerhalb der beiden Stränge dieser Alternanz inhaltlich ergiebige Analysen möglich sind. Dazu werden nun die Einstellungspaare nicht aufeinanderfolgender Einstellungen des gleichen Hauptortes untersucht und zwar als Erstes die Einstellungspaare des diegetischen Ortes »Berg«.

**T03 – T11**

Beide Einstellungen zeigen den gleichen Ort, wobei sich die gezeigten Ausschnitte überlappen, die räumliche Relation kann daher als *contiguous – overlapping* klassifiziert werden. Da dem Beobachter nicht bekannt gemacht wird, wie lange die Arbeitsschritte in der Werkstatt dauern, wird die zeitliche Lücke als *indefinite* klassifiziert. Es liegt hier also als Relation [P10] vor:

```
[P10] [non-projecting;
      hypotactic – extending;
      diegetic – event – ellipsis – forward
                                – indefinite
                                – contiguous – overlapping
      – specific]
```

**T11 – T16**

Hier liegt die gleiche Situation wie beim Einstellungspaar T03 – T11 vor und somit auch die paradigmatische Relation [P10].

**T16 – T36**

Die zeitliche Lücke ist ebenfalls von unbestimmter Dauer. Die Einstellungsgröße der Einstellung T16 ist »HT«, die der Einstellung T36 »HN«. Daher ist die räumliche Relation *contiguous – inclusion – narrowing*. Es liegt als Relation [P11] vor:

```
[P11] [non-projecting;
      hypotactic – extending;
      diegetic – event – ellipsis – forward
                                – indefinite
                                – contiguous – inclusion – narrowing
      – specific]
```

**T36 – T41**

Die gezeigten Ausschnitte überlappen sich, die zeitliche Lücke zwischen den Einstellungen ist unbestimmt. Es liegt daher die Relation [P10] vor.

Für die Einstellungen des diegetischen Ortes »Berg« liegen somit die in Abb. 2.25 gezeigten paradigmatischen Relationen vor.

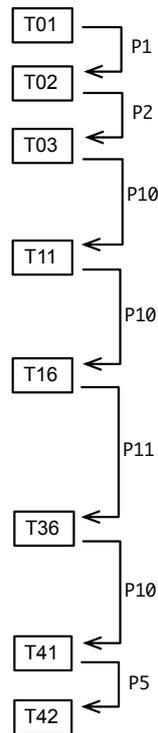


Abb. 2.25: Paradigmatische Relationen der Einstellungen des diegetischen Ortes »Berg«

Es folgen nun die Einstellungspaare des diegetischen Ortes »Werkstatt«.

#### **T10 – T12; T15 – T17; T35 – T37**

Bei den jeweiligen Übergängen liegt ein Wechsel von einer Station zu einer anderen Station vor. Es ist jeweils die gleiche paradigmatische Relation gültig wie bei einem »direkten« Stationswechsel, also ist die »Itinerar-Relation« [P6] klassifikatorisch anzuwenden.

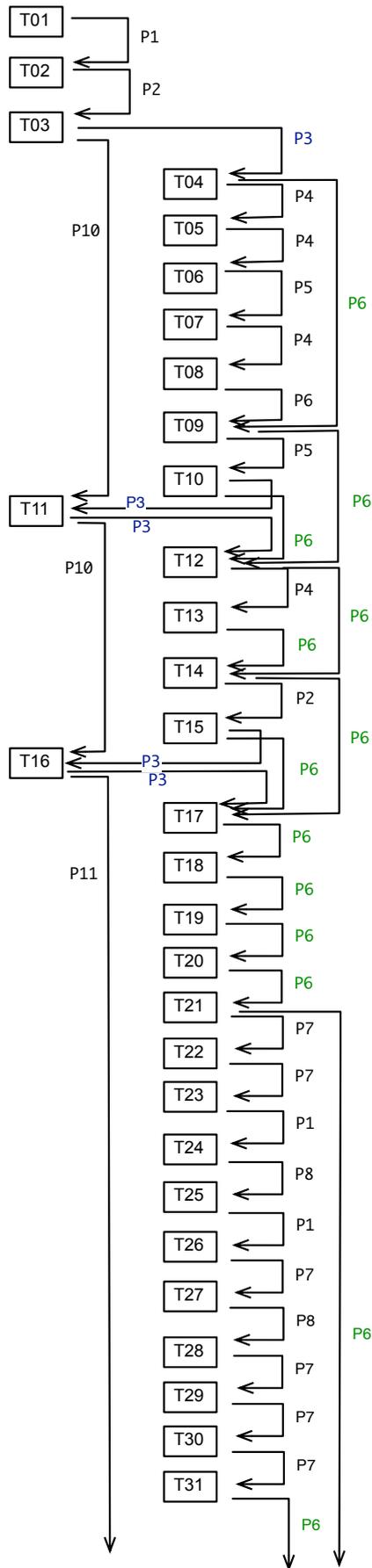
#### **Analyse der Relationen zwischen Einstellungen in der Werkstatt, die nicht aufeinanderfolgend sind und bei denen ein Stationswechsel vorliegt**

#### **T04 – T09; T09 – T12 ; T12 – T14; T14 – T17; T21 – T32; T32 – T37**

Bei allen Einstellungsübergängen liegt hier die Relation [P6] vor.

#### **Gesamtergebnis**

Abb. 2.26 zeigt eine Übersicht aller paradigmatischen Relationen des Videos »Snowboard«. Deutlich erkennbar ist die »Wechsel-Relation« [P3] und innerhalb des diegetischen Ortes Werkstatt die »Itinerar-Relation« [P6].



- [P1] [non-projecting;  
hypotactic - extending;  
diegetic - event - ellipsis - forward  
- measurable  
- contiguous - inclusion - narrowing  
- specific]
- [P2] [non-projecting;  
hypotactic - extending;  
diegetic - event - continuous  
- contiguous - inclusion - broadening  
- specific]
- [P3] [non-projecting; »Wechsel-Relation«  
paratactic - external  
- contrast;  
diegetic - event - ellipsis - forward  
- indefinite  
- non-contiguous - distant  
- specific]
- [P4] [non-projecting;  
hypotactic - extending;  
diegetic - event - continuous  
- contiguous - inclusion - narrowing  
- specific]
- [P5] [non-projecting;  
hypotactic - extending;  
diegetic - event - continuous  
- contiguous - overlapping  
- specific]
- [P6] [non-projecting; »Itinerar-Relation«  
hypotactic - extending;  
diegetic - event - ellipsis - forward  
- indefinite  
- non-contiguous - prolongation  
- specific]
- [P7] [non-projecting;  
hypotactic - extending;  
diegetic - event - ellipsis - forward  
- measurable  
- contiguous - inclusion - identity  
- specific]
- [P8] [non-projecting;  
hypotactic - extending;  
diegetic - event - ellipsis - forward  
- measurable  
- contiguous - inclusion - broadening  
- specific]
- [P9] [non-projecting;  
hypotactic - extending;  
diegetic - event - ellipsis - forward  
- measurable  
- contiguous - overlapping  
- specific]
- [P10] [non-projecting;  
hypotactic - extending;  
diegetic - event - ellipsis - forward  
- indefinite  
- contiguous - overlapping  
- specific]
- [P11] [non-projecting;  
hypotactic - extending;  
diegetic - event - ellipsis - forward  
- indefinite  
- contiguous - inclusion - narrowing  
- specific]

Abb. 2.26: Übersicht der paradigmatischen Relationen des Videos »Snowboard«

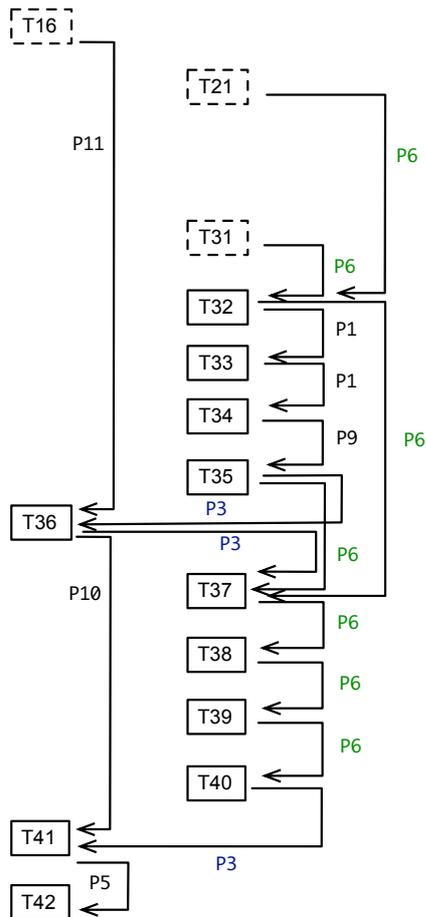


Abb. 2.27: Übersicht der paradigmatischen Relationen (Fortsetzung)

Im untersuchten Video wurden elf verschiedene paradigmatische Relationen klassifiziert. Wie zu erwarten war, kommen nicht alle möglichen Optionen des Netzwerks der »Grande Paradigmatique« in diesem Video vor. Abb. 2.28 zeigt die im Video vorhandenen Optionen im Netzwerk, die nicht vorhandenen Optionen sind in grau dargestellt.

Es gibt keinen Widerspruch zu den Erwartungen, die man an einen »Gebrauchsfilm« hat – im Gegenteil, es wird sichtbar, dass das diegetische Eigenleben der Agenten keine Rolle spielt, da Projektionen nicht vorkommen, Erzählweisen keine Rolle spielen, da die Relation *internal* nicht vorkommt, ferner nicht-diegetische Einstellungen nicht vorhanden sind, sondern auf spezifische Ereignisse fokussiert wird, die ohne Rückblenden mit allen räumlichen Relationen, die die Paradigmatik bietet, dargestellt werden.

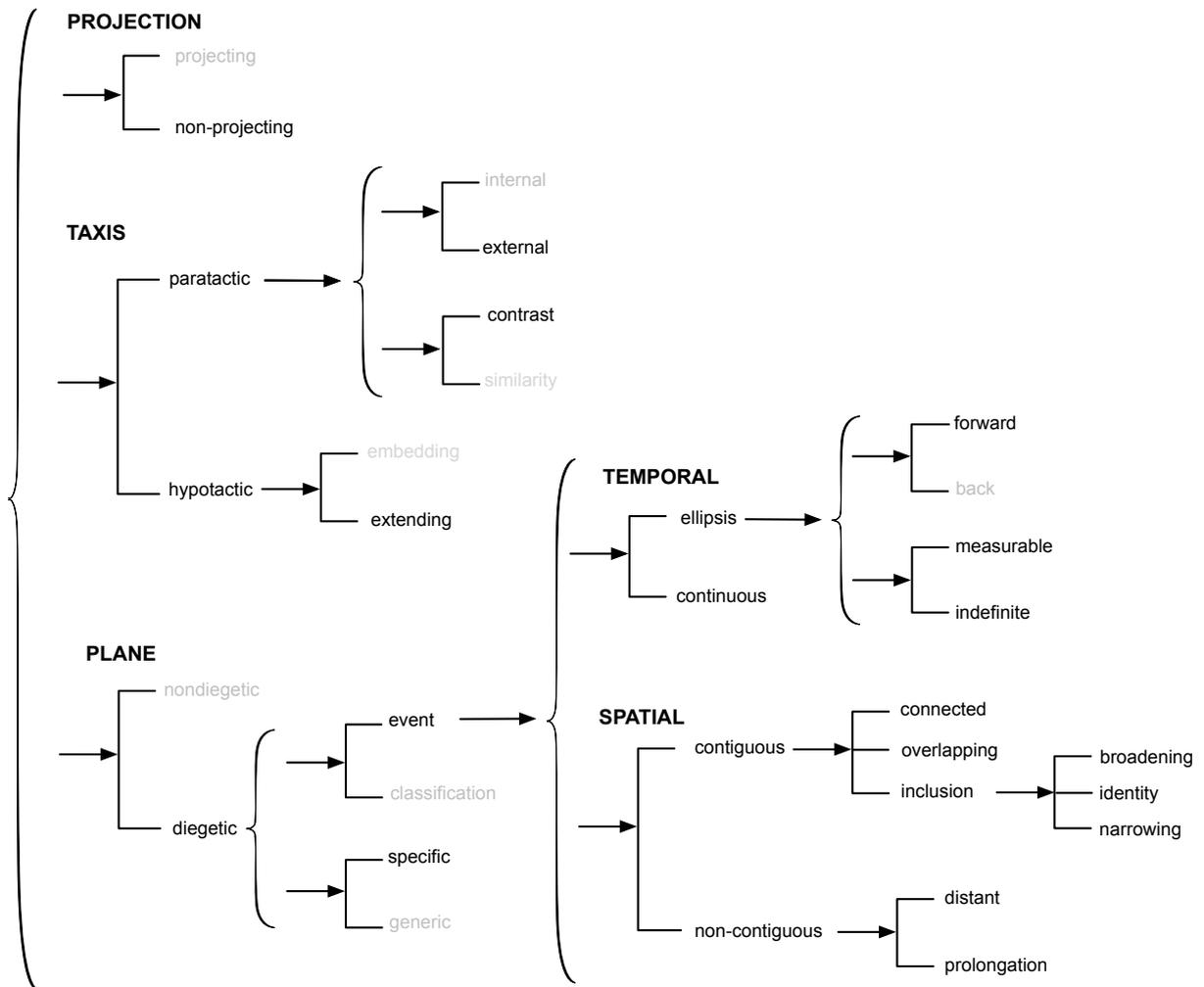


Abb. 2.28: Spezialisiertes Netzwerk der paradigmatischen Relationen des Videos »Snowboard«

### 2.5.3 Syntagmatische Analyse

Zur syntagmatischen Analyse eines Films ist es notwendig, jeweils ein maximales Segment zu finden, auf das die Definition eines der verschiedenen Syntagmen anwendbar ist [Bateman, Schmidt 2011, S. 252].

»[...] we will trying to find a syntagmatically maximal segment to which a syntagmatic class can be assigned.«

Im Beispielfideo »Snowboard« lassen sich, wie in der Wechseltabelle (Tab. 2.3) und der paradigmatischen Analyse dargestellt, die beiden folgenden Raumzeitumgebungen in den Einstellungen konzeptionalisieren:

- (1) Einstellungen mit Johanna auf dem Berg,
- (2) Einstellungen an verschiedenen Arbeitsstationen in einer Werkstatt.

Als erstes werden diese beiden Einstellungsmengen isoliert untersucht. Anschließend wird überprüft, ob es sich bei der gesamten Einstellungsmenge um ein alternantes Segment handelt.

#### Die Einstellungen mit Johanna auf dem Berg

Für die Einstellungen des Segmentes  $S_B = (T01, T02, T03, T11, T16, T36, T41, T42)$  kann beobachterseits der gleiche Raum, der Berg mit Johanna, angenommen werden. Zwischen einzelnen Einstellungen existieren zeitliche Lücken. Diese Einstellungsmenge erfüllt somit die Voraussetzungen von Definition 2.8 und kann als Sequenz klassifiziert werden.

Da zwischen  $T02$  und  $T03$  sowie zwischen  $T41$  und  $T42$  keine zeitlichen Lücken erkennbar sind, können die Einstellungsmengen  $S_{B\_01} = (T02, T03)$  und  $S_{B\_02} = (T41, T42)$  jeweils als Szene (nach Definition 2.7) klassifiziert werden. Insgesamt liegt also eine Sequenz vor, die zwei Szenen als Unterelemente enthält. Abb. 2.29 zeigt das Ergebnis der syntagmatischen Analyse von Segment  $S_B$ .

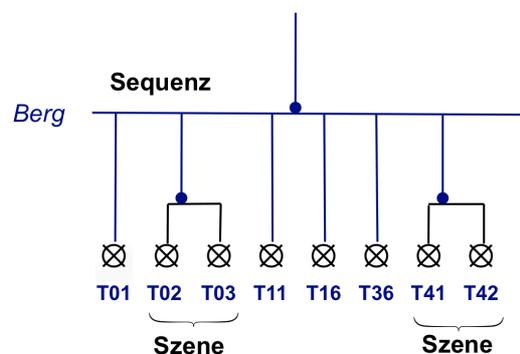


Abb. 2.29: Syntagmatische Analyse der Einstellungen auf dem Berg

### Die Einstellungen an verschiedenen Arbeitsstationen in einer Werkstatt

Die 14 Arbeitsstationen in der Werkstatt werden nun als räumlich getrennte Bereiche betrachtet. Einstellungen, zwischen denen eine spatiale paradigmatische Relation mit dem Wert *contiguous* vorliegt, werden zu einem Segment zusammengefasst. Somit ergeben sich die folgenden Segmente:

$$S_{W_{01}} = (T04, T05, T06, T07, T08), S_{W_{02}} = (T09, T10), S_{W_{03}} = (T12, T13),$$

$$S_{W_{04}} = (T14, T15),$$

$$S_{W_{05}} = (T17), S_{W_{06}} = (T18), S_{W_{07}} = (T19), S_{W_{08}} = (T20),$$

$$S_{W_{09}} = (T21, T22, T23, T24, T25, T26, T27, T28, T29, T30, T31),$$

$$S_{W_{10}} = (T32, T33, T34, T35),$$

$$S_{W_{11}} = (T37), S_{W_{12}} = (T38), S_{W_{13}} = (T39), S_{W_{14}} = (T40)$$

Zwischen den Einstellungen der Segmente  $S_{W_{01}}$ ,  $S_{W_{02}}$ ,  $S_{W_{03}}$  und  $S_{W_{04}}$  sind keine zeitlichen Lücken (paradigmatisch: *continuous*) erkennbar; daher können diese Segmente als Szenen (nach Definition 2.7) klassifiziert werden.

Zwischen den Einstellungen der Segmente  $S_{W_{09}}$  und  $S_{W_{10}}$  existieren zeitliche Lücken (paradigmatisch: *ellipsis*). Diese Segmente können daher als Sequenz (nach Definition 2.8) klassifiziert werden.

Die restlichen Einstellungen  $T17$ ,  $T18$ ,  $T19$ ,  $T20$ ,  $T37$ ,  $T38$ ,  $T39$ ,  $T40$  (Segmente  $S_{W_{05}}$ ,  $S_{W_{06}}$ ,  $S_{W_{07}}$ ,  $S_{W_{08}}$ ,  $S_{W_{11}}$ ,  $S_{W_{12}}$ ,  $S_{W_{13}}$ ,  $S_{W_{14}}$ ) können keinem der obigen Syntagmen zugeordnet werden, daher werden sie zunächst als Planszenen klassifiziert.

Es folgt also, dass jedes Segment  $S_{W_i}$ ,  $1 \leq i \leq 14$ , als narrativ basales Syntagma klassifiziert werden kann. Für das Gesamtsegment  $S_W = (S_{W_{01}}, \dots, S_{W_{14}})$ , bestehend aus 14 Teilsegmenten, existiert also eine Partitionierung in eben jene 14 Teilsegmente, die narrativ basale Syntagmen realisieren. Somit sind für das Gesamtsegment  $S_W$  die Voraussetzungen von Definition 2.10 erfüllt, das Segment  $S_W$  ist räumlich 14-partit.

Die paradigmatische Analyse zeigt, dass zwischen diesen Teilsegmenten (genauer: zwischen der letzten Einstellung des einen Segmentes und der ersten Einstellung des folgenden) jeweils die paradigmatische Relation [P6] vorliegt:

[P6] [non-projecting;  
hypotactic – extending;  
diegetic – event – ellipsis – forward  
– indefinite  
– non-contiguous – prolongation  
– specific]

Die temporale paradigmatische Relation hat den Wert *ellipsis - forward*, d. h. jede der genannten Einstellungen liegt zeitlich vor der folgenden Einstellung, für die die diegetischen Zeiten der Einstellungen liegt die Allen'sche Relation *before* vor.

Somit erfüllt die Menge  $S_W$  die Voraussetzungen einer narrativen Folge (Definition 2.11). Genauer gesagt handelt es sich hier um den Subtyp »Itinerar«, bei dem im Zeitverlauf eine »Figur«, in diesem Fall das Snowboard, verschiedene Orte, hier die Stationen in der Werkstatt, durchläuft.

Bei dieser narrativen Folge liegt folgende spezielle Situation vor: Zu Beginn eines Segmentes  $S_{W_i}$  liegt das Werkstück in einem Zustand  $Z_k$  (Eingangszustand) vor. Es wird gezeigt, wie das Werkstück verändert und ein Ausgangszustand  $Z_{k+1}$  erzeugt wird. Dieser Ausgangszustand ist nun Eingangszustand für den nächsten Arbeitsschritt, der dann im darauffolgenden Segment gezeigt wird. Abb. 2.30 visualisiert diese Zustandswechsel in Form eines Petri-Netzes.<sup>1</sup>

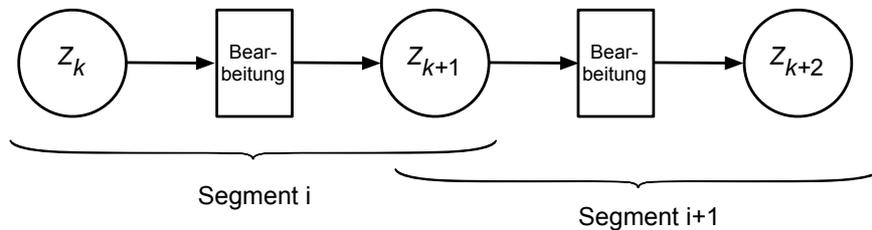


Abb. 2.30: Zustandswechsel in der Werkstatt darstellt als Petri-Netz

Abb. 2.31 zeigt nun das Ergebnis der syntagmatischen Analyse der Einstellungen in der Werkstatt.

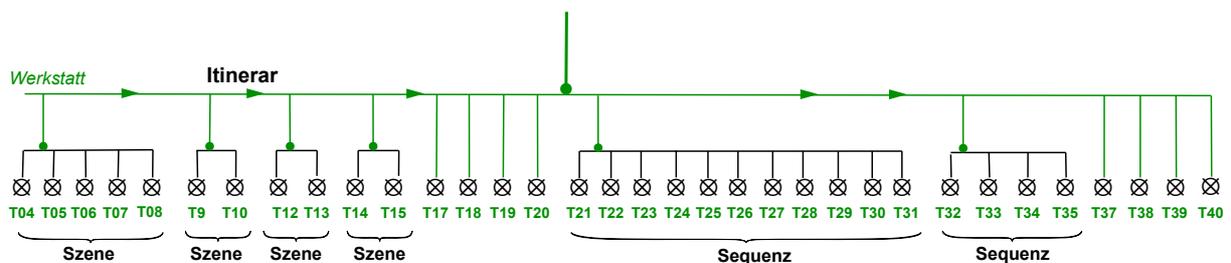


Abb. 2.31: Syntagmatische Analyse der Einstellungen in der Werkstatt

<sup>1</sup> Vgl. dazu auch Abb. 4.21: In dieser werden die Zustandswechsel des gesamten Videos »Snowboard« dargestellt.

### Die Wechsel zwischen Berg und Werkstatt

Die Arbeitsstationen  $S_i, 1 \leq i \leq 14$ , bilden räumlich getrennte Bereiche. Aufgrund der dargestellten Gesamtsituation kann jedoch davon ausgegangen werden, dass sich alle Stationen in einem Raum, einer Werkstatt, befinden. Die Werkstatt kann somit als ein zusammenhängender Raum klassifiziert werden und es gilt:

$$\text{Werkstatt} = \bigcup_{i=1}^{14} S_i$$

der Werkstatttraum stellt die Vereinigungsmenge der Arbeitsstationen dar.

Im gegebenen Video wird regelmäßig zwischen Einstellungen am Berg und Einstellungen in der Werkstatt gewechselt, siehe Tab. 2.5. Alle paradigmatischen Relationen zwischen den Teilsegmenten<sup>1</sup> wurden als *paratactic - external | contrast* klassifiziert.

Berg	Werkstatt
SegBerg1 (T01 – T03)	
	SegWerkstatt1 (T04 – T10)
SegBerg2 (T11)	
	SegWerkstatt2 (T12 – T15)
SegBerg3 (T16)	
	SegWerkstatt3 (T17 – T35)
SegBerg4 (T36)	
	SegWerkstatt4 (T37 – T40)
SegBerg5 (T41 – T42)	

Tab. 2.5: Wechsel zwischen Berg und Werkstatt

Es liegt die Vermutung nahe, dass ein 2-alternantes (bi-partites) Segment gegeben ist. Daher sind nun die Voraussetzungen von Definition 2.13 zu prüfen:

Zu (1): Das Segment ist monochron. Dies ist erfüllt.

Zu (2): Das Segment ist räumlich bi-partit. Für das Segment existiert eine Partition in zwei disjunkte Mengen, wobei eine die Einstellungen auf dem Berg, die andere die Einstellungen in der Werkstatt enthält. Jedes Teilsegment kann zudem, da es zeitliche Lücken enthält, als Sequenz, also als ein narrativ basales Syntagma, klassifiziert werden.

Zu (3): Existenz einer Kohärenzrelation. Nach Definition 2.13 (3) muss eine Relation  $R \subseteq \text{SpaceTime}_1 \times \text{SpaceTime}_2$  vorliegen, die symmetrisch ist, oder zusätzlich eine Relation  $R'$  existieren, für die die Disjunktion  $R \vee R'$  symmetrisch ist.

Für den vorliegenden Fall können für alle Transitionen die folgenden kohärenzstiftenden Relationen konzeptionalisiert werden:

<sup>1</sup> Genauer gesagt: die Relationen zwischen der letzten Einstellung eines Segmentes zur ersten Einstellung des folgenden Segmentes.

$R(\text{Johanna, Snowboard})$ : »Johanna gibt das Snowboard in die Werkstatt (zurück)«,

$R'(\text{Johanna, Snowboard})$ : »Johanna erhält das Snowboard zum Test«,

wobei gilt, dass  $R \vee R'$  symmetrisch ist.

Auch die weiteren Voraussetzungen von Definition 2.13 sind erfüllt:

Zu (4): Die Raumzeiturbildpaare (Berg, Werkstatt) sind disjunkt.

Zu (5): Es gibt mindestens drei Transitionen, in diesem Fall sogar acht.

Somit sind alle Voraussetzungen von Definition 2.13 erfüllt und die Vermutung, dass die gesamte Einstellungsmenge beobachterrelational als bi-partites Segment konzeptionalisiert werden kann, hat sich bestätigt.

Tab. 2.5 zeigt die Wechsel zwischen Berg und Werkstatt. Ergänzend zu dieser Darstellung sind in Tab. 2.6 zusätzlich die genauen Stationen innerhalb der Werkstatt vermerkt.

Berg	Werkstatt
T01   T03	
	T04 – S <sub>1</sub>   T10 – S <sub>2</sub>
T11	
	T12 – S <sub>3</sub>   T15 – S <sub>4</sub>
T16	
	T17 – S <sub>5</sub>   T35 – S <sub>10</sub>
T36	
	T37 – S <sub>11</sub>   T40 – S <sub>14</sub>
T41   T42	

Tab. 2.6: Wechsel zwischen Berg und Werkstatt mit Arbeitsstationen

In Tab. 2.6 ist gut das folgende Prinzip erkennbar:  $S_k \rightarrow \text{Berg} \rightarrow S_{k+1}$  für  $k = 2, 4, 10$ . Dies bedeutet: Beim Wechsel der Darstellung zum Berg wird eine Station  $S_k$  verlassen. Die Darstellung im Video kehrt jedoch nicht zu dieser Station zurück, sondern zur darauf folgenden Station  $S_{k+1}$ . Dies macht auch Sinn, da der Produktionsprozess fortschreitet. Die Berg-Segmente stellen »Zwischenschnitte« dar, die zur Aufmerksamkeitsweckung der jugendlichen Zuschauer gedacht sind.

### **Gesamtergebnis**

Das Gesamtergebnis der syntagmatischen Analyse wird in Abb. 2.32 dargestellt. Es existieren zwei diegetische Hauptorte: »Berg« und »Werkstatt«. Zwischen den Berg-Segmenten und den Werkstatt-Segmenten ist eine Alternanz konzeptionalisierbar. Als basale Syntagmen lassen sich sechs Szenen (zwei auf dem Berg, vier in der Werkstatt) und drei Sequenzen (eine auf dem Berg, zwei in der Werkstatt) identifizieren. Zwölf Einstellungen können keinem der Syntagmen Szene bzw. Sequenz zugeordnet werden und werden als Planszenen klassifiziert. Der diegetische Raum Werkstatt enthält 14 Teilräume, die einzelnen Arbeitsstationen. Syntagmatisch gesehen handelt es sich bei den Einstellungen in der Werkstatt um eine narrative Folge, genauer gesagt um den Subtyp »Itinerar«.

In Abb. 2.33 (in Anlehnung an [Bateman, Schmidt 2011, S. 286]) werden alle Ergebnisse der syntagmatischen Analyse in detaillierterer Form wiedergegeben. Jede Einstellung wird – wie im Transkriptionsprotokoll – durch einen Screenshot repräsentiert. Die verschiedenen diegetischen Hauptorte werden in der Grafik dargestellt durch eine klare Trennung in einen »oberen« und einen »unteren Bereich, um so die Alternanz zu visualisieren. Der »Weg« des Snowboards durch die Werkstatt, das Itinerar, ist erkennbar, insbesondere auch, weil Einstellungen, die innerhalb der Werkstatt an einer Arbeitsstation stattfinden, innerhalb eines Rechtecks mit abgerundeten Ecken dargestellt werden, um so ihre räumliche Zusammengehörigkeit zu kennzeichnen. Die Klassifikation einer Einstellungsfolge als Szene bzw. Sequenz ist textuell ergänzt.

Durch die syntagmatische Analyse ist geklärt, welche logischen Relationen zwischen Teilen des Dokumentes bestehen. Durch diese logischen Relationen werden allen Transformationen des Beispielfideos (s. Kapitel 4) Randbedingungen auferlegt.

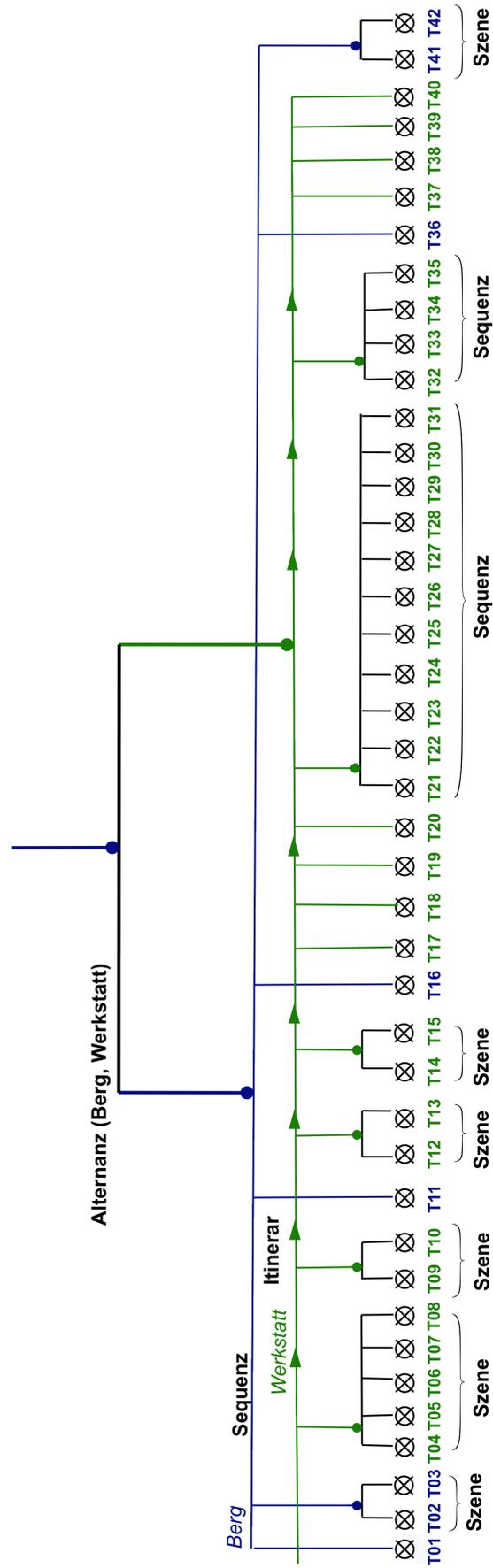


Abb. 2.32: Syntagmatische Analyse des Videos »Snowboard«



Abb. 2.33: Detaillierte Darstellung der syntagmatischen Analyse des Videos »Snowboard«

#### 2.5.4 Analyse der Text-Bild-Relationen

Das in dieser Arbeit untersuchte Video »Snowboard« ist vollständig (in jeder Einstellung) mit Sprechertext unterlegt. Ein männlicher Sprecher aus dem OFF kommentiert die Handlungen. Der sprachliche Duktus wird durch den Schnitt beherrscht. In vielen Fällen enden die Satzkonstruktionen jeweils mit den Schnitten, d. h. die Betextung passt sich der Montage an. In wenigen Fällen wird durch Sprache eine Verbindung zwischen zwei Einstellungen geschaffen, hier liegt ein sogenannter einstellungsübergreifender Text vor. Es kann davon ausgegangen werden, dass der Sprecher die Bildfolge genau kennt und die Betextung nach dem Zusammenfügen der Einstellungen erfolgte (vgl. hierzu auch [Stötzel 1990]).

Die Dauer der Einstellungen im Video variiert zwischen 1 und 35 Sekunden, im Durchschnitt beträgt sie 10,2 Sekunden. Mit zunehmender Einstellungslänge ist es tendenziell unwahrscheinlicher, dass zwischen dem filmischen Bild und dem zugehörigen Sprechertext nur ein einziger Relationstyp vorliegt. Um den Fall mehrerer Relationen für eine Einstellung repräsentieren zu können, wird der Sprechertext jeder Einstellung grundsätzlich in Segmente (Textspannen) eingeteilt. Die Textsegmente werden möglichst klein gewählt (Sätze oder vollständige Teilsätze), um eine möglichst detaillierte Analyse zu ermöglichen; gleichzeitig wird darauf geachtet, dass jedes Textsegment eine vollständige Sinneinheit darstellt. So wird beispielsweise der Satz »Die sollen die Schwingungen abfedern, wenn man später damit fährt« als ein Textsegment ausgezeichnet, weil der zweite Teil ohne den ersten nicht verständlich ist.<sup>1</sup> Die Segmentierung ist im ganzen Dokument gut möglich, da der Text ein »Text zum Hören« ist und die Zielgruppe Kinder sind. Die Sätze des Sprechertextes sind daher relativ kurz und komplizierte Schachtelsätze kommen nicht vor. Einstellungsübergreifende Textsegmente werden bei der Analyse gesondert markiert.

Im Folgenden wird die Analyse der Text-Bild-Relationen des Videos anhand einiger Einstellungen exemplarisch erläutert. Eine vollständige tabellarische Übersicht befindet sich in 6.2.

Die Einstellungen T01 bis T04 zeigen ein Mädchen auf einem schneebedeckten Berg.

T01\_01: »Da kommt Johanna.«

Hier wird der Name des Mädchens genannt, eine Information, die im Bild natürlich nicht ersichtlich ist. Daher wird diese Relation als *expansion – extension – collection* klassifiziert.

---

<sup>1</sup> Bei einer RST-Analyse (Rhetorical Structure Theory) würde der Satz in zwei Einheiten zerlegt werden.

T01\_02: »Ausgerüstet wie der Teufel mit Sturzhelm und Schneebrille,«

Durch diesen Text soll Aussage des Bildes ergänzt werden (unabhängig von der Stichhaltigkeit des Vergleichs). Daraus ergibt sich eine Klassifikation der Relation als *expansion – elaboration – specification*.

T01\_03: »und die will den Berg 'runterfahren«

Hier weiß der Sprecher mehr, als auf dem Bild zu sehen ist, er gibt quasi einen Ausblick auf das, was Johanna geplant hat. Dies dient der Weckung der Neugier und Aufmerksamkeit der Zuschauer und verstärkt die Motivation, das Video weiter zu betrachten. Daher wird diese Relation als *expansion – enhancement – motivation<sup>1</sup>* klassifiziert.

T01\_04: »Aber womit? Vielleicht will sie Ski fahren?«

Bei diesem Textsegment handelt es sich um das sprachliche Gestaltungsmittel der »*Hypophora*«. Eine *Hypophora* ist eine Frage, die vom Autor bzw. Redner im Folgenden selbst beantwortet wird. Durch dieses Gestaltungsmittel wird (erneut) die Aufmerksamkeit der Leser bzw. Zuhörer geweckt. Die *Hypophora* wird als neuer Relationstyp in die Taxonomie aufgenommen und als Untertyp zu *decoration* eingeordnet.

T01\_05: »und dazu hat sie erstmal« – T02\_01: »ein Brett dabei«

Durch den einstellungsübergreifenden Sprechertext wird hier eine Verbindung zwischen den beiden Einstellungen T01 und T02 geschaffen. Der Text formuliert das, was der Zuschauer auch sieht. Der Relationstyp ist daher *expansion – elaboration – explanation*.

T02\_02: »Aber zum Skifahren gehören doch zwei Bretter?«

Wie im Textsegment T01\_04 stellt der Sprecher eine Frage, die er im Folgenden beantwortet. Es liegt also wieder der Relationstyp *decoration – hypophora* vor.

T03\_01: »Es gibt aber doch auch diese Schneebretter "Snowboards". Da braucht man ein Brett.«

Im Textsegment T03\_01 gibt der Sprechertext eine neue Information, daher liegt der Relationstyp *expansion – extension – collection* vor.

T03\_02: »Vielleicht ist es so 'was?«

Hier liegt wieder der Relationstyp *decoration – hypophora* vor.

T03\_03: »Aber funktionieren tut das nicht, auch wenn sie sich quer draufstellt.«

Der Sprechertext enthält die Information, die im Bild zu sehen ist. Durch »auch wenn sie sich quer draufstellt« wird diese jedoch weiter präzisiert, daher wird die Relation als *expansion – elaboration – specification* klassifiziert.

---

<sup>1</sup> In einer RST-Analyse wird eine Relation als *motivation* klassifiziert, wenn der Satellit die Funktion »*information intended to increase the reader's desire to perform the action*« erfüllt [Mann, Taboada 2014].

T04\_01: »Also müssen wir 'mal dahingehen, wo solche Dinger hergestellt werden.«

Der Sprechertext gibt die Information, dass jetzt der Handlungsort gewechselt wird. Wie in T01\_03 ist diese Bemerkung wahrnehmungsleitend und aufmerksamkeitsweckend und verstärkt somit die Motivation, das Video weiter zu betrachten. Daher wird die Relation als *expansion – enhancement – motivation* klassifiziert.

Die folgenden Einstellungen wurden in einer Werkstatt gedreht.

T05\_01: »Da werden ja mehrere solcher Bretter genommen.«

Der Text beschreibt genau die gezeigte Handlung, daher liegt der Relationstyp *expansion – elaboration – explanation* vor.

T06\_01: »Bestimmte Anzahl.«

Es wird zwar nicht die genaue Anzahl genannt, jedoch stellt auch »Bestimmte Anzahl« eine Präzisierung dar, daher liegt der Relationstyp *expansion – elaboration – specification* vor.

T06\_02: »Die kommen in so 'nen Vorratsstapel.«

Der Sprechertext gibt die Information, die im Bild zu sehen ist, es liegt wieder der Relationstyp *expansion – elaboration – explanation* vor.

T06\_03: »Dann wird unten von diesem Vorratsstapel immer ein Brett abgezogen.«

T06\_04: »Das wandert dann unter diesem Rad dahinten durch, nach hinten.«

In den Segmenten T06\_03 und T06\_04 wird dem Zuschauer durch das Demonstrativpronomen »diesem« ein direkter Verweis vom Text zum Bild gegeben. Daher werden beide Relationen als *expansion – elaboration – identification* klassifiziert.

T07\_01: »Und danach in so 'ne grüne Kiste.«

T07\_02: »Vorne rein und hinten wieder raus,«

Die Textsegmente T07\_01 und T07\_02 beschreiben die gezeigte Handlung. Es liegt der Relationstyp *expansion – elaboration – explanation* vor.<sup>1</sup>

T07\_03: »und immer drei Bretter gehören zusammen.«

Hier gibt der Text eine Präzisierung an, daher liegt der Relationstyp *expansion – elaboration – specification* vor.

T07\_04: »Und der Strich bedeutet nicht, dass da abgeschnitten werden soll, sondern da ist später hinten.«

Der Sprechertext gibt an, welchen Zweck bzw. welche Zielsetzung die ausgeführte Handlung hat. Aus diesem Grund wird die Relation als *expansion – enhancement – purpose*<sup>2</sup> klassifiziert.

---

<sup>1</sup> Versteht man T07\_01 »in so ne grüne Kiste« als direkten Verweis auf das Bild, dann würde die Relation als *expansion – elaboration – identification* klassifiziert.

<sup>2</sup> RST-Relation *purpose*: »the intent behind the situation« [Mann, Taboada 2014].

*T10\_02: »An dieser Stelle muss man darauf achten, dass der Kleber gut in die Zwischenräume zwischen die Bretter kommt, also in die Ritzen.«*

Der Sprechertext gibt genaue Hinweise, wie die gezeigte Handlung ausgeführt werden soll. Es liegt die Relation *expansion – enhancement – manner* vor.

Die Einstellung *T11* spielt wieder auf dem Berg.

*T11\_01: »So, das sieht doch jetzt schon viel besser aus.«*

Hier kommentiert der Sprecher die gezeigte Situation, vermutlich ist dieser Kommentar auch subjektiv. Für diese Art der »Kommentierung« wird der neue Relationstyp »comment« eingeführt und als Untertyp von *decoration* eingeordnet.

*T11\_02: »Das Brett ist breiter, man kann in der Mitte auch besser drauf stehen.«*

Im Bild wird Johanna mit dem Brett in der Hand und anschließend stehend auf dem Brett gezeigt. Der Text gibt dazu präzisere Informationen über das Brett und dessen Verwendbarkeit. Daher wird die Relation als *expansion – elaboration – specification* klassifiziert.

*T11\_03: »Mal sehen, ob das geht.«*

Dieser Text wirkt wie eine Vorbereitung oder Überschrift für das Folgende. Für entsprechende Relationen wird ein weiterer Relationstyp benötigt. In Anlehnung an die RST-Analyse wird dieser Relationstyp als »preparation« bezeichnet<sup>1</sup> und als Untertyp zu »decoration« eingeordnet.

*T11\_04: »Naaaa. Vielleicht doch nicht so,«*

Es liegt der Relationstyp *decoration – comment* vor, da es sich um einen Kommentar des Sprechers handelt.

*T11\_05: »Vielleicht, wenn man sich quer draufstellt?«*

Dieser Sprechertext wirkt auf die Zuhörer bzw. Zuschauer wie eine Aufforderung an Johanna (und nicht wie in *T11\_03* als Wendung an die Zuschauer), die von Johanna auch sofort befolgt wird. Für eine solche Relation wird der Typ *invitation* als Untertyp von *decoration* eingeführt.

*T11\_06: »Bisschen rutschen tut's.«*

Dieser Text beschreibt genau das, was der Zuschauer sieht, daher wird die Relation als *expansion – extension – explanation* klassifiziert.

*T11\_07: »Vielleicht auch allen Vieren mit anstoßen?«*

Wie im Textsegment *T11\_05* wird eine Aufforderung an Johanna formuliert. Die Relation ist vom Typ *decoration – invitation*.

*T11\_08: »Aber so der wahre Jakob ist das auch noch nicht.«*

Dies ist wieder ein Sprecherkommentar, also liegt der Relationstyp *decoration – comment* vor.

---

<sup>1</sup> RST-Relation *preparation*: »text which prepares the reader to expect and interpret the text to be presented« [Mann, Taboada 2014].

Zum Abschluss werden noch einige Relationen zu Einstellungen in der Werkstatt analysiert, die zum einen die bereits genutzten Relationen noch einmal exemplifizieren, zum anderen eine weitere einführen.

T17\_01: »Als nächstes kommen in alle Löcher solche Metallteile.«

T17\_02: »Das sind Metallgewinde, «

T18\_01: »Das sind Kunststoffbahnen mit Blümchen drauf.«

Die Relationen der Textsegmente T17\_01, T17\_02, T18\_01 zum Filmbild könnten zunächst als *specification* klassifiziert werden, wenn man die Informationen »Metallteile«, »Metallgewinde«, »Kunststoffbahnen« im Sinne von *expansion – elaboration – specification* zur Präzisierung der gezeigten Situation interpretiert. Andererseits werden hier Zusatzinformationen über verwendete Materialien gegeben, bei denen nicht unbedingt davon ausgegangen werden kann, dass jeder Zuschauer diese anhand des Bildes erkennen kann. Von daher erscheint eine Klassifikation als *expansion – extension – collection* sinnvoller.

T21\_02: »weil das die Lauffläche ist, kommt die natürlich als erstes in die Form – soll ja schließlich unten sein.«

Hier wird im Sprechertext eine Begründung für die gezeigte Handlung gegeben. Eine Relation »*reason*« fehlt in der Taxonomie von Kong. Diese Relation wird in die Taxonomie eingefügt und als Untertyp von *enhancement* eingeordnet, da sie für Zusatzinformationen steht.

### Zusammenfassung

Abb. 2.34 zeigt die Anzahl der verschiedenen Text-Bild-Relationen sowie die Verteilung auf die beiden Orte »Berg« und »Werkstatt«.<sup>1</sup> Die in die Taxonomie neu aufgenommenen Relationen sind mit »\*« gekennzeichnet.

Als Untertyp zu *expansion – enhancement* wurde eine Relation neu eingeführt:

*reason*: Begründung für die gezeigte Handlung.

Als Untertypen zu *decoration* wurden folgende Relationen neu eingeführt:

*hypophora*: Frage, die vom Autor bzw. Redner im Folgenden selbst beantwortet wird,

*comment*: Kommentierung (evtl. subjektiv),

*invitation*: direkte Aufforderung an Johanna,

*preparation*: »Überschrift« für das Folgende.

Der Typ *expansion* ist der am häufigsten vorkommende Relationstyp: 94 der 114 Relationen sind von diesem Typ, was 82,5% entspricht. Vom Typ *decoration* sind 20 Relationen und damit 17,5%. Der deutlich höhere Anteil des Relationstyps *expansion* passt zur Intention des Videos, eine »Sachgeschichte« darzustellen. Jedoch scheint der Anteil des Relationstyps *decoration* mit

<sup>1</sup> Die Auswertung wurde mit dem XSLT-Stylesheet `textImage_report.xml` (s. 4.5.5) generiert.

17,5% für eine »Sachgeschichte« ungewöhnlich hoch, da diese für Texte mit »wenig oder keinem Informationswert« stehen. Unter dem Aspekt, dass die Aufgabe dieser Textsegmente darin besteht, die Aufmerksamkeit der Leser bzw. Zuschauer, im Beispielvideo also von Kindern und Jugendlichen, zu wecken, ist dieser hohe Anteil erklärbar.

Der häufigste Untertyp von *expansion* ist der Typ *elaboration* mit einer Anzahl von 64 Relationen, was einem Anteil von 68,1% entspricht. Danach folgt der Typ *enhancement* mit 17 Relationen (18,1%) und der Typ *extension* mit 13 Relationen (13,8%). Dies bedeutet, dass bei etwa zwei Dritteln der Textsegmente der Text »lediglich« die gezeigten Handlungen beschreibt und keine neuen bzw. zusätzlichen Informationen zum Bild liefert. Der Anteil an Textsegmenten, die zusätzliche Informationen liefern, scheint mit 13,8% relativ gering. Zum einen liegt dieser relativ geringe Anteil sicherlich darin begründet, dass durch das Filmbild schon sehr viele Informationen übermittelt werden (s. 2.4), zum anderen vermutlich daran, dass die Autoren die junge Zielgruppe nicht mit zu vielen technischen Details überfordern wollten.

Etliche Relationstypen der dritten Untertyp-Ebene kommen gar nicht vor (z. B. *exemplification*, *variation*, *sequence*, *cause*, *effect*, usw.). Neu hinzugefügt wurde als Untertyp von *enhancement* der Relationstyp »*reason*«, der vorliegt, wenn im Text eine Begründung für die gezeigte Handlung gegeben wird.

Betrachtet wird nun die Verteilung der beiden Haupttypen *expansion* und *decoration* innerhalb der beiden diegetischen Orte »Werkstatt« und »Berg«. Beim diegetischen Ort »Werkstatt« sind 79, dies entspricht 95% der Textsegmente, vom Typ *expansion* und somit lediglich 4 (5%) vom Typ *decoration*. Beim diegetischen Ort »Berg« sieht die Verteilung deutlich anders aus: Hier sind etwas mehr als die Hälfte der Relationen ist vom Typ *decoration*, nämlich 16 (52%) und dementsprechend 15 (48%) vom Typ *expansion*. In der Werkstatt wird der Produktionsprozess des Snowboards gezeigt, daher ist klar, dass hier der Typ *expansion* dominieren muss. Die Berg-Segmente stellen Einschübe (Zwischenschnitte) dar, die wahrnehmungsleitend und aufmerksamkeitserweckend sein sollen, weshalb hier auch ein entsprechend hoher Anteil des Relationstyps *decoration* zu erwarten ist.

## Auswertung Text-Bild-Relationen

Anzahl gesamt: 114 (Berg: 31, Werkstatt: 83)

————— **expansion: 94** (Berg: 15, Werkstatt: 79)  
 ————— **elaboration: 64** (Berg: 10, Werkstatt: 54)  
 ————— **explanation: 27** (Berg: 5, Werkstatt: 22)  
 ————— exemplification: 0  
 ————— **specification: 24** (Berg: 5, Werkstatt: 19)  
 ————— **identification: 13** (Berg: 0, Werkstatt: 13)  
 ————— **extension: 13** (Berg: 2, Werkstatt: 11)  
 ————— **collection: 13** (Berg: 2, Werkstatt: 11)  
 ————— variation: 0  
 ————— sequence: 0  
 ————— response: 0  
 ————— alternation: 0  
 ————— **enhancement: 17** (Berg: 3, Werkstatt: 14)  
 ————— **spatio-temporal: 2** (Berg: 0, Werkstatt: 2)  
 ————— **manner: 1** (Berg: 0, Werkstatt: 1)  
 ————— cause: 0  
 ————— effect: 0  
 ————— condition: 0  
 ————— means: 0  
 ————— **purpose: 9** (Berg: 2, Werkstatt: 7)  
 ————— justification: 0  
 ————— concession: 0  
 ————— **motivation: 2** (Berg: 1, Werkstatt: 1)  
 ————— enablement: 0  
 ————— restatement: 0  
 ————— summary: 0  
 ————— **reason\*: 3** (Berg: 0, Werkstatt: 3)  
 ————— projection: 0  
 ————— speech: 0  
 ————— thought: 0  
 ————— **decoration: 20** (Berg: 16, Werkstatt: 4)  
 ————— **hypophora\*: 6** (Berg: 3, Werkstatt: 3)  
 ————— **comment\*: 6** (Berg: 6, Werkstatt: 0)  
 ————— **invitation\*: 4** (Berg: 4, Werkstatt: 0)  
 ————— **preparation\*: 4** (Berg: 3, Werkstatt: 1)

Abb. 2.34: Auswertung der Text-Bild-Relationen des Videos »Snowboard«

## 3 Algebraische Handlungssemantik

### 3.1 Einleitung

In den Definitionen der basalen Syntagmen »Planszene«, »Szene«, »Sequenz« (s. Kapitel 2) wurde bewusst auf eine handlungstheoretische Fundierung verzichtet. In diesem Kapitel steht nun die cinematographische Repräsentation von Handlungen im Focus. Um dies zu formalisieren, wird auf Ergebnisse der formalen Semantik zurückgegriffen. Speziell werden Verben als einstellige Prädikate über Ereignisse dargestellt und die Partizipanten des Ereignisses durch zweistellige Hilfsrelationen, den sogenannten thematischen Relationen, mit dem Ereignis verbunden. Anhand bestimmter Eigenschaften der thematischen Relation in Kombination mit den Eigenschaften des betroffenen Objektes lässt sich der Handlungstyp der Gesamtaussage (nach der Vendler'schen Klassifikation [Vendler 1957, 1967]) als *State*, *Activity*, *Achievement* oder *Accomplishment* bestimmen. Zur Übertragung dieser Ergebnisse auf cinematographische Dokumente werden einfache Modellierungskategorien auf Basis des MPEG-7 Standards definiert und segment-thematische Relationen eingeführt, die die Beziehung zwischen filmischen Segmenten (auch unterhalb der Einstellungsebene) und instanziierten Ereignissen ausdrücken. Darauf aufbauend wird algebraisch formalisiert, unter welchen Voraussetzungen eine Handlung in einem cinematographischen Dokument als Handlungstyp *Accomplishment* konzeptionalisierbar ist. Da für cinematographische Dokumente wichtig ist, in welcher Art und Weise Handlungen in bzw. auf (verschiedene) Einstellungen aufgeteilt/verteilt sind, werden nun drei Realisationstypen unterschieden:

- Eine Handlungsinstanz und eine Einstellung entsprechen einander (wie in einer Planszene);
- mehrere Handlungsinstanzen finden sich in einer Einstellung;
- eine Handlungsinstanz wird auf mehrere Einstellungen verteilt.

Deren filmische Umsetzung wird anhand von Beispielen aus dem Video »Snowboard« analysiert.

## 3.2 Grundlegende Begriffe und Definitionen

### 3.2.1 Klassen, Instanzen und ihre Relationen

Als Sammelbegriff für Objekte, Prozesse, Funktionen, Strukturen, Orte, Zeiten wird Entität (*entity*) verwendet. Innerhalb dieser Entitäten lassen sich zwei spezielle Teilgesamtheiten unterscheiden: Instanzen und Klassen [Smith, Rosse 2004, S. 444].

Instanzen sind Individuen (*individuals, particulars, tokens*) einer speziellen Sorte. Klassen (*universals, kinds, types*) existieren durch ihre zugehörigen Instanzen. Zwischen Instanzen und Klassen bestehe eine binäre Relation der Instanziierung »*inst*«. An erster Stelle dieser Relation steht immer eine Instanz, an zweiter Stelle die Klasse, die die Instanz instanziiert.

So lässt sich eine Klasse definieren als alles, was instanziiert werden kann [Smith 2005, S. 511]:

$$\text{class}(e) =_{\text{def}} \exists f \text{ inst}(f, e)$$

Umgekehrt ist eine Instanz alles, was eine Klasse instanziiert:

$$\text{instance}(f) =_{\text{def}} \exists e \text{ inst}(f, e)$$

Es gelten die folgenden Axiome:

(1) Jede Klasse hat mindestens eine Instanz:

$$\forall e (\text{class}(e) \rightarrow \exists f (\text{instance}(f) \wedge \text{inst}(f, e)))$$

(2) Nichts kann sowohl Instanz als auch Klasse sein:

$$\neg \exists e (\text{class}(e) \wedge \text{instance}(e))$$

Instanzen sind vom Typ IND (*Token*), Klassen vom Typ  $\Lambda$  (*Types*).

Für Instanzen vom Typ IND ist eine (strenge) Teil-von-Relation »*part*« vorzusehen, für die folgende Eigenschaften gelten:

**Irreflexivität:**  $\forall x: \text{IND} (\neg \text{part}(x, x))$

**Asymmetrie:**  $\forall x, y: \text{IND} (\text{part}(x, y) \rightarrow \neg \text{part}(y, x))$

**Transitivität:**  $\forall x, y, z: \text{IND} ((\text{part}(x, y) \wedge \text{part}(y, z)) \rightarrow \text{part}(x, z))$

Ferner gibt es die folgende Operation zur Summenbildung (» $\oplus$ «):

**Summenoperation  $\oplus$ :**

Für alle  $x, y, z: \text{IND}$  gibt es eine Summe  $x \oplus y$ , welche idempotent, kommutativ und assoziativ ist, d. h. es gilt  $x \oplus x = x$ ,  $x \oplus y = y \oplus x$  und  $x \oplus (y \oplus z) = (x \oplus y) \oplus z$ .

Zwei Elemente  $x, y: \text{IND}$  überlappen, wenn sie in der Überlappungsrelation (» $\otimes$ «) stehen:

**Überlappungsrelation  $\otimes$ :**

Für alle  $x, y: \text{IND}$  gilt:  $x \otimes y =_{\text{def}} \exists z: \text{IND} (z \oplus x = x \wedge z \oplus y = y)$ .

Zudem gilt folgendes Restprinzip:

**Restprinzip:**  $\forall x, y: \text{IND} (part(x, y) \rightarrow \exists! z: \text{IND} (\neg(x \otimes z) \wedge x \oplus z = y)).$

Als Abkürzung für  $part(x, y)$  wird  $x < y$  für  $x, y: \text{IND}$  verwendet.

Zu einer Klasse kann es speziellere Klassen geben. Für die spezielleren Klassen, die auch als Unterklassen oder Subklassen bezeichnet werden, gilt, dass ihre Instanzen auch Instanzen der Superklasse (Oberklasse) sind. Die Relation zwischen Unterklasse und Oberklasse wird als **is\_a**-Beziehung oder **Subsumption** bezeichnet:

**Subsumption:** Für  $A, B: \Lambda$  gilt:  $A \text{ is\_a } B =_{def} \forall x: \text{IND} (inst(x, A) \rightarrow inst(x, B)).$

Die Menge aller Klassen und die zwischen ihnen bestehenden *is\_a*-Beziehungen bilden eine Klassenhierarchie.

Häufig diskutiert wird die Frage, ob die Teil-von-Relation (bzw. ihre inverse Relation) angewandt auf Klassen immer transitiv ist. Ein Beispiel: Ein Arm hat eine Hand. Eine Hand hat Finger. Folgt nun: Hat ein Arm Finger? [Schulz, Hahn 2002, S. 387], [Winston 1987, S. 434].

Folgend [Schulz, Hahn 2002] definieren daher [Smith, Rosse 2004, S. 445] für Klassen die Relation **part\_of** basierend auf den Relationen **part\_for** und **has\_part**.

**part\_for:** Für  $A, B: \Lambda$  gilt:

$A \text{ part\_for } B =_{def} \forall x: \text{IND} (inst(x, A) \rightarrow \exists y: \text{IND} (inst(y, B) \wedge part(x, y)))$

d. h. Instanzen von  $A$  existieren nur als Teile von Instanzen von  $B$  oder anders ausgedrückt: Als notwendige Ganze für Instanzen von  $A$  fungieren Instanzen von  $B$ .

**has\_part:** Für  $A, B: \Lambda$  gilt:

$B \text{ has\_part } A =_{def} \forall y: \text{IND} (inst(y, B) \rightarrow \exists x: \text{IND} (inst(x, A) \wedge part(x, y)))$

d. h. Instanzen von  $B$  haben notwendig Instanzen von  $A$  als Teile.

**part\_of:** Für  $A, B: \Lambda$  gilt:

$A \text{ part\_of } B =_{def} A \text{ part\_for } B \wedge B \text{ has\_part } A$

$A \text{ part\_of } B$  verlangt also, dass zum einen Instanzen von  $A$  nur als Teile von Instanzen von  $B$  existieren und zum anderen  $B$  strukturell so organisiert ist, dass zu jeder Instanz von  $B$  eine Instanz von  $A$  existiert, die Teil der Instanz von  $B$  ist. Die Relation *part\_of* ist irreflexiv, asymmetrisch und transitiv. Bei geeigneter Formalisierung kann ein Arm aufgrund der Transitivität dann Finger haben.

## Notation

Um Klassennamen typografisch zu kennzeichnen, werden Kapitälchen verwendet. Instanzen werden grundsätzlich klein geschrieben.

### Ein Beispiel

Ein (fahrbereites) Snowboard hat eine Bindung, daher gilt auf Klassenebene sinnvollerweise: »SNOWBOARD *has\_part* BINDUNG«. Die Umkehrung »BINDUNG *part\_for* SNOWBOARD« gilt nicht von vornherein, da nicht jede Bindung Teil eines Snowboards ist – eine (Ski-)Bindung etwa ist Teil eines Skis. Betrachtet man die Unterklasse SNOWBOARD-BINDUNG mit »SNOWBOARD-BINDUNG *is\_a* BINDUNG«, dann gilt »SNOWBOARD *has\_part* SNOWBOARD-BINDUNG« und »SNOWBOARD-BINDUNG *part\_for* SNOWBOARD« und insgesamt folgt »SNOWBOARD-BINDUNG *part\_of* SNOWBOARD«.

### 3.2.2 Segmente in Film und Video

Für die Ausführungen dieses Kapitel werden über die Definitionen von Kapitel 2 hinausgehend weitere Begriffe und Definitionen für Videosegmente und (bestimmte) Teile von Videosegmenten (auch unterhalb der Einstellungsebene) benötigt.

Die nun eingeführten Begriffe orientieren sich am MPEG-7 Standard. Der MPEG-7 Standard, definiert von der *Moving Picture Expert Group* (MPEG), ist ein Metadatenschema zur Beschreibung und Annotation multimedialer Inhalte. Der Standard besteht aus 10 Teilen. Die folgenden Definitionen stammen aus dem Teil 5 »*Multimedia Description Scheme* (MDS)« [ISO/IEC 15938-5:2003], einer generischen Bibliothek von Beschreibungsstrukturen für Multimediadaten. Die Syntax und Semantik der einzelnen Einheiten wird durch sogenannte »*description schemes*« beschrieben. MPEG-7 basiert auf XML-Schema. Die *description schemes* sind im Schema als komplexe Datentypen definiert und in einer Typhierarchie organisiert. Diese XML-Schema-Typhierarchie ist entsprechend der Klassenhierarchie der damit beschriebenen Konzepte (*is\_a*-Relation) organisiert. Abb. 1 zeigt den für die folgenden Ausführungen relevanten Ausschnitt der Typhierarchie als UML-Klassendiagramm.

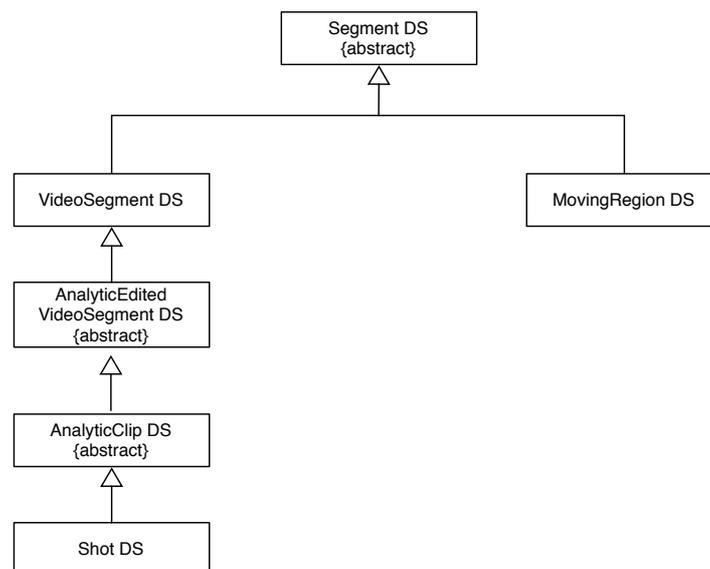


Abb. 3.1: Typhierarchie von MPEG-7 (Ausschnitt) dargestellt als UML-Klassendiagramm

An oberster Stelle der Hierarchie steht das *description scheme* Segment. Segment ist ein abstrakter Datentyp<sup>1</sup> und enthält die Attribute (z. B. Titel, Ersteller, Erstellungsdatum), die für alle von ihm abgeleiteten Typen gültig sind. Davon abgeleitet sind unter anderem VideoSegment<sup>2</sup> und MovingRegion. Ein VideoSegment ist ein beliebiges Fragment von Videodaten. Dies kann ein einzelner Frame, eine Folge von beliebig vielen Frames oder sogar die ganze Videosequenz sein. Es wird nicht verlangt, dass ein VideoSegment zeitlich zusammenhängend ist. Eine MovingRegion wiederum beschreibt einen raum-zeitlichen Bereich von Videodaten in einem VideoSegment, also eine beliebige Menge von Pixeln in einer beliebigen Folge von Frames. Eine MovingRegion kann also ein Pixel in einem Frame, einen ganzen Frame oder auch ein ganzes VideoSegment umfassen. Für eine MovingRegion wird nicht verlangt, dass sie räumlich oder zeitlich zusammenhängend ist. Abb. 3.2 visualisiert den Zusammenhang zwischen VideoSegment und MovingRegion (Quelle: [ISO/IEC 15938-5:2003], S. 252, Figure 25, eigene Ergänzungen).

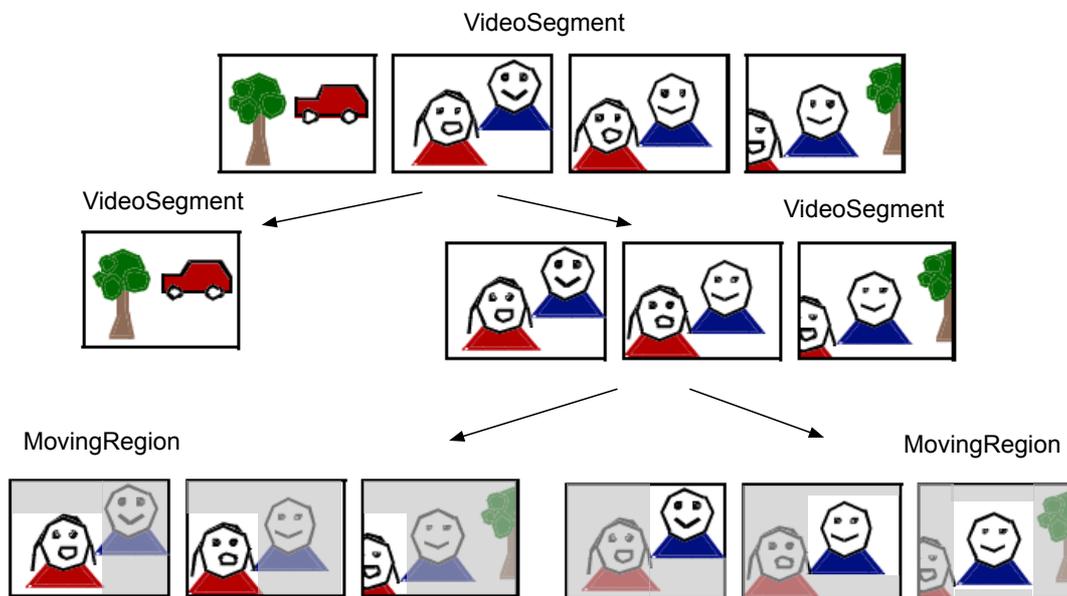


Abb. 3.2: Zusammenhang zwischen VideoSegment und MovingRegion

Eine Einstellung (*shot*) ist eine Folge von Frames, die durch eine Blende/Schnitt begrenzt wird. In den folgenden Ausführungen werden nur VideoSegmente im obigen Sinne, die Einstellungen umfassen, betrachtet. Mit dem Großbuchstaben »S« (»Große Syntagmatik für große Segmente«) werden speziell solche Einstellungsmengen, die wenigstens eine Einstellung enthalten, bezeichnet. Diese Definition von VideoSegment ist verträglich mit Definition 2.3, in der ein Segment (eines cinematographischen Dokumentes) als »nichtleere Teilfolge in der layoutierten zeitlichen Folge aller Einstellungen für einen gegebenen Layoutprozess« definiert wurde.

Ein Beobachter sieht i. Allg. eine Einstellung nicht als eine isolierte Einheit, sondern versucht Verbindungen v. a. zwischen benachbarten Einstellungen zu erkennen. Dazu orientiert er sich an

<sup>1</sup> Von einem abstrakten Typ können andere Typen abgeleitet, jedoch keine Instanzen erzeugt werden.

<sup>2</sup> Normterme werden eingedeutscht und in »camel case«-Schreibweise verwendet.

Pixelmengen, die in benachbarten Einstellungen für ihn identisch sind oder zumindest gemeinsame Teilbereiche haben. Nach dem MPEG-7 Vokabular sind dies MovingRegions. Diese stellen für den Zuschauer also visuelle Anker dar, die die Kohärenz (s. 2.2.3) eines Films gewährleisten können. In Abb. 3.2 kann jede Figur als MovingRegion erkannt und demzufolge als eine potentiell kohärenzerzeugende Region von einem Beobachter genutzt werden. MovingRegions werden im Folgenden mit dem Kleinbuchstaben  $s$  bezeichnet.

Häufig wird nur der Teil einer MovingRegion betrachtet, der in einer einzelnen Einstellung enthalten ist. Der Einfachheit halber wird dieser Teil als »ShotRegion« bezeichnet. Abb. 3.3 visualisiert den Zusammenhang zwischen MovingRegion, Einstellung (*shot*) und ShotRegion.

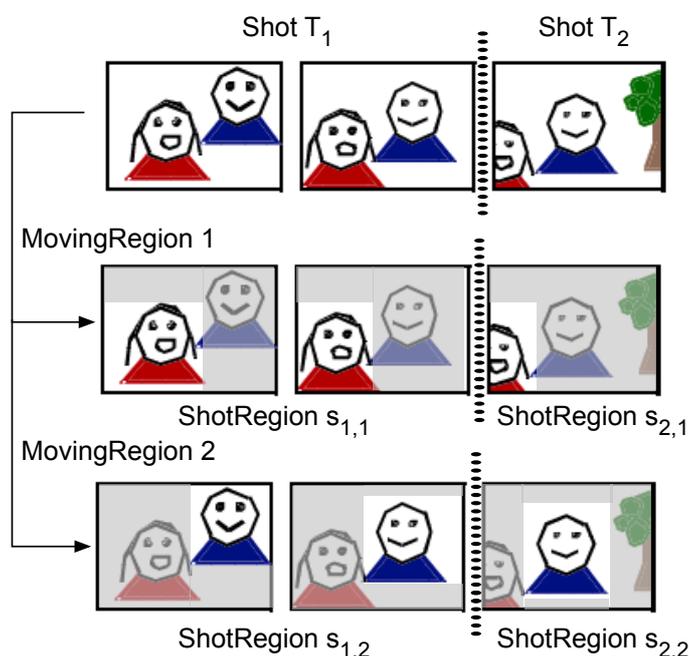


Abb. 3.3: Zusammenhang zwischen MovingRegion, Shot und ShotRegion

### Labelling

Betrachtet ein Beobachter  $B$  aus einer Beobachtermenge  $\mathcal{B}$  die Pixelmenge eines Bildes oder Films, versucht er aufgrund seiner Perzeption (auch unbewusst) in dieser Pixelmenge Objekte aus seiner Vorstellungswelt zu erkennen. Dies bedeutet, dass er diese Pixelmengen, genauer gesagt die MovingRegions, mit Begriffen in Zusammenhang bringt und »benennt«. Dieser Prozess wird im Folgenden als »Labelling« bezeichnet und die verwendete Benennung des Beobachters als »Label«.

Formal betrachtet ist ein Labelling einer MovingRegion  $s$  immer eine Instanziierung<sup>1</sup> einer Labelklasse  $L:\Lambda$ . Bei der Vergabe eines Labels erfolgt eine Instanziierung. Soll die Beobachtermenge  $\mathcal{B}$ , die eine solche Instanziierung zum Labelling durchführt, explizit gekennzeichnet werden, schreibt man » $inst_B$ «, für einen einzelnen Beobachter » $inst_B$ «,  $B \in \mathcal{B}$ .

<sup>1</sup> Der Beobachter sieht ein (konkretes) Objekt oder eine (konkrete) Person.

Weiterhin kann davon ausgegangen werden, dass auch für die Labelklassen eine *is\_a*-Hierarchie vorliegt, die analog zu einer Begriffshierarchie einer diegetischen Welt aufgebaut ist.

Wird eine MovingRegion gelabelt, so verfügen i. Allg. die ShotRegions, die Teil dieser MovingRegion sind, über das gleiche Label. Daher wird im Folgenden davon ausgegangen, dass ein Labelling der ShotRegions der Einstellungen eines Segmentes erfolgt. Allerdings kann auch folgende Situation eintreten: Zu Beginn eines Segmentes, o. B. d. A. in Einstellung 1, erkennt eine Beobachtermenge in einer MovingRegion z. B. »eine Frau«, diese wird während der Darstellungszeit, o. B. d. A. innerhalb der Einstellung 2, als »Johanna« identifiziert. In diesem Fall wird die ShotRegion der ersten Einstellung mit »eine Frau« gelabelt, die der zweiten Einstellung vollständig mit »Johanna«. Dies stellt für die Modellierung im Weiteren keine wesentliche Einschränkung dar, das Labelling der MovingRegion muss nur mit dem Labelling der ShotRegion »verträglich« sein. In diesem Beispiel ist dies der Fall, da  $inst_{\mathcal{B}}(\text{eineFrau}, \text{FRAU})$  und  $inst_{\mathcal{B}}(\text{Johanna}, \text{FRAU})$  gilt. Die Folge dieser Vorgehensweise ist, dass hier auf diese Weise Erkenntnisprozesse auf Einstellungsebene granularisiert werden, was aber für die folgenden Ausführungen ausreicht.

Zur Beschreibung einer ShotRegion  $s$  bedarf es für die Beobachtermenge  $\mathcal{B}$  einer Beobachtungssprache. Diese kann beliebig reichhaltig sein. Um jedoch die Problematiken der natürlichen Sprachen, die durch den Gebrauch von Synonymen und durch Mehrdeutigkeiten von Benennungen entstehen, zu vermeiden, gelten in diesem Kontext folgende Beschränkungen:

- (1) Es ist für dieselbe ShotRegion  $s$  und zwei verschiedene Label  $l_s \neq l'_s$  die gleichzeitige Gültigkeit von  $inst_{\mathcal{B}}(l_s, L)$  und  $inst_{\mathcal{B}}(l'_s, L)$  für ein  $L: \Lambda$  ausgeschlossen. Für eine ShotRegion gebe es nur ein Label für eine gegebene Labelklasse (»unique labels«), Synonyme sind nicht erlaubt.
- (2) Wenn die Label für zwei Segmente gleich sind, sind auch die von den Segmenten eines Filmbildes denotierten Objekte für die Beobachter ebenfalls gleich. Die gleich gelabelten Segmente können natürlich verschieden sein, insbesondere in zwei verschiedenen Einstellungen (semantische Eindeutigkeit [Goodman 1995, S. 144]).

Zur formalen Beschreibung des Labellings gelten nun folgende **Bezeichnungen**:

Sei  $\mathcal{L} = \{L_i \mid L_i: \Lambda, i \in \mathbb{N}\}$  eine Menge von Labelklassen.  $\mathcal{L}$  wird im Weiteren kurz als »Labelmenge« bezeichnet. Sei  $S(\mathcal{L}, \mathcal{B})$  die Menge der mittels  $\mathcal{L}$  durch die Beobachtermenge  $\mathcal{B}$  gelabelten ShotRegions  $s$  in  $S$ , also:

$$S(\mathcal{L}, \mathcal{B}) = \{s^l \mid s^l < S \text{ mit } l: \text{IND und } inst_{\mathcal{B}}(l, L) \text{ für ein } L \in \mathcal{L}\}.$$

Für ein Segment, bestehend aus  $n$  Einstellungen  $T_i, 1 \leq i \leq n$ , gibt es also für jede Einstellung  $i_k$  ShotRegions  $s_{i,j}, 1 \leq j \leq i_k, i_k \in \mathbb{N}$ , und somit für jede Einstellung  $i_k$  für jede Labelklasse eindeutige Label  $l_{i,j}: \text{IND}$ .

### Beispiel 3.1

Als Beispiel für ein Labelling wird das Segment  $S_{W_{02}} = (T09, T10)$  aus dem Beispielvideo »Snowboard« betrachtet. Syntagmatisch gesehen ist dies eine Szene (s. 2.5.3). Sie zeigt eine häufig vorkommende Montageform: Die zweite Einstellung zeigt die Fortsetzung der ersten, allerdings aus einer anderen Perspektive und in einer anderen Einstellungsgröße. Der Akteur, die Bretter auf dem Tisch, die Klebepistole, das Gewicht und der Hintergrund stellen MovingRegions für die Beobachter dar. Die Beobachter labeln die ShotRegions mit Instanzen zu  $\mathcal{L} = \{\text{AKTEUR, 3BRETTER, WERKZEUG, GEGENSTAND, HINTERGRUND, REST}\}$  (s. Abb. 3.4), so dass gilt:

$$S(\mathcal{L}, \mathcal{B}) = \{s^l \mid s^l < (T09, T10) \text{ mit genau einem } l: \text{IND und } inst_{\mathcal{B}}(l, L) \text{ für ein } L \in \{\text{AKTEUR, 3BRETTER, WERKZEUG, GEGENSTAND, HINTERGRUND, REST}\}\}$$

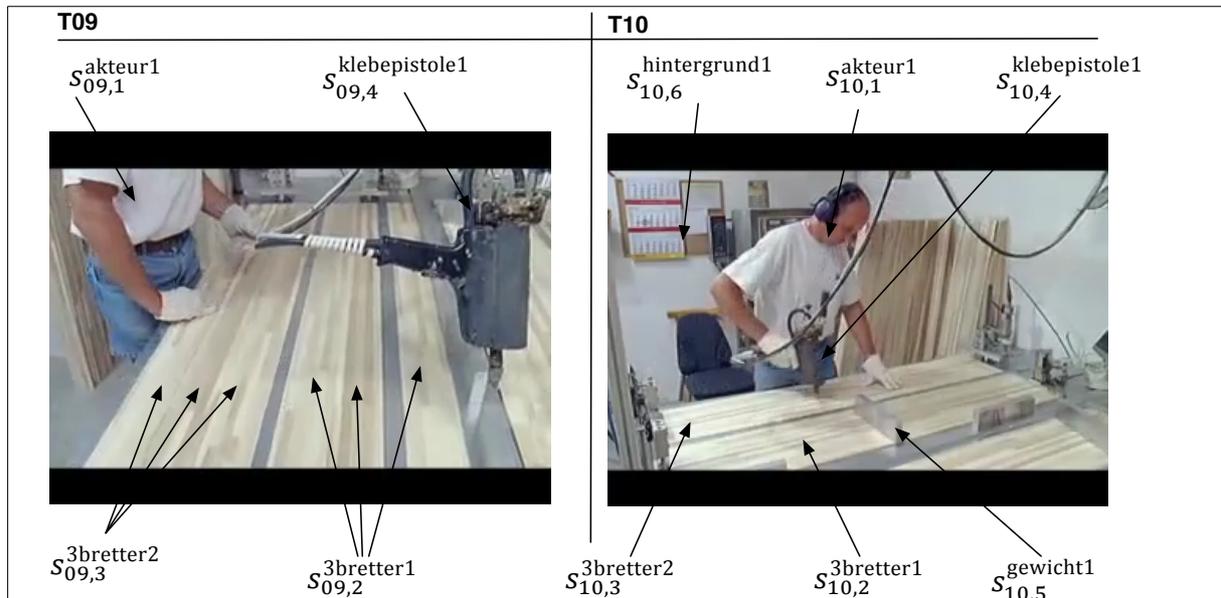


Abb. 3.4: Labelling des Segmentes  $S_{W_{02}} = (T09, T10)$

In beiden Einstellungen ist ein Akteur zu sehen. Zwar ist in  $T09$  nur ein kleiner Ausschnitt dieser Person zu erkennen, dennoch kann ein Beobachter im Default annehmen (auch weil er die vorhergehenden Einstellungen gesehen hat), dass es sich in beiden Einstellungen um den gleichen

Akteur handelt. Daher liegen gelabelte ShotRegions  $s_{09,1}^{\text{akteur1}} < T09$  und  $s_{10,1}^{\text{akteur1}} < T10$  mit  $inst_{\mathcal{B}}(\text{akteur1}, \text{AKTEUR})$  vor.

Es werden zweimal drei Bretter bereitgelegt, so dass in  $T09$  die entsprechend gelabelten ShotRegions  $s_{09,2}^{3\text{bretter1}}$  und  $s_{09,3}^{3\text{bretter2}}$ ,  $inst_{\mathcal{B}}(3\text{bretter1}, 3\text{BRETTER})$  und  $inst_{\mathcal{B}}(3\text{bretter2}, 3\text{BRETTER})$ , vorliegen. Im Default ist anzunehmen, dass dieselben Bretter auch in Einstellung  $T10$  zu sehen sind, daher werden diese ShotRegions ebenfalls mit »3bretter1« bzw. »3bretter2« gelabelt, so dass die ShotRegions  $s_{10,2}^{3\text{bretter1}} < T10$  und  $s_{10,3}^{3\text{bretter2}} < T10$  existieren.

In beiden Einstellungen erkennen die Beobachter eine Klebepistole und labeln die ShotRegions mit »klebepistole1« mit  $inst_{\mathcal{B}}(\text{klebepistole1}, \text{KLEBEPISTOLE})$  und  $\text{KLEBEPISTOLE } is\_a \text{ WERKZEUG}$ , so dass die ShotRegions  $s_{09,4}^{\text{klebepistole1}} < T09$  und  $s_{10,4}^{\text{klebepistole1}} < T10$  vorliegen.

In der Einstellung  $T09$  füllen Akteur, Bretter und die Klebepistole bis auf einen Restbereich  $s_{09,5}^{\text{rest1}}$ ,  $inst_{\mathcal{B}}(\text{rest1}, \text{REST})$ , das ganze Filmbild aus, daher gilt:

$$T09 = s_{09,1}^{\text{akteur1}} \oplus s_{09,2}^{3\text{bretter1}} \oplus s_{09,3}^{3\text{bretter2}} \oplus s_{09,4}^{\text{klebepistole1}} \oplus s_{09,5}^{\text{rest1}}$$

In Einstellung  $T10$  ist noch ein Gewicht zu sehen. Diese ShotRegion wird daher mit »gewicht1«,  $inst_{\mathcal{B}}(\text{gewicht1}, \text{GEWICHT})$  und  $\text{GEWICHT } is\_a \text{ GEGENSTAND}$ , gelabelt. In dieser Einstellung ist die Kameraeinstellung »HN«, dadurch wird für die Beobachter auch ein Ausschnitt der Werkstatt sichtbar, diesen labeln die Beobachter mit »hintergrund1« mit  $inst_{\mathcal{B}}(\text{hintergrund1}, \text{HINTERGRUND})$ , daher folgt mit dem Restbereich  $s_{10,7}^{\text{rest2}}$ ,  $inst_{\mathcal{B}}(\text{rest2}, \text{REST})$ :

$$T10 = s_{10,1}^{\text{akteur1}} \oplus s_{10,2}^{3\text{bretter1}} \oplus s_{10,3}^{3\text{bretter2}} \oplus s_{10,4}^{\text{klebepistole1}} \oplus s_{10,5}^{\text{gewicht1}} \oplus s_{10,6}^{\text{hintergrund1}} \oplus s_{10,7}^{\text{rest2}}$$

Insgesamt gilt somit für das Segment  $S_{W\_02}$ :

$$S_{W\_02} = (T09, T10) = (s_{09,1}^{\text{akteur1}} \oplus s_{09,2}^{3\text{bretter1}} \oplus s_{09,3}^{3\text{bretter2}} \oplus s_{09,4}^{\text{klebepistole1}} \oplus s_{09,5}^{\text{rest1}}, \\ s_{10,1}^{\text{akteur1}} \oplus s_{10,2}^{3\text{bretter1}} \oplus s_{10,3}^{3\text{bretter2}} \oplus s_{10,4}^{\text{klebepistole1}} \oplus s_{10,5}^{\text{gewicht1}} \oplus s_{10,6}^{\text{hintergrund1}} \\ \oplus s_{10,7}^{\text{rest2}})$$

Ein Labelling ist immer von der jeweiligen Beobachtermenge abhängig. So könnte eine andere Beobachtermenge alle sechs Bretter zusammen als eine ShotRegion konzeptionalisieren und damit würden die gelabelten ShotRegions  $s_{09,2}^{6\text{bretter1}} < T09$  und  $s_{10,2}^{6\text{bretter1}} < T10$  mit  $inst_{\mathcal{B}}(6\text{bretter1}, 6\text{BRETTER})$  vorliegen. In diesem Fall würde für das Segment  $S_{W\_02}$  gelten:

$$S_{W\_02} = (T09, T10) = (s_{09,1}^{\text{akteur1}} \oplus s_{09,2}^{6\text{bretter1}} \oplus s_{09,3}^{\text{klebepistole1}} \oplus s_{09,4}^{\text{rest1}}, \\ s_{10,1}^{\text{akteur1}} \oplus s_{10,2}^{6\text{bretter1}} \oplus s_{10,3}^{\text{klebepistole1}} \oplus s_{10,4}^{\text{gewicht1}} \oplus s_{10,5}^{\text{hintergrund1}} \oplus s_{10,6}^{\text{rest2}})$$

### 3.3 Thematische Relationen und ihre Verankerung im Filmbild

Bisher wurde lediglich eine Repräsentation zur Konzeptionalisierung statischer Objekte wie »Akteur«, »Bretter« oder »Werkzeug« entwickelt. Dieser und die folgenden Abschnitte behandeln die Repräsentation und formale Darstellung von Ereignissen bzw. Handlungen in filmischen Segmenten. Dazu wird auf Definitionen und Ergebnisse aus der Linguistik (insbesondere [Krifka 1989a], [Krifka 1989b], [Mollá-Aloid 1997]) zurückgegriffen, die dann übertragen werden.

Unter einem **Ereignis** versteht man eine Zustandsänderung, für die es einen Ort und eine Zeitdauer oder einen Zeitpunkt gibt. Ereignisse, die durch ein belebtes Objekt absichtsvoll ausgelöst werden, nennen wir **Handlungen**. Das auslösende Objekt heißt **Agent**, auch Akteur [Reimer 1991, S. 21f].

In Produktionsprozessen werden im Zuge der Automatisierung vermehrt Arbeitsschritte mit Hilfe einer Maschine oder sogar von einer Maschine allein, ohne weitere direkte Einflussnahme eines Menschen, ausgeführt. Daher werden im Folgenden auch Ereignisse, die von Maschinen ausgelöst werden, als Handlungen bezeichnet. Ebenso wird eine auslösende Maschine als Agent angesehen.

Die in Verben durchgeführte Repräsentation von Ereignissen modelliert [Krifka 1989a] mit einer Ereignis-Semantik. In dieser werden Verben als einstellige Prädikate über Ereignisse dargestellt. So entspricht z. B. das Prädikat *essen* der Menge der *Essens*-Ereignisse oder das Prädikat *fräsen* der Menge der *Fräsen*-Ereignisse. Die Partizipanten eines Ereignisses werden mit zweistelligen Hilfsrelationen mit dem Ereignis verbunden. Diese Hilfsrelationen entsprechen semantischen Rollen<sup>1</sup> wie Agens (AG) für den Verursacher einer Handlung oder Patiens (PAT) für den Betroffenen einer Handlung. [Krifka 1989b, S. 238] bezeichnet diese Relationen als **thematische Relationen**. Für den Beispielsatz: »Akteur1 fräst ein Brett« ergibt sich<sup>2</sup>:

$$\exists e (\text{fräsen}(e) \wedge \text{AG}(\text{akteur1}, e) \wedge \text{PAT}(\text{brett}, e)).$$

Da Ereignisklassen sich von anderen Konzeptklassen deutlich unterscheiden, wird hier für Ereignisklassen ein eigener Typ, *EVT*, eingeführt. Mit  $\mathcal{E}$  wird eine Menge von Ereignisklassen bezeichnet:  $\mathcal{E} = \{E \mid E: EVT\}$ .

In einer cinematographischen Realisation sieht man keine natürlich-sprachlichen Sätze, sondern Bildsegmente. Eine Beobachtermenge erkennt z. B. eine Pixelmenge als »Brett« und labelt sie mit »brett1«,  $\text{inst}_{\mathcal{B}}(\text{brett1}, \text{BRETT})$ . Zudem erkennt sie, dass das mit »brett1«, gelabelte Segment erkennbar verändert wird, z. B. wird ein Teil abgesägt (z. B. in Einstellung T37 im Beispielvideo),

<sup>1</sup> Krifka versucht nicht, die thematischen Relationen zu charakterisieren. [Fillmore 1968] liefert hierzu einen Ansatz [Mollá-Aliod 1997, S. 67]. [Bußmann 2002, S. 697] merkt an: »Die Bezeichnung, Zahl und inhaltliche Bestimmung der T.R. sind in der Forschung ein ständig neu aufgegriffenes und kontroverses Thema.«

<sup>2</sup> Die Argumente der Relationen werden in dieser Arbeit in umgekehrter Reihenfolge als z. B. bei [Krifka 1989a] notiert.

oder das Brett wird um weitere Teile ergänzt (z. B. in Einstellung *T17* im Beispielvideo). Des Weiteren konzeptualisiert sie durch diese Veränderungen bestimmte Handlungen, wie etwa »(ab)sägen« oder »Metallteile montieren«.

Zur Modellierung thematischer Beziehungen zwischen Segmenten aus  $S(\mathcal{L}, \mathcal{B})$  und instanziierten Ereignissen von  $E$ ,  $E \in \mathcal{E}$ , für eine von den Beobachtern zu konzeptionalisierende Ereignismenge  $\mathcal{E}$  werden hier **segment-thematische Relationen**  $R_{S(\mathcal{L}, \mathcal{B}), \mathcal{E}}$  mit

$$R_{S(\mathcal{L}, \mathcal{B}), \mathcal{E}} = \{(s^l, e) | s^l \in S(\mathcal{L}, \mathcal{B}), \text{inst}_{\mathcal{B}}(e, E), E: EVT, E \in \mathcal{E}\}$$

eingeführt.

Die ShotRegions sind qua Zugehörigkeit zu  $S(\mathcal{L}, \mathcal{B})$  gelabelt. Wenn das Label im jeweiligen Zusammenhang unerheblich ist, wird – um die Notation einfach zu halten – nur  $s$  statt  $s^l$  geschrieben. Ist die Menge der Ereignisse und die Menge der gelabelten ShotRegions  $S(\mathcal{L}, \mathcal{B})$  unerheblich oder aus dem Kontext ersichtlich, wird nur  $R$  verwendet.

Auf Basis der segment-thematischen Relation ergeben sich die folgenden Definitionen:

Steht eine Ereignisklasse in einer segment-thematischen Relation zu einem Segment, spricht man von einer Verankerung der Ereignisklasse im Segment, genauer definiert:

**Definition 3.1a:** Eine Ereignisklasse  $E: EVT$  wird in einem VideoSegment  $S$  von einer Beobachtermenge  $\mathcal{B}$  **unter  $R$**  für eine vorgegebene Labelmenge  $\mathcal{L}$  **verankert** oder  **$R$ -verankert**, wenn es ein  $s < S$  gibt, so dass  $R_{S(\mathcal{L}, \mathcal{B}), \{E\}}$  nicht leer ist.

Das Wort »Verankerung« ist mit Bedacht gewählt, da sich die in diesem Sinne thematische Verknüpfung von Segmenten von Pixelmengen leicht als (ggf. mehrfacher) »Link« interpretieren lässt (s. auch Abb. 3.6).

**Definition 3.1b:** Eine Ereignisklasse  $E: EVT$  wird in einem VideoSegment  $S$  von einer Beobachtermenge  $\mathcal{B}$  **unter  $R$**  für eine vorgegebene Labelmenge  $\mathcal{L}$  **mehrfach verankert** oder **mehrfach  $R$ -verankert**, wenn es für  $E \in \mathcal{E}$  Instanzen  $e', e''$  gibt, sowie ein  $s'$  mit  $(s', e') \in R_{S(\mathcal{L}, \mathcal{B}), \{E\}}$  und ein  $s''$  mit  $(s'', e'') \in R_{S(\mathcal{L}, \mathcal{B}), \{E\}}$ , so dass  $e', e''$  sich nicht überlappen, also  $\neg(e' \otimes e'')$  gilt.

Die Definition einer zählenden mehrfachen » $m$ -fachen« Verankerung für  $m \geq 2$  ist kanonisch:

**Definition 3.1c:** Eine Ereignisklasse  $E: EVT$  wird in einem VideoSegment  $S$  von einer Beobachtermenge  $\mathcal{B}$  **unter  $R$**  für eine vorgegebene Labelmenge  $\mathcal{L}$   **$m$ -fach verankert** oder  **$m$ -fach- $R$ -verankert**, wenn es für  $E \in \mathcal{E}$  Instanzen  $e^{(1)}, e^{(2)}, \dots, e^{(m)}$  mit  $e^{(i)} \neq e^{(j)}$  für  $1 \leq i, j \leq m$  gibt und jeweils ein  $s^{(i)}$  mit  $(s^{(i)}, e^{(i)}) \in R_{S(\mathcal{L}, \mathcal{B}), \{E\}}$  und ein  $s^{(j)}$  mit  $(s^{(j)}, e^{(j)}) \in R_{S(\mathcal{L}, \mathcal{B}), \{E\}}$  gibt, wobei  $e^{(i)}$  und  $e^{(j)}$  sich nicht überlappen, also  $\neg(e^{(i)} \otimes e^{(j)})$  gilt.

In der Definition wird nicht verlangt, dass sich die Segmente  $s^{(i)}$  nicht überlappen, daher kann  $s^{(i)} \otimes s^{(j)}$  gelten für  $1 \leq i, j \leq m$ . Dagegen hat die Anforderung, dass sich die Ereignisse paarweise nicht überlappen, folgenden Hintergrund: Sind z. B.  $e^{(1)}$  und  $e^{(2)}$  zwei verschiedene Instanzen von  $E$ , also  $inst_{\mathcal{B}}(e^{(1)}, E)$ ,  $inst_{\mathcal{B}}(e^{(2)}, E)$  und  $e^{(1)} \neq e^{(2)}$  und es gelte  $\neg(e^{(1)} \otimes e^{(2)})$ , dann ist die Ereignisklasse  $E$  2-mal in einem geeigneten Segment verankert bei Erfüllung der sonstigen Bedingungen der Definition. Schließt man aber die Überlappung von Ereignissen nicht aus und gelte  $e^{(1)} < f$  und  $e^{(2)} < f$  für  $inst_{\mathcal{B}}(f, F)$ ,  $E < F$  mit  $F: EVT$ , dann würde für die Ereignisklasse  $F$  folgen, dass diese 2-mal verankert ist, sie sollte jedoch nur 1-mal verankert sein, da es sich konzeptionell nur um eine Instanz handelt.

Für cinematographische Dokumente ist im Rahmen einer großen Syntagmatik der Fall wesentlich, dass ein Ereignis in wenigstens zwei verschiedenen Einstellungen verankert ist: Dazu wird zusätzlich verlangt, dass wenigstens zwei der geforderten Teilsegmente  $s'$  und  $s''$  auch in zwei verschiedenen Einstellungen  $T'$  und  $T''$  liegen, also  $s' < T'$  und  $s'' < T''$  gilt. Eine einzelne Ereignisklasse  $E: EVT$  wird in einem Segment  $S$  mit wenigstens zwei verschiedenen Einstellungen  $T'$  und  $T''$  von einer Beobachtermenge  $\mathcal{B}$  für eine vorgegebene Labelmenge  $\mathcal{L}$  in verschiedenen Einstellungen  $R$ -verankert, wenn für  $E \in \mathcal{E}$  eine (ggf. mehrfache) Verankerung vorliegt und es für Instanzen  $e', e''$  (im mehrfachen Falle mit  $e' \neq e''$  und  $\neg(e' \otimes e'')$ ) ein  $s' < T'$  mit  $(s', e') \in R_{S(\mathcal{L}, \mathcal{B}), \{E\}}$  und ein  $s'' < T''$  mit  $(s'', e'') \in R_{S(\mathcal{L}, \mathcal{B}), \{E\}}$  gibt. Damit ergibt sich:

**Definition 3.1d:** Eine Ereignisklasse  $E: EVT$  wird in einem VideoSegment  $S$  von einer Beobachtermenge  $\mathcal{B}$  **unter  $R$  für eine vorgegebene Labelmenge  $\mathcal{L}$  in  $n$  Einstellungen ( $m$ -fach) verankert** oder **in  $n$  Einstellungen ( $m$ -fach)  $R$ -verankert**, wenn sie in jeder der  $n$  Einstellungen wenigstens einmal (und insgesamt  $m$ -fach) unter  $R$  verankert ist für  $n, m \geq 1$ .

Wird die Anzahl der Verankerungen gezählt, wird im Weiteren kurz von einer **( $n, m$ )- $R$ -Verankerung** der jeweiligen Ereignisklasse in dem jeweiligen VideoSegment gesprochen.

Liegt in einem VideoSegment mit  $n$  Einstellungen eine  $(n, n)$ -Verankerung einer Ereignisklasse vor, bedeutet dies, dass die Beobachter in jeder Einstellung genau einmal diese Ereignisklasse konzeptionalisieren.

Ein wichtiger Anwendungsfall für die obige Definition ist der, dass man sich alle Einstellungen eines cinematographischen Dokumentes für eine Ereignisklasse »zusammensucht«.

**Definition 3.1e:** Sei  $D$  ein cinematographisches Dokument,  $E: EVT$  eine Ereignisklasse. Die Menge aller Einstellungen, in denen  $E$   $R$ -verankert ist, wird als  **$R$ -Träger von  $E$  in  $D$**  bezeichnet.

**Beispiel 3.2**

Betrachtet wird, wie in Beispiel 3.1, das Segment  $S_{W_{02}} = (T09, T10)$  aus dem Beispielvideo »Snowboard«. Im Segment  $S_{W_{02}}$  werden folgende Handlungen konzeptionalisiert (s. Abb. 3.5, Zeile 1 des Bildtextes): In Einstellung  $T09$  nimmt ein Akteur (akteur1) zweimal drei Bretter vom Stapel und legt diese auf den Tisch (»BRETTER\_LEGEN«). Der Akteur greift zur Klebepistole, gibt Kleber auf die ersten drei Bretter (»KLEBER\_AUFTRAGEN«). Dann folgt ein Schnitt und in  $T10$  trägt der Akteur wieder Kleber auf die ersten drei Bretter auf. Anschließend legt er ein Gewicht auf die ersten drei Bretter (»GEWICHT\_AUFLEGEN«) und trägt Kleber auf die zweiten drei Bretter auf.

<b>T09</b>				
				
1	bretter legen_1		bretter legen_2	
2	bretter_vom_Stapel- _nehmen_1	bretter_zurecht legen_1	bretter_vom_Stapel- _nehmen_2	bretter_zurecht legen_2
3	bretter_.....			

<b>T10</b>				
				
1	kleber_auftragen_1	kleber_auftragen_i	gewicht_auflegen_1	kleber_auftragen_j
2	kleber_auftragen_1	kleber_auftragen_i	gewicht_auflegen_1	kleber_auftragen_j
3	....._zusammenkleben_1			

Abb. 3.5: Handlungen in  $S_{W_{02}} = (T09, T10)$

Es werden also Instanzen der Ereignisklassen BRETTER\_LEGEN, KLEBER\_AUFTRAGEN und GEWICHT\_AUFLEGEN konzeptionalisiert. Bei allen konzeptionalisierten Instanzen ist akteur1 der Agent. Für die Konzeptionalisierung der Instanzen von KLEBER\_AUFTRAGEN sind beobachterseits auf Basis des Filmmaterials die in Abb. 3.6 dargestellten Fälle möglich.

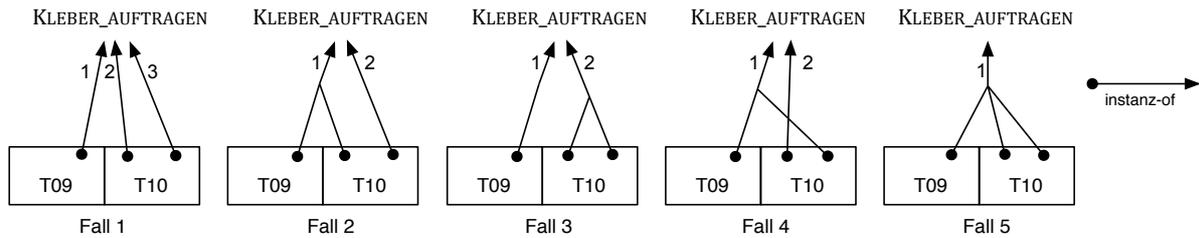


Abb. 3.6: Mögliche Fälle für die Konzeptionalisierung der Instanzen von KLEBER\_AUFTRAGEN

*Fall 1:* Ein Beobachter konzeptionalisiert drei verschiedene, sich nicht überlappende Instanzen der Ereignisklasse KLEBER\_AUFTRAGEN: kleber\_auftragen\_1 in T09, kleber\_auftragen\_2 und kleber\_auftragen\_3 in T10.

*Fall 2:* Ein Beobachter konzeptionalisiert zwei verschiedene, sich nicht überlappende Instanzen der Ereignisklasse KLEBER\_AUFTRAGEN. Die Instanz kleber\_auftragen\_1 beginnt in T09 und wird in T10 fortgesetzt. Da zwischen den Einstellungen keine zeitliche Lücke vorhanden ist, der Akteur am gleichen Bretterstapel weiterarbeitet und ein Beobachter erkennen kann, dass die Fortsetzung der Handlung lediglich aus einer anderen Kameraperspektive gezeigt wird, kann ein Beobachter die Handlung, trotz des Schnittes, als eine Instanz konzeptionalisieren. Die nach der Handlung gewicht\_auflegen\_1 konzeptionalisierte Instanz von KLEBER\_AUFTRAGEN wird als eine neue, zweite Instanz, kleber\_auftragen\_2, wahrgenommen, auch weil hier ein anderes Objekt, 3Bretter2, von der Handlung betroffen ist.

*Fall 3:* Ein Beobachter konzeptionalisiert zwei verschiedene, sich nicht überlappende Instanzen der Ereignisklasse KLEBER\_AUFTRAGEN, davon eine Instanz in T09 und die zweite Instanz in T10.

*Fall 4:* Ein Beobachter konzeptionalisiert eine Instanz der Ereignisklasse KLEBER\_AUFTRAGEN in T09. In T10 konzeptionalisiert er eine zweite Instanz und dann in dieser Einstellung eine Fortsetzung der ersten Instanz.

*Fall 5:* Ein Beobachter konzeptionalisiert nur eine Instanz der Ereignisklasse KLEBER\_AUFTRAGEN. Ob das Auftragen des Klebers auf die ersten oder zweiten der drei Bretter erfolgt, ist für ihn nicht von Bedeutung. Die Instanz der Ereignisklasse GEWICHT\_AUFLEGEN wird vom Beobachter lediglich als eine Unterbrechung der Handlung KLEBER\_AUFTRAGEN wahrgenommen.

An dieser Stelle sieht man sehr schön, dass die Zerlegung einer Handlung in Teilhandlungen oder die Annahme mehrerer Handlungen gleichen Typs vom Filmbild überhaupt nicht nahe gelegt werden muss und oft eine reine Beobachterleistung ist, die sich in der weiteren Interpretation bewähren muss.

Für die segment-thematische Relation  $AG_{S(\mathcal{L},\mathcal{B}),\mathcal{E}}$  gilt für alle fünf Fälle:

$$\begin{aligned}
 AG_{S(\mathcal{L},\mathcal{B}),\mathcal{E}} &= \{(s, e) \mid s \in S(\mathcal{L}, \mathcal{B}), inst_{\mathcal{B}}(e, E), E \in \mathcal{E}\} \\
 &= \{(s, e) \mid s \in S(\{\text{PERSON}\}, \mathcal{B}), inst_{\mathcal{B}}(e, E), \\
 &\quad E \in \{\text{BRETTER\_LEGEN}, \text{KLEBER\_AUFTRAGEN}, \text{GEWICHT\_AUFLEGEN}\}\} \\
 &= \{(s_{09,1}^{\text{akteur1}}, \text{bretter\_legen\_1}), (s_{09,1}^{\text{akteur1}}, \text{bretter\_legen\_2}), \\
 &\quad (s_{09,1}^{\text{akteur1}}, \text{kleber\_auftragen\_1}), (s_{10,1}^{\text{akteur1}}, \text{kleber\_auftragen\_i}), \\
 &\quad (s_{10,1}^{\text{akteur1}}, \text{gewicht\_auflegen\_1}), (s_{10,1}^{\text{akteur1}}, \text{kleber\_auftragen\_j}) \\
 &\quad \mid inst_{\mathcal{B}}(\text{akteur1}, \text{PERSON}), \\
 &\quad inst_{\mathcal{B}}(\text{bretter\_legen\_1}, \text{BRETTER\_LEGEN}), \\
 &\quad inst_{\mathcal{B}}(\text{bretter\_legen\_2}, \text{BRETTER\_LEGEN}), \\
 &\quad inst_{\mathcal{B}}(\text{kleber\_auftragen\_1}, \text{KLEBER\_AUFTRAGEN}), \\
 &\quad inst_{\mathcal{B}}(\text{kleber\_auftragen\_i}, \text{KLEBER\_AUFTRAGEN}), \\
 &\quad inst_{\mathcal{B}}(\text{kleber\_auftragen\_j}, \text{KLEBER\_AUFTRAGEN}), \\
 &\quad inst_{\mathcal{B}}(\text{gewicht\_auflegen\_1}, \text{GEWICHT\_AUFLEGEN})\}
 \end{aligned}$$

mit Fall 1:  $i = 2, j = 3$ ; Fall 2:  $i = 1, j = 2$ ; Fall 3:  $i = 2, j = 2$ ;

Fall 4:  $i = 2, j = 1$ ; Fall 5:  $i = 1, j = 1$ .

Betroffene der konzeptionalisierten Ereignisklassen BRETTER\_LEGEN, KLEBER\_AUFTRAGEN und GEWICHT\_AUFLEGEN sind 3bretter1 bzw. 3bretter2. Es gilt für die segment-thematische Relation  $PAT_{S(\mathcal{L},\mathcal{B}),\mathcal{E}}$ :

$$\begin{aligned}
 PAT_{S(\mathcal{L},\mathcal{B}),\mathcal{E}} &= \{(s, e) \mid s \in S(\mathcal{L}, \mathcal{B}), inst_{\mathcal{B}}(e, E), E \in \mathcal{E}\} \\
 &= \{(s, e) \mid s \in S(\{\text{3BRETTER}\}, \mathcal{B}), inst_{\mathcal{B}}(e, E), \\
 &\quad E \in \{\text{BRETTER\_LEGEN}, \text{KLEBER\_AUFTRAGEN}, \text{GEWICHT\_AUFLEGEN}\}\} \\
 &= \{(s_{09,2}^{\text{3bretter1}}, \text{bretter\_legen\_1}), (s_{09,3}^{\text{3bretter2}}, \text{bretter\_legen\_2}), \\
 &\quad (s_{10,2}^{\text{3bretter1}}, \text{kleber\_auftragen\_1}), (s_{10,2}^{\text{3bretter1}}, \text{kleber\_auftragen\_i}), \\
 &\quad (s_{10,2}^{\text{3bretter1}}, \text{gewicht\_auflegen\_1}), (s_{10,3}^{\text{3bretter2}}, \text{kleber\_auftragen\_j}) \\
 &\quad \mid inst_{\mathcal{B}}(\text{3bretter1}, \text{3BRETTER}), inst_{\mathcal{B}}(\text{3bretter2}, \text{3BRETTER}), \\
 &\quad inst_{\mathcal{B}}(\text{bretter\_legen\_1}, \text{BRETTER\_LEGEN}), \\
 &\quad inst_{\mathcal{B}}(\text{bretter\_legen\_2}, \text{BRETTER\_LEGEN}), \\
 &\quad inst_{\mathcal{B}}(\text{kleber\_auftragen\_1}, \text{KLEBER\_AUFTRAGEN}), \\
 &\quad inst_{\mathcal{B}}(\text{kleber\_auftragen\_i}, \text{KLEBER\_AUFTRAGEN}), \\
 &\quad inst_{\mathcal{B}}(\text{kleber\_auftragen\_j}, \text{KLEBER\_AUFTRAGEN}), \\
 &\quad inst_{\mathcal{B}}(\text{gewicht\_auflegen\_1}, \text{GEWICHT\_AUFLEGEN})\}
 \end{aligned}$$

mit Fall 1:  $i = 2, j = 3$ ; Fall 2:  $i = 1, j = 2$ ; Fall 3:  $i = 2, j = 2$ ;

Fall 4:  $i = 2, j = 1$ ; Fall 5:  $i = 1, j = 1$ .

Die Beziehung einer Handlung zu ihren Partizipanten sowie die Beziehungen zwischen Klassen und Instanzen können in Form eines semantischen Netzes dargestellt werden. Abb. 3.8 zeigt dies beispielhaft für das Ereignis `kleber_auftragen_1` und die thematischen Rollen (bzw. segment-thematischen Relationen) Agens und Patiens.

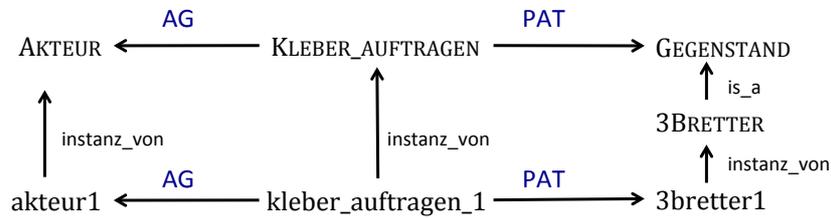


Abb. 3.7: Semantisches Netz: AG, PAT-Relationen zur Instanz `kleber_auftragen_1`

Zum Auftragen des Klebers auf die Bretter verwendet der Akteur eine Klebepistole. Bei dieser Handlung gibt es daher neben den thematischen Relationen AG und PAT eine weitere Relation, die die Beziehung der Handlung zu der Klebepistole ausdrückt. Diese lässt sich als thematische Relation *Instrument*<sup>1</sup> (INST) charakterisieren. Auch die Instanz `gewicht1` steht in der thematischen Relation INST zur Handlung `gewicht_auflegen_1`.

Daher gilt für die segment-thematische Relation  $INST_{S(\mathcal{L},\mathcal{B}),\mathcal{E}}$ :

$$\begin{aligned}
 INST_{S(\mathcal{L},\mathcal{B}),\mathcal{E}} &= \{(s, e) \mid s \in S(\mathcal{L}, \mathcal{B}), inst_{\mathcal{B}}(e, E), E \in \mathcal{E}\} \\
 &= \{(s, e) \mid s \in S(\{\text{WERKZEUG}, \text{GEGENSTAND}\}, \mathcal{B}), \\
 &\quad inst_{\mathcal{B}}(e, E), E \in \{\text{KLEBER\_AUFTRAGEN}, \text{GEWICHT\_AUFLEGEN}\}\} \\
 &= \{(s_{09,4}^{\text{klebepistole1}}, \text{kleber\_auftragen}_1), (s_{09,4}^{\text{klebepistole1}}, \text{kleber\_auftragen}_i), \\
 &\quad (s_{10,5}^{\text{gewicht1}}, \text{gewicht\_auflegen}_1), (s_{10,4}^{\text{klebepistole1}}, \text{kleber\_auftragen}_j) \\
 &\quad \mid inst_{\mathcal{B}}(\text{klebepistole1}, \text{KLEBEPISTOLE}), \text{KLEBEPISTOLE } is\_a \text{ WERKZEUG}, \\
 &\quad inst_{\mathcal{B}}(\text{kleber\_auftragen}_1, \text{KLEBER\_AUFTRAGEN}), \\
 &\quad inst_{\mathcal{B}}(\text{kleber\_auftragen}_i, \text{KLEBER\_AUFTRAGEN}), \\
 &\quad inst_{\mathcal{B}}(\text{kleber\_auftragen}_j, \text{KLEBER\_AUFTRAGEN}), \\
 &\quad inst_{\mathcal{B}}(\text{gewicht1}, \text{GEWICHT}), \text{GEWICHT } is\_a \text{ GEGENSTAND}, \\
 &\quad inst_{\mathcal{B}}(\text{gewicht\_auflegen}_1, \text{GEWICHT\_AUFLEGEN})\}
 \end{aligned}$$

mit Fall 1:  $i = 2, j = 3$ ; Fall 2:  $i = 1, j = 2$ ; Fall 3:  $i = 2, j = 2$ ;

Fall 4:  $i = 2, j = 1$ ; Fall 5:  $i = 1, j = 1$ .

<sup>1</sup> »die Rolle des unbelebten Verursachers einer Handlung bzw. des Objektes mit dessen Hilfe eine Handlung vollzogen wird« [Bußmann 2002, S. 333]

Abb. 3.8 zeigt nun das um die thematische Rolle INST erweiterte semantische Netz zur Instanz kleber\_auftragen\_1.

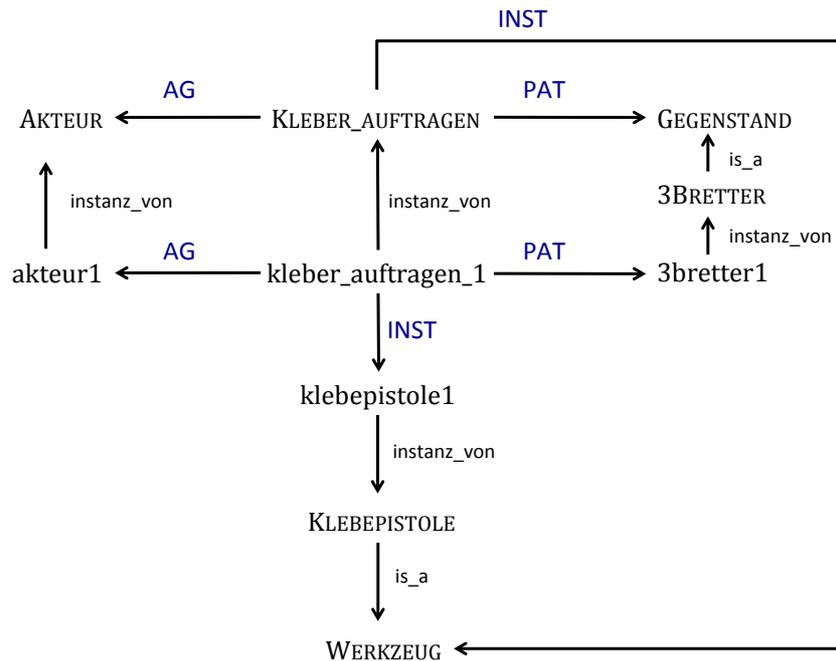


Abb. 3.8: Semantisches Netz: AG, PAT, INST-Relationen zur Instanz kleber\_auftragen\_1

Aus den oben dargestellten segment-thematischen Relationen lässt sich direkt die Verankerung der verschiedenen Ereignisklassen im Segment  $S_{W_{02}} = (T09, T10)$  ablesen. Es gilt:

- BRETTGERÄTE\_LEGEN: (1,1)-AG-verankert, (1,2)-PAT-verankert,
- GEWICHT\_AUFLEGEN: (1,1)-verankert für die Relationen AG, PAT und INST,
- KLEBER\_AUFTRAGEN:
  - Fall 1: (2,2)-AG-verankert, (2,3)-PAT-verankert, (2,2)-INST-verankert,
  - Fall 2: (2,2)-AG-verankert, (2,3)-PAT-verankert, (2,2)-INST-verankert,
  - Fall 3: (2,2)-AG-verankert, (2,3)-PAT-verankert, (2,2)-INST-verankert,
  - Fall 4: (2,2)-AG-verankert, (2,3)-PAT-verankert, (2,2)-INST-verankert,
  - Fall 5: (2,1)-AG-verankert, (2,1)-PAT-verankert, (2,1)-INST-verankert.

Bei der Ereignisklasse KLEBER\_AUFTRAGEN unterscheiden sich die Fälle 1 bis 4 nicht hinsichtlich ihrer AG-, PAT- und INST-Verankerung, obwohl die Anzahl der Instanzierungen im Fall 1 verschieden ist zu der in den Fällen 2 bis 4. In den Fällen 1 bis 4 ist in den beiden Einstellungen  $T09$  und  $T10$  die handelnde Person zweimal verankert, das behandelte Objekt dreimal verankert. Im Unterschied dazu finden sich im Fall 5 für die AG-, PAT- und INST-Relation definitionsgemäß in beiden Einstellungen keine mehrfachen Verankerungen, da nur eine Instanzierung des Ereignisses durchgeführt wird. Somit wird die Ereignisklasse nur einmal verankert.

Aus den segment-thematischen Relationen können auch die verschiedenen Träger der Ereignisklassen im Segment  $S_{W\_02}$  abgelesen werden. Hier gilt:

- BRETTER\_LEGEN: (T09) für die Relationen AG und PAT,
- GEWICHT\_AUFLEGEN: (T10) für die Relationen AG, PAT und INST,
- KLEBER\_AUFTRAGEN: (T09, T10) für die Relationen AG, PAT und INST.

Da diese Ereignisklassen nur in diesem Segment vorkommen, sind die obigen Träger auch die Träger für das gesamte Dokument.

Wie bereits für das Labelling in Beispiel 3.1 ausgeführt, ist jede Konzeptionalisierung von der Beobachtermenge abhängig. Die bishere Betrachtung geht davon aus, dass die Beobachter die Handlungen relativ detailliert erfassen. Es wäre auch eine noch detailliertere Sicht möglich, wie in Abb. 3.5, Zeile 2 des Bildtextes dargestellt. So könnten statt der Handlung BRETTER\_LEGEN auch zwei (Teil-)Handlungen BRETTER\_VOM\_STAPEL\_NEHMEN und BRETTER\_ZURECHT\_LEGEN konzeptionalisiert werden.

Eine weitere Beobachtermenge,  $B'$ , könnte dagegen eine viel gröbere Sicht haben (s. Abb. 3.5, Zeile 3 des Bildtextes). Für sie geht es in dieser Szene lediglich um die Handlung BRETTER\_ZUSAMMENKLEBEN. In diesem Fall gilt für die segment-thematischen Relationen AG, PAT und INST:

$$\begin{aligned}
 AG_{S(\mathcal{L}, \mathcal{B}'), \mathcal{E}} &= \{(s, e) \mid s \in S(\mathcal{L}, \mathcal{B}'), inst_{\mathcal{B}'}(e, E), E \in \mathcal{E}\} \\
 &= \{(s, e) \mid s \in S(\{\text{PERSON}\}, \mathcal{B}'), \\
 &\quad inst_{\mathcal{B}'}(e, E), E \in \{\text{BRETTER\_ZUSAMMENKLEBEN}\}\} \\
 &= \{(s_{09,1}^{\text{akteur1}}, \text{bretter\_zusammenkleben\_1}), \\
 &\quad (s_{10,1}^{\text{akteur1}}, \text{bretter\_zusammenkleben\_1}) \\
 &\quad \mid inst_{\mathcal{B}'}(\text{akteur1}, \text{PERSON}), \\
 &\quad inst_{\mathcal{B}'}(\text{bretter\_zusammenkleben\_1}, \text{BRETTER\_ZUSAMMENKLEBEN})\}
 \end{aligned}$$

$$\begin{aligned}
 PAT_{S(\mathcal{L}, \mathcal{B}'), \mathcal{E}} &= \{(s, e) \mid s \in S(\mathcal{L}, \mathcal{B}'), inst_{\mathcal{B}'}(e, E), E \in \mathcal{E}\} \\
 &= \{(s, e) \mid s \in S(\{\text{6BRETTER}\}, \mathcal{B}'), \\
 &\quad inst_{\mathcal{B}'}(e, E), E \in \{\text{BRETTER\_ZUSAMMENKLEBEN}\}\} \\
 &= \{(s_{09,2}^{\text{6bretter1}}, \text{bretter\_zusammenkleben\_1}), \\
 &\quad (s_{10,2}^{\text{6bretter1}}, \text{bretter\_zusammenkleben\_1}) \\
 &\quad \mid inst_{\mathcal{B}'}(\text{6bretter1}, \text{6BRETTER}), \\
 &\quad inst_{\mathcal{B}'}(\text{bretter\_zusammenkleben\_1}, \text{BRETTER\_ZUSAMMENKLEBEN})\}
 \end{aligned}$$

$$\begin{aligned}
 \text{INST}_{S(\mathcal{L}, \mathcal{B}'), \mathcal{E}} &= \{(s, e) \mid s \in S(\mathcal{L}, \mathcal{B}'), \text{inst}_{\mathcal{B}'}(e, E), E \in \mathcal{E}\} \\
 &= \{(s, e) \mid s \in S(\{\text{WERKZEUG}, \text{GEGENSTAND}\}, \mathcal{B}'), \\
 &\quad \text{inst}_{\mathcal{B}'}(e, E), E \in \{\text{KLEBER\_AUFTRAGEN}, \text{GEWICHT\_AUFLEGEN}\}\} \\
 &= \{(s_{09,3}^{\text{klebepistole1}}, \text{bretter\_zusammenkleben\_1}), \\
 &\quad (s_{10,3}^{\text{klebepistole1}}, \text{bretter\_zusammenkleben\_1}), \\
 &\quad (s_{10,4}^{\text{gewicht1}}, \text{bretter\_zusammenkleben\_1}) \\
 &\quad \mid \text{inst}_{\mathcal{B}'}(\text{klebepistole1}, \text{KLEBEPISTOLE}), \text{KLEBEPISTOLE is}_a \text{ WERKZEUG}, \\
 &\quad \text{inst}_{\mathcal{B}'}(\text{gewicht1}, \text{GEWICHT}), \text{GEWICHT is}_a \text{ GEGENSTAND}, \\
 &\quad \text{inst}_{\mathcal{B}'}(\text{bretter\_zusammenkleben\_1}, \text{BRETTER\_ZUSAMMENKLEBEN})\}
 \end{aligned}$$

### 3.4 Objekt-Eindeutigkeit segment-thematischer Relationen

In der Linguistik ist nach [Krifka 1989b, S. 241] eine thematische Relation  $R$  **objekt-eindeutig**, wenn jedes Ereignis  $e$  höchstens zu einem Objekt  $x$  in der  $R$ -Beziehung steht. Ein Beispiel liegt bei der Handlung »ein Glas Wein trinken« vor. Das Ereignis »trinken« ist über die PAT-Relation mit diesem Glas Wein verbunden und mit keinem anderen. Nicht objekt-eindeutig ist dagegen das Ereignis »ein Haus sehen«, da das Sehen des Hauses auch das Sehen eines davorstehenden Baumes, der nicht Teil des Hauses ist, miteinschließen kann. Nach [Mollá-Aloid 1997, S. 116] liegt die Eigenschaft der Objekt-Eindeutigkeit immer dann vor, wenn ein Objekt erstellt oder zerstört wird. Allerdings kann sie auch vorliegen, wenn keine dieser beiden Situationen der Fall ist, z. B. beim Ereignis »ein Buch lesen«. Zu beachten ist allerdings, dass beim erneuten Lesen des Buches ein neues Ereignis vorliegt.

Die segment-thematischen Relationen  $R_{S(\mathcal{L}, \mathcal{B}), \mathcal{E}}$ , wie  $\text{AG}_{S(\mathcal{L}, \mathcal{B}), \mathcal{E}}$  und  $\text{PAT}_{S(\mathcal{L}, \mathcal{B}), \mathcal{E}}$ , beschreiben die Beziehungen zwischen Teilsegmenten  $s < S$  und allen im VideoSegment  $S$  von einer Beobachtermenge  $\mathcal{B}$  konzeptionalisierten Ereignisse  $E \in \mathcal{E}$ . Für die folgenden Überlegungen wird davon ausgegangen, dass sich die Beobachtermenge auf eine einzelne Ereignisklasse  $E$  fokussiert, d. h. die segment-thematischen Relationen werden auf  $R_{S(\mathcal{L}, \mathcal{B}), \{E\}}$  bzw.  $\text{AG}_{S(\mathcal{L}, \mathcal{B}), \{E\}}$  und  $\text{PAT}_{S(\mathcal{L}, \mathcal{B}), \{E\}}$  eingeschränkt.

Überträgt man obige Definition der Objekt-Eindeutigkeit auf filmische Repräsentationen, ist eine segment-thematische Relation  $R$  objekt-eindeutig, wenn die zu einem Ereignis in der  $R$ -Relation stehende ShotRegion eindeutig ist. Daher wird definiert:

**Definition 3.2a:** Eine segment-thematische Relation  $R_{S(\mathcal{L}, \mathcal{B}), \{E\}}$  ist **für eine Einstellung  $T$**  und eine Ereignisklasse  $E$  für eine Beobachtermenge  $\mathcal{B}$  **objekt-eindeutig**, wenn für ein  $e$  mit  $\text{inst}_{\mathcal{B}}(e, E)$  und  $s, s' < T$  mit  $(s, e) \in R_{S(\mathcal{L}, \mathcal{B}), \{E\}}$  und  $(s', e) \in R_{S(\mathcal{L}, \mathcal{B}), \{E\}}$  die Gleichheit  $l_s = l_{s'}$  folgt.

In der Definition wird nicht die Gleichheit der Segmente, sondern die Gleichheit der Label gefordert. Diese Anforderung ist für den Film passend, da sich, z. B. durch Schwenk oder Zoom, die Pixelmenge für »ein- und dasselbe Objekt« schon innerhalb einer Einstellung ändern kann. Aus der Gleichheit der Label folgt, da semantische Eindeutigkeit vorausgesetzt wird, die Gleichheit der Objekte in der Diegese.

### Beispiel 3.3:

Wird für die segment-thematische Relation  $PAT_{S(\mathcal{L},\mathcal{B}),\{\text{KLEBER\_AUFTRAGEN}\}}$  in der Einstellung  $T_{10}$  (s. Beispiel 3.1, 3.2) Objekt-Eindeutigkeit unterstellt, dann folgt, dass die gezeigten Bretter als EIN Objekt konzeptionalisiert werden.

Objekt-Eindeutigkeit für ein VideoSegment  $S$  heißt damit, dass für ein konzeptionalisiertes Ereignis  $E$  und für die gleiche Beobachtermenge  $\mathcal{B}$  mit gleicher Menge von Labelklassen  $\mathcal{L}$  die zugehörigen Segmente in allen Einstellungen gleich gelabelt werden und deshalb in deren Diegese das gleiche Objekt referenzieren.

**Definition 3.2b:** Eine segment-thematische Relation  $R_{S(\mathcal{L},\mathcal{B}),\{E\}}$  ist **für ein VideoSegment**  $S = (T_1, T_2, \dots, T_n)$  für eine Beobachtermenge  $\mathcal{B}$  **objekt-eindeutig**, wenn sie in den  $n$  Einstellungen jeweils objekt-eindeutig und die zugehörigen Segmente in den sie verankernden Einstellungen jeweils gleich gelabelt sind:

$$\forall s_i, s_j, T_i, T_j < S \text{ mit } s_i < T_i, s_j < T_j \text{ und } (s_i, e), (s_j, e) \in R_{S(\mathcal{L},\mathcal{B}),\{E\}} \text{ mit } i \neq j \text{ gilt:}$$

$$l_{s_i} = l_{s_j} \text{ für alle } 1 \leq i, j \leq n.$$

## 3.5 Relationierung von Subobjekten und Subereignissen

In den linguistischen Theoriebildungen finden sich zwei weitere Konzepte: die Subobjekt-Abbildbarkeit (d. h. ein Ereignis ist auf Subobjekte abbildbar) und die Subereignis-Abbildbarkeit (d. h. ein Objekt ist auf Subereignisse abbildbar).<sup>1</sup>

Eine thematische Relation  $R$  heißt **subereignis-abbildbar**, wenn gilt: »Falls ein Ereignis und ein Objekt in der Relation zueinander stehen, so entspricht jedem Teil des Objektes ein Teil des Ereignisses, sodaß diese Teile ebenfalls in Relation zueinander stehen.« [Krifka 1989b, S. 241]. Das Ereignis »ein Glas Wein trinken« ist subereignis-abbildbar, da jeder Teil des Glases Wein auf einen Teil des Ereignisses »trinken« abgebildet werden kann (s. Abb. 3.9, nach [Mollá-Aliod 1997, S. 120]).

Eine thematische Relation  $R$  ist **subobjekt-abbildbar**, wenn »es bei Bestehen einer thematischen Relation zu jedem Teil eines Ereignisses einen Teil des Objektes gibt, sodaß die Teile in derselben

<sup>1</sup> Bei [Krifka 1989a] lediglich: Objekt-Abbildbarkeit bzw. Ereignis-Abbildbarkeit, engl: *mapping to objects, mapping to events*.

thematischen Relation zueinander stehen« [Krifka 1989b, S. 241]. Ein Beispiel ist wiederum das Ereignis »ein Glas Wein trinken«, da jedes Teilereignis von »trinken« auf einen Teil des Glases Wein abgebildet wird (s. Abb. 3.9<sup>1</sup>).

### Visualisierung durch Raum-Zeit-Diagramme

Zur Visualisierung der Beziehung zwischen Ereignis und Objekt schlägt [Krifka 1989a, S. 159f] Raum-Zeit-Diagramme vor (wie es sich in Abb. 3.9 fast schon zeigt).

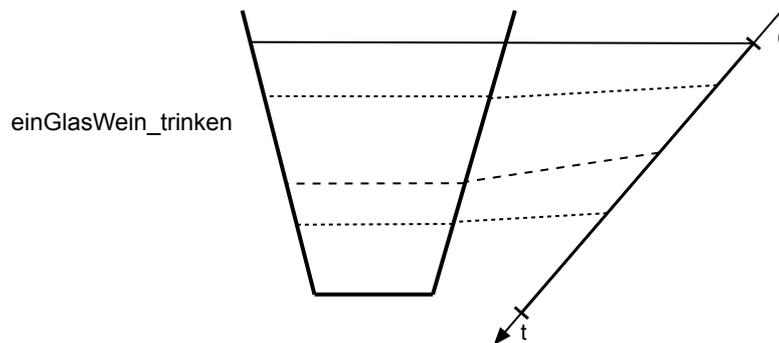


Abb. 3.9: Subereignis-Abbildbarkeit und Subobjekt-Abbildbarkeit von »ein Glas Wein trinken«

In einem Raum-Zeit-Diagramm repräsentiert eine Koordinate den Raum, die andere die Zeit. Objekte werden als Linie bzw. Band, entsprechend ihrer räumlichen Ausdehnung, dargestellt. Ereignisse werden zeitlich lokalisiert und entsprechend ihrer zeitlichen Ausdehnung auf der Zeitachse abgebildet. Auf diese Weise wird das Zusammenwirken von Objekt und Ereignis visualisiert, insbesondere wird sichtbar, ob und wenn ja, wie sich die räumliche Größe/Ausdehnung im Zeitverlauf ändert. Abb. 3.10 zeigt zwei Raum-Zeit-Diagramme. Das linke Diagramm zeigt eine Situation, bei der das Objekt dem Ereignis nach und nach unterworfen wird und dabei verschwindet, wie dies z. B. bei »einen Apfel essen« der Fall ist. Das rechte Diagramm zeigt eine Situation, in der das Objekt wächst. Ein Beispiel hierzu ist »ein Haus bauen«.

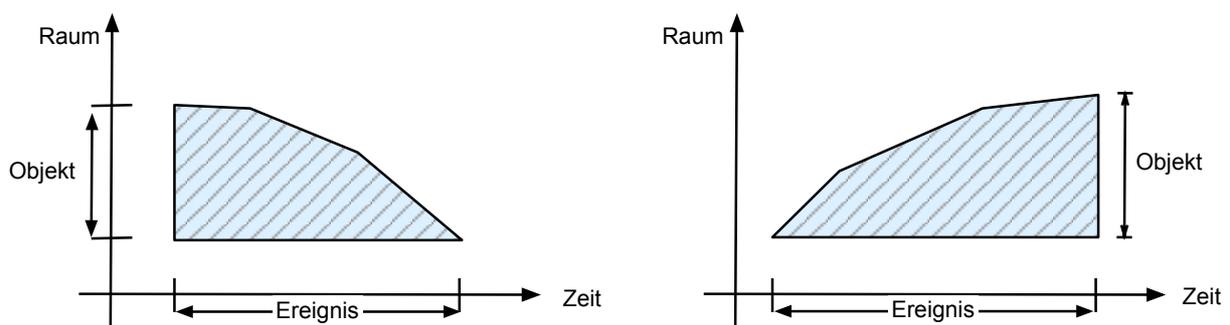


Abb. 3.10: Beispiele für Raum-Zeit-Diagramme

<sup>1</sup> Da »ein Glas Wein trinken« abbauend ist, muss der Nullpunkt der Zeitachse oben liegen.

### Anwendung auf cinematographische Dokumente

Für segment-thematische Relationen werden nun die Eigenschaften der Subereignis-Abbildbarkeit und der Subobjekt-Abbildbarkeit eingeführt. Letztere wird aus dem Spezialfall der Subsegment-Abbildbarkeit abgeleitet.

Eine segment-thematische Relation  $R$  ist auf ein Subereignis abbildbar oder subereignis-abbildbar, wenn beobachterseits zu jedem Teil eines gelabelten Segmentes ein Teil von  $e$  angenommen werden kann, für den wiederum  $R$  gilt. Es wird definiert:

**Definition 3.3:** Eine segment-thematische Relation  $R_{S(\mathcal{L},\mathcal{B}),\{E\}}$  ist für ein VideoSegment  $S$  eines cinematographischen Dokumentes für eine Beobachtermenge  $\mathcal{B}$  für eine vorgegebene Labelmenge  $\mathcal{L}$  **subereignis-abbildbar**, wenn aus  $(s, e) \in R_{S(\mathcal{L},\mathcal{B}),\{E\}}$  für ein  $s' < s < S$  mit  $l_s = l_{s'}$  folgt, dass ein  $e' < e$  existiert mit  $(s', e') \in R_{S(\mathcal{L},\mathcal{B}),\{E\}}$ .

Diese Definition gilt gleichermaßen für Vorgänge, die »abbauen« (z. B. »Apfel essen« ) oder »aufbauen« (z. B. »Haus bauen«).

Eine segment-thematische Relation  $R$  ist auf Subsegmente eines VideoSegmentes  $S$  abbildbar oder subsegment-abbildbar, wenn es für  $R$  zu jedem beobachterseits konzeptionalisierbaren Teil eines Ereignisses einen gleich gelabelten Teil eines gelabelten Segmentes in  $S$  gibt, das wiederum in der  $R$ -Beziehung zu dem Teilereignis steht. Definitorisch ergibt sich:

**Definition 3.4:** Eine segment-thematische Relation  $R_{S(\mathcal{L},\mathcal{B}),\{E\}}$  ist in einem VideoSegment  $S$  eines cinematographischen Dokumentes für eine Beobachtermenge  $\mathcal{B}$  für eine vorgegebene Labelmenge  $\mathcal{L}$  **subsegment-abbildbar**, wenn aus  $(s, e) \in R_{S(\mathcal{L},\mathcal{B}),\{E\}}$  und  $e' < e$  folgt, dass es ein  $s' < s$  gibt mit  $(s', e') \in R_{S(\mathcal{L},\mathcal{B}),\{E\}}$  und  $l_s = l_{s'}$ .

Ist eine segment-thematische Relation subsegment-abbildbar, bedeutet dies also, dass für ein fortschreitendes Ereignis auch immer eine entsprechende Pixelmenge angegeben werden kann, in der dieser Fortschritt sichtbar ist. In dieser Definition finden sich offenbar zwei Beschränkungen. Für ein fortschreitendes Ereignis kann nur dann eine entsprechende Pixelmenge angegeben werden, wenn es sich um eine vollständige Aufzeichnung einer Handlung handelt. Es wird also der häufige Fall einer filmischen Raffung nicht berücksichtigt. Die zweite Beschränkung liegt darin, dass obige Definition  $s' < s$  verlangt, d. h. dass eine echte Verkleinerung der Segmente gefordert wird (die Relation »<« ist irreflexiv). Fälle, in denen sich Objekte der Diegese und die dementsprechende Pixelmenge im Filmbild in anderer Weise verändern, werden in dieser Definition nicht berücksichtigt.

Aus diesen Gründen muss nun eine allgemeinere Definition der Subobjekt-Abbildbarkeit entwickelt werden. Zur kompakten Darstellung wird dazu folgende Notation zusätzlich verwendet:

Treten mehr als zwei Veränderungen in Folge auf, werden zur Kennzeichnung Zahlen in Klammern verwendet:  $s_{i,(j)}^{\text{label}}$ .

Eine »veränderte« ShotRegion umfasst nicht notwendigerweise die gesamte Einstellungslänge, sie kann auch auf ein bestimmtes zeitliches Intervall eingegrenzt sein.

### Beispiel 3.4

Ein instruktives Beispiel zeigt Abb. 3.11, die einen Ausschnitt aus Einstellung  $T17$  enthält. Die Veränderungen an der mit »brett1« gelabelten ShotRegion sind deutlich erkennbar: Es wird jeweils ein Metallteil ergänzt. Im letzten Bild zoomt die Kamera weg, wodurch sich die fokussierte ShotRegion verkleinert.



Abb. 3.11: Beispiel für das Labelling bei veränderten ShotRegions

Für den fortschreitenden Vorgang des Ereignisses »X\_METALLTEILE\_MONTIEREN« kann jeweils eine entsprechende »Brett-Pixelmenge« in  $T17$  angegeben werden, wobei das Verhältnis der Segmente  $s_{17,(1)}^{\text{brett1}}$ ,  $s_{17,(2)}^{\text{brett1}}$ ,  $s_{17,(3)}^{\text{brett1}}$  und  $s_{17,(4)}^{\text{brett1}}$  zueinander nicht inkrementell im Sinne von  $s_{17,(1)}^{\text{brett1}} < s_{17,(2)}^{\text{brett1}} < s_{17,(3)}^{\text{brett1}} < s_{17,(4)}^{\text{brett1}}$  anzunehmen ist. Dies ist ein Beispiel dafür, dass die oben genannte zweite Beschränkung selbst dann zu restriktiv ist, wenn keine filmische Raffung vorliegt.

Bei einer filmischen Raffung wird allgemein auf eine vollständige Aufzeichnung verzichtet oder bei der filmischen Montage werden Teile einer Handlung herausgeschnitten. Dies hat zur Folge, dass es nicht mehr möglich ist, für jeden Teil einer Handlung ein passendes Segment anzugeben (s. Abb. 3.12).

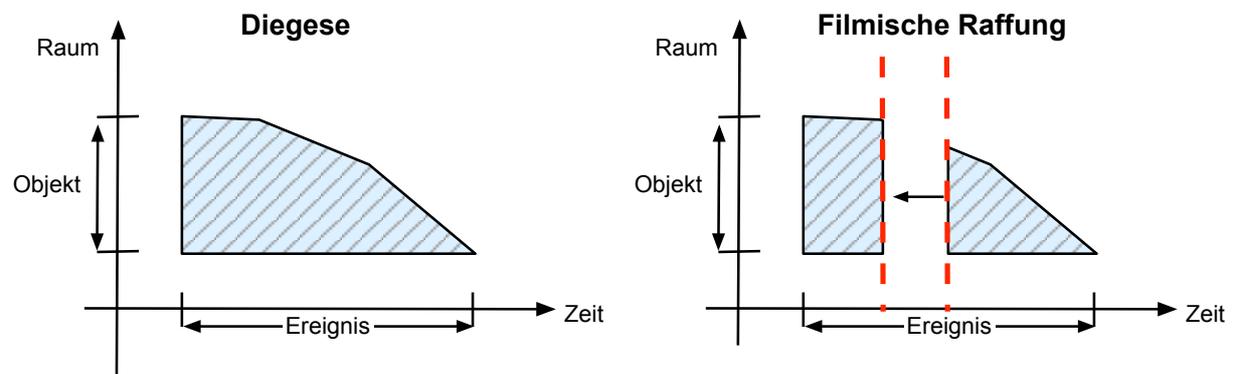


Abb. 3.12: Fehlendes Segment durch »Filmische Raffung«

An dieser Stelle setzt die filmische Repräsentation auf die Interpolationsfähigkeiten des Beobachters. Sichtbare Veränderungen im Filmbild lassen ihn Rückschlüsse auf die Veränderungen in der Diegese ziehen, insbesondere wenn er weitere Teilereignisse eines instanziierten Ereignisses erkennen kann.

Um dies zu modellieren, wird o. B. d. A. von folgender Situation ausgegangen: Es gebe  $k \geq 1$  gelabelte Segmente, die im Filmbild durch ein Ereignis  $e$  mit Teilereignissen  $e^{(v)} < e, 1 \leq v \leq k$ , sichtbar betroffen sind:

$$(s_{i_1,(1)}^{\text{label}}, e^{(1)}), (s_{i_2,(2)}^{\text{label}}, e^{(2)}), \dots, (s_{i_k,(k)}^{\text{label}}, e^{(k)}).$$

Für die  $k$  Teilereignisse gilt oft nicht  $e^{(1)} < e^{(2)} < \dots < e^{(k)}$ ; in filmischen Situationen gilt vielmehr häufig  $e^{(1)} \hookrightarrow e^{(2)} \hookrightarrow e^{(3)} \dots \hookrightarrow e^{(k)}$  im Sinne von »folgen aufeinander«.

Jedem Segment  $s_{i_v,(v)}^{\text{label}}, 1 \leq v \leq k$ , werde beobachterseits jeweils ein diegetisches Teilobjekt  $objekt_{(v)}$  zugeordnet. Dies geschieht durch die bereits in 2.2.2 eingeführte Abbildung *dieST*, die Segmente für eine Beobachtermenge in eine diegetische Raumzeit<sup>1</sup> abbildet:

$$\begin{aligned} dieST: \mathcal{B} \times Seg(D) &\rightarrow DieSpace \times DieTime \\ (B, S) &\mapsto dieS(B, S) \times dieT(B, S) \end{aligned}$$

Für Segmente  $s_{i_v,(v)}^{\text{label}}, 1 \leq v \leq k, B \in \mathcal{B}$ , gilt:

$$dieST(B, s_{i_v,(v)}^{\text{label}}) = objekt_{(v)} \subseteq DieSpace \times DieTime,$$

wobei für die räumliche Komponente  $dieS(B, s_{i_v,(v)}^{\text{label}}) = o_{(v)} \subseteq DieSpace$  und für das zeitliche Verhalten  $dieT(B, s_{i_v,(v)}^{\text{label}}) = t_{(v)} \subseteq DieTime$  gilt.

Insgesamt liegt skizziert die folgende Situation vor:

Teilereignis	$e^{(1)}$	$e^{(2)}$	...	$e^{(\mu-1)}$	$e^{(\mu)}$	$e^{(\mu+1)}$	...	$e^{(k)}$
Objekt der Diegese (räumliche Komponente)	$o_{(1)}$			$o_{(\mu-1)}$	$o_{(\mu)}$	$o_{(\mu+1)}$		$o_{(k)}$
Segment	$s_{i_1,(1)}^{\text{label}}$	$s_{i_2,(2)}^{\text{label}}$		$s_{i_{\mu-1},(\mu-1)}^{\text{label}}$	$s_{i_{\mu},(\mu)}^{\text{label}}$	$s_{i_{\mu+1},(\mu+1)}^{\text{label}}$		$s_{i_k,(k)}^{\text{label}}$

Aufgrund einer filmischen Raffung »fehle« das Segment  $s_{i_{\mu},(\mu)}^{\text{label}}$ . An der Stelle  $\mu$  schätze der Beobachter ein Teilereignis mit der zugehörigen räumlichen Komponente  $o_{(\mu)}$  eines Objektes  $objekt_{(\mu)}$ . Basis dieser Schätzung ist, dass wenigstens  $s_{i_{\mu-1},(\mu-1)}^{\text{label}}$  und/oder  $s_{i_{\mu+1},(\mu+1)}^{\text{label}}$  beobachtet werden und damit  $o_{\mu-1}$  und/oder  $o_{\mu+1}$  angenommen werden können.

<sup>1</sup> »universeller Bildraum für die Urbilder filmischer Messungen« [Schmidt 2008, S. 221]

Wenn nun aus  $(s_{i_{\mu-1},(\mu-1)}^{\text{label}}, e^{(\mu-1)}) \in R_{S(\mathcal{L},\mathcal{B}),\{E\}}$  oder  $(s_{i_{\mu+1},(\mu+1)}^{\text{label}}, e^{(\mu+1)}) \in R_{S(\mathcal{L},\mathcal{B}),\{E\}}$  folgt, dass beobachterseits jeweils auch ein Teilobjekt in der Diegese mit räumlicher Komponente  $o_\mu$  mit  $o_{\mu-1} \subseteq o_\mu$  oder  $o_\mu \subseteq o_{\mu+1}$  angenommen werden kann, dem ein fehlendes Segment  $\hat{s}_o$  im Filmbild entspricht, also  $dieS(\mathcal{B}, \hat{s}_o) = o_\mu$  mit  $s_{i_{\mu-1},(\mu-1)}^{\text{label}} < \hat{s}_o$  oder  $\hat{s}_o < s_{i_{\mu+1},(\mu+1)}^{\text{label}}$  und  $(\hat{s}_o, e^{(\mu)}) \in R_{S(\mathcal{L},\mathcal{B}),\{E\}}$ , dann liegt eine Form der Subobjekt-Abbildbarkeit vor.

Auch der Fall  $s_{i_{\mu-1},(\mu-1)}^{\text{label}} = \hat{s}_o$  oder  $\hat{s}_o = s_{i_{\mu+1},(\mu+1)}^{\text{label}}$  kann vorkommen. Dies ist der Fall, wenn in der Diegese ein Ereignis eintritt, bei dem im zeitlichen Ablauf im Filmbild keine räumlichen Veränderungen sichtbar sind, z. B. wenn ein (rotationssymmetrischer) Gegenstand rotiert wird. Hier sind alle geschätzten Segmente in der Diegese und auch die zugehörigen Segmente im Filmbild gleich groß. Gleichwohl kann man auch hier, wenn die zeitliche Komponente hinzugenommen wird, von Subobjekt-Abbildbarkeit sprechen: Jedem Teil des Ereignisses rotieren,  $inst_{\mathcal{B}}(\text{rotieren}, \text{ROTIEREN})$ , kann in der Diegese ein Stück in der Diegese zugeordnet werden.

Damit wird Subobjekt-Abbildbarkeit kompatibel zu Subsegment-Abbildbarkeit definiert, Subsegment-Abbildbarkeit ist somit ein »vollständiger« Spezialfall der Subobjekt-Abbildbarkeit.

**Definition 3.5:** Eine segment-thematische Relation  $R_{S(\mathcal{L},\mathcal{B}),\{E\}}$  ist in einem VideoSegment  $S$  eines cinematographischen Dokumentes von einer Beobachtermenge  $\mathcal{B}$  für eine vorgegebene Labelmenge  $\mathcal{L}$  **subobjekt-abbildbar**,

- wenn aus  $(s, e) \in R_{S(\mathcal{L},\mathcal{B}),\{E\}}$  und  $e' < e$  folgt, dass es tatsächlich ein  $s' < s$  gibt mit  $(s', e') \in R_{S(\mathcal{L},\mathcal{B}),\{E\}}$  und  $l_s = l_{s'}$ , so dass  $dieS(\mathcal{B}, s') = o' \subset o = dieS(\mathcal{B}, s)$  gilt, oder
- wenn aus  $(s_{i_{\mu-1},(\mu-1)}^{\text{label}}, e^{(\mu-1)}) \in R_{S(\mathcal{L},\mathcal{B}),\{E\}}$  oder  $(s_{i_{\mu+1},(\mu+1)}^{\text{label}}, e^{(\mu+1)}) \in R_{S(\mathcal{L},\mathcal{B}),\{E\}}$  folgt, dass beobachterseits jeweils auch ein Teilobjekt in der Diegese mit räumlicher Komponente  $o_\mu$  und  $o_{\mu-1} \subseteq o_\mu$  oder  $o_\mu \subseteq o_{\mu+1}$  so angenommen werden kann, dass von  $B \in \mathcal{B}$  ein fehlendes Segment  $\hat{s}_o$  mit  $s_{i_{\mu-1},(\mu-1)}^{\text{label}} < \hat{s}_o$  oder  $\hat{s}_o < s_{i_{\mu+1},(\mu+1)}^{\text{label}}$  oder  $s_{i_{\mu-1},(\mu-1)}^{\text{label}} = \hat{s}_o = s_{i_{\mu+1},(\mu+1)}^{\text{label}}$ , so dass  $dieS(\mathcal{B}, \hat{s}_o) = o_\mu$  und  $(\hat{s}_o, e^{(\mu)}) \in R_{S(\mathcal{L},\mathcal{B}),\{E\}}$  gilt.

Die in der Definition formulierte zweite Anforderung ist strenger als die bloße Annahme eines zusätzlichen diegetischen Teilobjektes: Die Sensordaten der abgebildeten Segmente liefern die Interpolationsbasis für die nicht abgebildeten Teile des Objektes in der Diegese, deren filmische Abbildung aber als möglich angenommen werden kann. Das Weglassen eines Teilsegmentes wird somit tatsächlich als filmische Raffung modelliert.

## Gradualität

Damit sind alle Mittel vorhanden, um eine filmische Repräsentation gradueller Handlungen auszuzeichnen. [Krifka 1989b, S. 245] führt »Gradualität« wie folgt ein:

*»Die Kombination von dreien solcher Eigenschaften ist besonders interessant, nämlich von Objekt-Eindeutigkeit, Objekt-Abbildbarkeit und Ereignis-Abbildbarkeit. Sie sei Gradualität genannt. Sie drückt aus, daß das Objekt dem Verb auf eine graduelle oder inkrementelle Weise unterworfen ist, das heißt in der Art und Weise, wie es in dem Raum-Zeit-Diagramm oben dargestellt wurde.«*

Da »ein Glas Wein trinken« objekt-eindeutig (s. 3.4), subobjekt-abbildbar und subereignis-abbildbar ist, ist ein so modellierbares Ereignis auch graduell.

Anlog zur Gradualität thematischer Relationen wird die Eigenschaft der Gradualität für segment-thematische Relationen definiert.

**Definition 3.6:** Eine segment-thematische Relation  $R_{S(\mathcal{L},\mathcal{B}),\{E\}}$  ist in einem Segment  $S$  eines cinematographischen Dokumentes für eine Beobachtermenge  $\mathcal{B}$  und eine vorgegebene Labelmenge  $\mathcal{L}$  **graduell**, wenn gilt:

- (1)  $R_{S(\mathcal{L},\mathcal{B}),\{E\}}$  ist objekt-eindeutig,
- (2)  $R_{S(\mathcal{L},\mathcal{B}),\{E\}}$  ist subobjekt-abbildbar,
- (3)  $R_{S(\mathcal{L},\mathcal{B}),\{E\}}$  ist subereignis-abbildbar.

Ist die Anzahl der Einstellungen eines Segmentes und die Anzahl der Verankerungen eines graduellen Ereignisses bekannt, lässt sich definieren:

**Definition 3.7:** Eine Ereignisklasse  $E: EVT$  ist in einem Segment  $S$  für eine Beobachtermenge  $\mathcal{B}$  für eine vorgegebene Labelmenge  $\mathcal{L}$  **durch  $R_{S(\mathcal{L},\mathcal{B}),\{E\}}$  graduell  $(n, m)$ -verankerbar** oder **graduell  $(n, m)$ -R-verankerbar** für  $n, m \geq 1$ , wenn gilt:

- (1) die Ereignisklasse  $E$  ist in  $S$   $(n, m)$ -R-verankert,
- (2)  $R_{S(\mathcal{L},\mathcal{B}),\{E\}}$  ist graduell.

### 3.6 Handlungstypen

Für Handlungen und entsprechend für verbale Ausdrücke/Prädikate, die Handlungen repräsentieren, hat als erstes [Garey 1957] die Distinktionen »**atelisch**« und »**telisch**« verwendet. Ein Verbalausdruck ist atelisch<sup>1</sup>, wenn er keinen »natürlichen« Endpunkt impliziert [Krifka 1989b, S. 236]. Beispiele sind *laufen, gehen, Äpfel essen*. Handlungen dieser Art können beliebig fortgesetzt werden, mit der Konsequenz, dass atelische Prädikate Ereignisse repräsentieren, deren Teile typischerweise selbst wieder unter dieses Prädikat fallen. Diese Eigenschaft bezeichnen [Bennet, Partee 1972] als »*subinterval property*«. Telische Prädikate repräsentieren dagegen Ereignisse mit natürlichem Endpunkt. Beispiele sind *einen Apfel essen* und *einen Kreis zeichnen*. [Moens, Steedman 1988, S. 18] beschreiben die Struktur eines telischen Ereignisses wie folgt: Es besteht aus einem vorbereitenden Prozess (Vorbereitungsphase), einem Kulminationspunkt und einem resultierenden Zustand (s. Abb. 3.13).

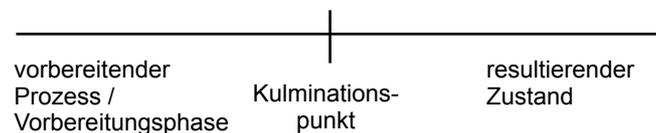


Abb. 3.13: Struktur eines telischen Ereignisses

[Vendler 1957], [Vendler 1967] unterscheidet vier »*time schemata*«, die nach [Krifka 1989a, S. 98] »*möglicherweise den gesamten Verbvorrat erschöpfend klassifizieren*«. Die vier Distinktionen, welche als Handlungstypen bezeichnet werden, sind *State, Achievement, Activity* und *Accomplishment*.

Ein **State** beschreibt einen Zustand, der eine gewisse Zeitdauer und keinen Endpunkt hat, also atelisch ist. Beispiele für Zustände sind *love somebody* und *have something*.

Für die Behandlung der Handlungstypen *Achievement, Activity* und *Accomplishment* sei im Folgenden  $I$  der Anfangspunkt, d. h. der Zeitpunkt zu dem das betrachtete Ereignis beginnt, und  $F_N$  der natürliche Endpunkt des Ereignisses.  $F_A$  sei ein beliebiger, willkürlicher (*arbitrary*) Endpunkt.

Ein **Achievement** ist ein Ereignis, das vollendet ist, sobald es begonnen hat. Es handelt sich dabei um eine unmittelbare, punktuelle Zustandsänderung, d. h. der Anfangspunkt  $I$  und der natürliche Endpunkt  $F_N$  fallen zusammen. Ein Beispiel ist *sehen*: Zu dem Zeitpunkt, zu dem man etwas sieht, hat man schon gesehen. Einem Ereignis können jedoch Prozesse vorausgehen (s. Abb. 3.14). [Singh, Singh 1997, S. 51] nennen dazu als Beispiel »ein Rennen gewinnen«. Der Agent eines solchen Ereignisses muss zumindest etwas laufen, bevor er gewinnen kann.

<sup>1</sup> von *telos* (griech.): Ziel

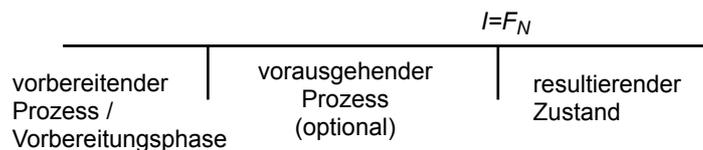


Abb. 3.14: Schema für Achievements

Eine **Activity** ist eine Handlung, die beliebig fortgesetzt werden kann, ohne ihren Charakter zu verändern. Sie ist im Zeitverlauf homogen und hat keinen natürlichen Endpunkt, ist also atelisch. Sie kann kleinere Ereignisse beinhalten, ein vorbereitender Prozess kann vorausgehen (s. Abb. 3.15, vgl. [Singh, Singh 1997, S. 51]). Beispiele sind *laufen* und *essen*.

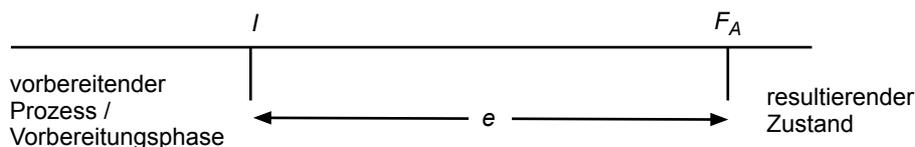


Abb. 3.15: (Einfaches) Schema für Activities

Ein **Accomplishment** ist eine Tätigkeit, die auf das Erreichen eines Ziels gerichtet ist. Sie ist stufenweise fortschreitend (graduell) und erreicht einen natürlichen Endpunkt. Beispiele sind *einen Kilometer laufen*, *einen Kreis zeichnen*, *einen Apfel essen*. Abb. 3.16 (vgl. [Singh, Singh 1997, S. 51]) zeigt ein (einfaches) Schema für Accomplishments.

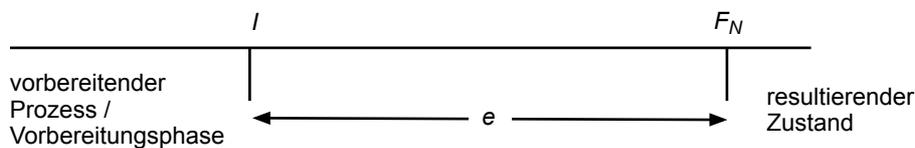


Abb. 3.16: (Einfaches) Schema für Accomplishments

Mit [Singh, Singh 1995, Fig. 2] wird in Abb. 3.17 eine genauere Skizze der Struktur von *Activities* und *Accomplishments* gezeigt, in der die vorbereitenden Prozesse  $p_1, p_2, \dots$  und die Teilereignisse  $e^{(1)}, e^{(2)}, \dots, e^{(k)}$  von  $e$  sichtbar werden.

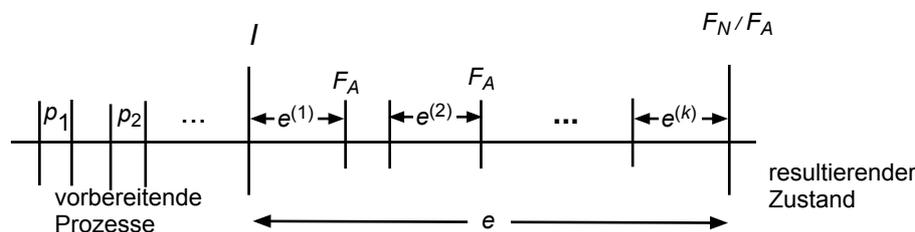


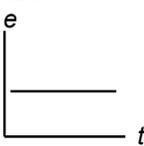
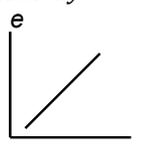
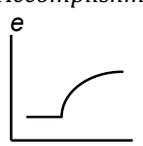
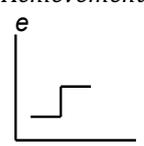
Abb. 3.17: Detailliertes Schema für Activities und Accomplishments

[Halliday, Matthiessen 2000, S. 470f] charakterisieren zur Abgrenzung der vier Vendler'schen Handlungstypen diese mit den drei unabhängigen Dimensionen:

- Veränderung (*change*) versus keine Veränderung (*no change*),

- (Zeitlich) begrenzt (*bounded*) versus unbegrenzt (*unbounded*)<sup>1</sup>,
- andauernd (*duration*) versus punktuell (*no duration*).

Tab. 3.1 zeigt eine diese drei Merkmalsdimensionen nutzende Anordnung der vier Handlungstypen in Tabellenform.

<i>no change</i>		<i>change</i>		
	<i>State</i> 	<i>Activity</i> 		<i>unbounded</i>
		<i>Accomplishment</i> 	<i>Achievement</i> 	<i>bounded</i>
<i>no duration</i>	<i>duration</i>		<i>no duration</i>	

Tab. 3.1: Systematik der Handlungstypen von Vendler

[Krifka 1989a, S. 34] führte noch weitere Eigenschaften für Prädikate ein: die Eigenschaften »kumulativ« und »gequantelt«. Allgemein gilt für sowohl für Prädikate über Objekte als auch für Prädikate über Ereignisse:

Ein Prädikat ist **kumulativ**, wenn gilt:

$$\forall P (\text{CUM}(P) \Leftrightarrow \forall x, y (P(x) \wedge P(y) \rightarrow P(x \oplus y)))$$

Dies bedeutet: Wenn auf zwei Entitäten das Prädikat  $P$  angewandt werden kann, dann kann es auch auf die Zusammenfassung angewandt werden [Krifka 1989b, S. 228]. »Äpfel« sind kumulativ in dem Sinne, dass »Äpfel« und »Äpfel« wieder »Äpfel« ergeben. Zwei »laufen«-Ereignisse fallen wieder unter »laufen«.

Ein Prädikat ist **gequantelt**, wenn gilt:

$$\forall P (\text{QUA}(P) \Leftrightarrow \forall x, y (P(x) \wedge y < x \rightarrow \neg P(y)))$$

Dies bedeutet: Wenn ein Prädikat auf zwei Entitäten angewendet werden kann, so ist es auf die Zusammenfassung nicht mehr anwendbar [Krifka 1989b, S. 228]. Beispielsweise sind »2Äpfel« und »1km\_laufen« gequantelt.

Zwischen Kumulativität und Gequanteltheit gilt folgende Beziehung [Singh, Singh 1995, S. 3]:

$$\text{CUM}(P) \rightarrow \neg \text{QUA}(P), \text{ wenn } |\{x | P(x)\}| \geq 2.$$

<sup>1</sup> Die Begriffe *bounded* bzw. *unbounded* entsprechen in der Sprechweise von [Garey 1957], [Krifka 1989a] den Begriffen telisch bzw. atelisch.

**Ein »Regelkatalog« zur Bestimmung des Handlungstyps**

Mit Hilfe der eingeführten Begriffe ist es für natürlich-sprachliche Sätze möglich, anhand der Eigenschaften des Verbs und den Eigenschaften des Objektes, das in der PAT-Relation zum konzeptionalisierten Ereignis steht, den Handlungstyp der Gesamtkonstruktion zu bestimmen.

Die Telizität einer Handlung ist abhängig von der Kumulativität des Objektes, das in der PAT-Relation zu dem konzeptionalisierten Ereignis steht, [Krifka 1989a, S. 158]:

*»Ein verbales Prädikat ist telisch, wenn es gequantelt ist. Die Atelizität wird umgekehrt durch die Kumulativität des verbalen Prädikates erfasst.«*

Das bedeutet: Da »laufen« kumulativ ist, ist »laufen« atelisch. Dagegen ist »1km laufen« gequantelt, also telisch.

Ob es sich bei einem Ereignis um ein *Accomplishment*, eine *Activity* oder ein *Achievement* handelt, wird sowohl durch die Eigenschaften des Verbs als auch die des Objektes bestimmt. [Singh, Singh 1992, S. 530] stellen die Regeln zur Bestimmung des Handlungstyps in folgender Kurznotation dar:

[Objekteigenschaft, PAT-Relations-Eigenschaft → Handlungstyp]

Es gilt:

- **Accomplishments:** [+QUA, +GRAD → ACCOMPLISHMENT]
- **Activities:** [+CUM, +GRAD → ACTIVITY]
- **Achievements:** [+QUA, - GRAD → ACHIEVEMENT]

Dies bedeutet: Ist das Ereignis mit einer graduellen PAT-Relation modellierbar und das Objekt gequantelt, liegt ein *Accomplishment* vor, ist das Objekt dagegen kumulativ liegt eine *Activity* vor. Ist das Objekt gequantelt und das Ereignis mit einer PAT-Relation modellierbar, die nicht graduell ist, handelt es sich um ein *Achievement*.

Da *Achievements* von der Identität des Anfangspunktes und des Endpunktes des Verbalvorgangs ausgehen, ist eine filmische Repräsentation typischerweise nur in einer Einstellung einzufangen (unter der Voraussetzung, dass auf eine Mehrfachrepräsentation desselben Ereignisses verzichtet wird). *Activities* und *Accomplishments* lassen sich dagegen auf natürliche Art und Weise auch auf mehrere Einstellungen verteilen. Für den Handlungstyp *Accomplishment* wird dies in 3.8 behandelt.

### 3.7 Repräsentation von Accomplishments im Beispielvideo

Im Beispielvideo »Snowboard« werden zwei Handlungsstränge präsentiert: der Bau des Snowboards in der Werkstatt und der Test des – teilweise fertiggestellten – Snowboards durch Johanna auf dem Berg.

Ein Snowboard kann als gequanteltes Objekt der Snowboardfertigung angesehen werden. Der Bau des Snowboards kann zweifelsfrei als *Accomplishment* konzeptionalisiert werden. Es wird gezeigt, wie der Bau Schritt für Schritt voranschreitet, bis das Snowboard fertiggestellt ist. Der eigentlichen Handlung eines *Accomplishments* kann eine Vorbereitungsphase vorausgehen (s. Abb. 3.16, Abb. 3.17). Im hier betrachteten Beispiel kann man davon ausgehen, dass die Vorbereitungsphase aus Tätigkeiten wie Werkzeuge und Materialien besorgen, diese an den Arbeitsstationen bereitstellen etc., besteht. Im Video wird die Vorbereitungsphase jedoch nicht gezeigt. Der Bau eines Snowboards besteht aus mehreren Teilprozessen/Subereignissen. Eine Übersicht zeigt Tab. 3.2.

Subereignis	Einstellungen	Handlungen
$e^{(1)}$	T04 – T08	Bretter fräsen
$e^{(2)}$	T09 – T10	Bretter zusammenkleben
$e^{(3)}$	T12 – T13	Bretterenden rund fräsen
$e^{(4)}$	T14 – T15	Löcher bohren
$e^{(5)}$	T17	Gewinde in Löcher einsetzen
$e^{(6)}$	T18	Kunststoffbahnen auf Bretter
$e^{(7)}$	T19	Glasfaserschicht auf Bretter
$e^{(8)}$	T20	Fräsen
$e^{(9)}$	T21 – T31	Gummistreifen, Stahlkappen, Fiberglasstreifen aufkleben
$e^{(10)}$	T32 – T35	Pressen
$e^{(11)}$	T37	Mit Stichsäge in Form bringen
$e^{(12)}$	T38 – T39	Stahlkanten schleifen
$e^{(13)}$	T40	Löcher aufbohren / Bindung montieren

Tab. 3.2: Übersicht: Subereignisse von »Bau des Snowboards«

Jeder Teilprozess endet mit einem vorab definierten Zustand, also einem natürlichen Endpunkt. Daher sind auch die Subereignisse *Accomplishments*. Wie in Abschnitt 2.5.3 bereits erwähnt und in Abb. 2.30 dargestellt, ist der Endzustand eines Prozessschrittes Eingangszustand des folgenden Schrittes. Abb. 3.18 zeigt das Schema des *Accomplishments* »Bau des Snowboards«.

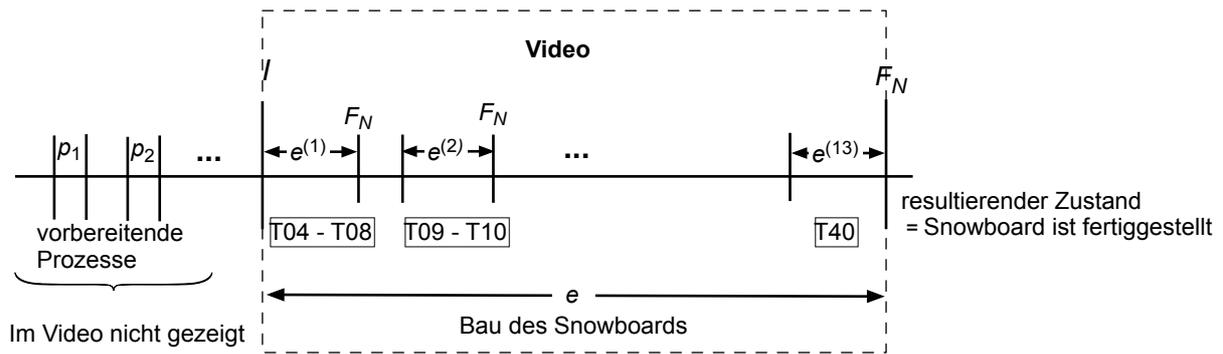


Abb. 3.18: Schema des Accomplishments »Bau des Snowboards«

Der zweite Handlungsstrang im Video besteht aus einzelnen Tests des Snowboards. Die Tests finden zu diegetisch unterschiedlichen Zeiten statt. Testobjekt ist jeweils ein Snowboard in einem Entwicklungszustand. Bei jedem Test ist dieses jedoch in einem anderen, etwas vollständigerem Zustand. Jede einzelne Testsituation könnte man als *Activity* auffassen: »ein bisschen auf dem Snowboard rumrutschen und schauen, ob es funktioniert«. Jeder Teil davon ist wieder ein »ein bisschen auf dem Snowboard rumrutschen und schauen, ob es funktioniert«. Dies wäre eine mögliche Interpretation. Sie könnte dann plausibel sein, wenn man das Dokument nicht in der layoutierten Ordnung, sondern die beiden Handlungsstränge getrennt betrachtet (s. dazu 4.9). Eine andere Interpretationsmöglichkeit ist: Jeder Test hat einen natürlichen Endpunkt. Der resultierende Zustand ist in diesem Fall das Ergebnis des Tests, als Werte sind »success« und »failure« möglich, je nachdem, ob das Snowboard zum Fahren geeignet ist oder nicht. Tab. 3.3 zeigt eine Übersicht über die fünf im Video gezeigten Tests (vgl. auch Tab. 4.5).

Subereignis	Einstellungen	Handlungen
$e^{(1)}$	$T_{01} - T_{03}$	Test 1 – Ergebnis: »failure«
$e^{(2)}$	$T_{11}$	Test 2 – Ergebnis: »failure«
$e^{(3)}$	$T_{16}$	Test 3 – Ergebnis: »failure«
$e^{(4)}$	$T_{36}$	Test 4 – Ergebnis: »failure«
$e^{(5)}$	$T_{41} - T_{42}$	Test 5 – Ergebnis: »success«

Tab. 3.3: Übersicht: Subereignisses von »Test des Snowboards«

Die ersten vier Tests haben als Ergebnis »failure«. Jede dieser Handlungen ist ein *Accomplishment*. Abb. 3.19 zeigt das Schema eines *Accomplishments* »Test« beispielhaft für den ersten Test.

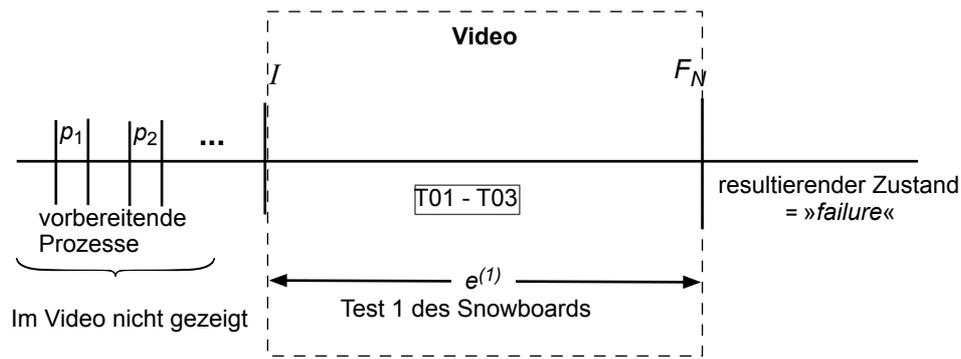


Abb. 3.19: Schema des Accomplishments »Test 1 des Snowboards«

Der letzte Test hat als Ergebnis »success«. Diese Handlung könnte daher statt als »Test« auch als »Fahren mit dem Snowboard« interpretiert werden, da das Fahren ja nun möglich ist. In der layoutierten Ordnung hat diese Lesart die größere Plausibilität.

Abb. 3.20 zeigt ein Raum-Zeit-Diagramm zur Handlung »Bau des Snowboards«. Auf der Ordinate wird in diesem Fall nicht der Raum, sondern der Fertigungszustand des Snowboards abgebildet. Die durchgezogenen Linien markieren zusätzlich zum Einstellungsende auch einen Ortswechsel.

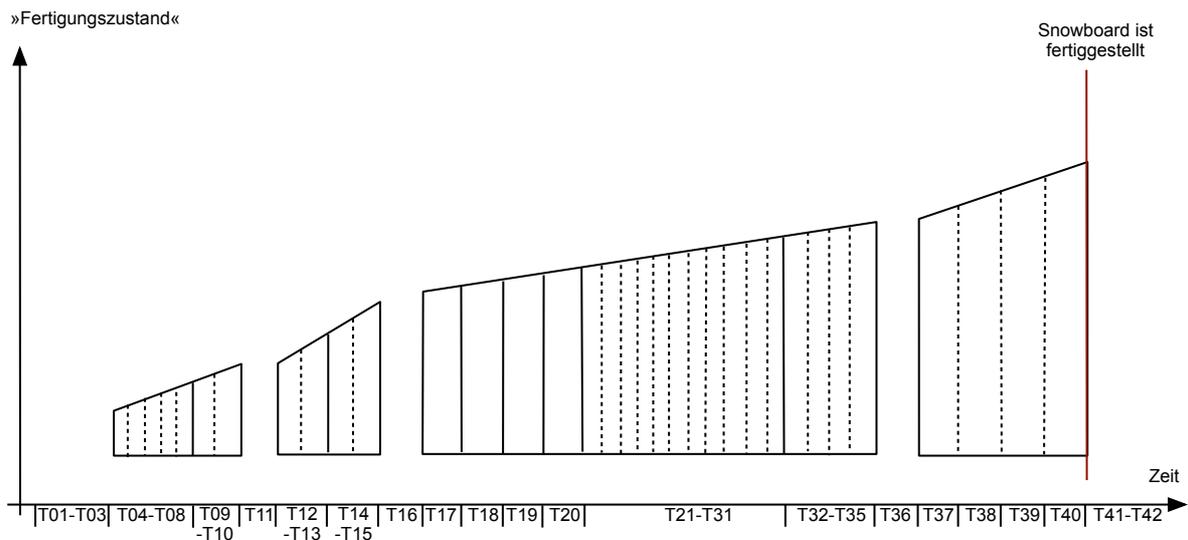


Abb. 3.20: Raum-Zeit-Diagramm zu »Bau des Snowboards«

Nimmt man den zweiten Handlungsstrang hinzu, ergibt sich das in Abb. 3.21 gezeigte Raum-Zeit-Diagramm für die alternante Struktur des gesamten Dokumentes. Dieses Raum-Zeit-Diagramm ist um 90 Grad gedreht, der zeitliche Fortschritt ist also bei einer Leserichtung von oben nach unten erkennbar. Im linken Bereich des Diagramms werden die Einstellungen der Tests auf dem Berg dargestellt. Hier ist visualisiert, dass sich der »Fertigungszustand« des Snowboards während eines Tests nicht ändert.<sup>1</sup> Im rechten Bereich wird wie in Abb. 3.20 das Raum-Zeit-Diagramm zu »Bau des Snowboards« dargestellt. Zusätzlich zu den Einstellungen sind die zeitli-

<sup>1</sup> Vgl. hierzu auch Abb. 4.21. Hier werden die Zustandswechsel des gesamten Videos »Snowboard in Form eines Petri-Netzes dargestellt und es ist ebenfalls erkennbar, dass sich der Zustand des Snowboards in den Berg-Segmenten nicht verändert.

chen Lücken zwischen ihnen und auch die Arbeitsstationen vermerkt, so dass das Diagramm auch die wesentlichen Informationen der syntagmatischen Analyse enthält.

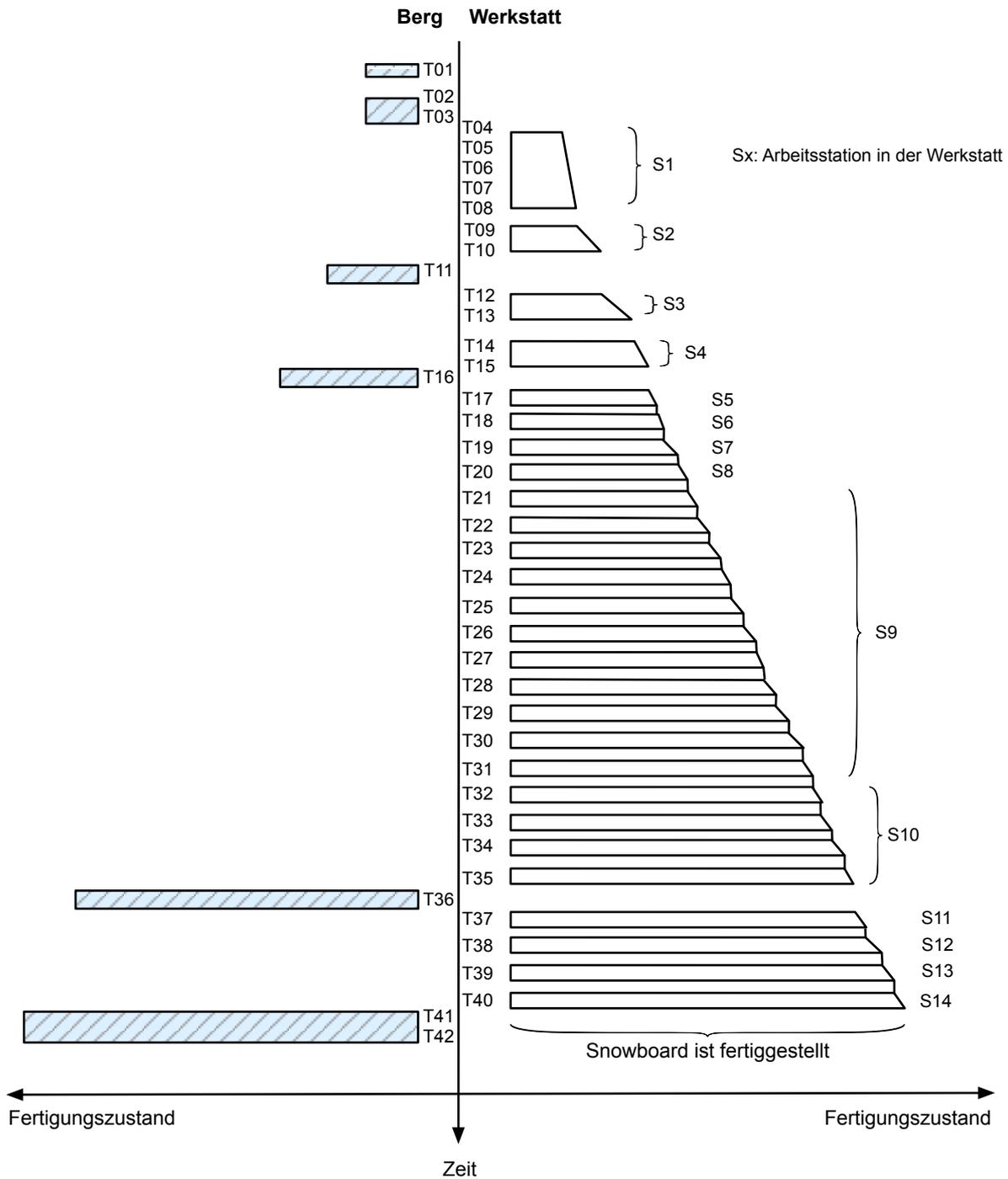


Abb. 3.21: Raum-Zeit-Diagramm des Videos »Snowboard«

### 3.8 Zur cinematographischen Realisierung von Accomplishments

Die bisherigen Ausführungen motivieren eine »Übertragung« der für diesen Kontext wichtigen Definitionen aus der Linguistik auf cinematographische Dokumente. Es ist in der folgenden Weise möglich, die »*Accomplishment*-Regel« [+QUA, +GRAD → ACCOMPLISHMENT] auf cinematographische Dokumente zu übertragen:

#### **Accomplishment-Regel für cinematographische Dokumente**

In einem VideoSegment  $S$  eines cinematographischen Dokumentes ist für eine Beobachtermenge  $\mathcal{B}$  ein Ereignis  $E$  als **Accomplishment** konzeptionalisierbar, wenn eine Instanz  $e$  mit  $inst_{\mathcal{B}}(e, E)$  existiert, die Ereignisklasse  $E, E \in \mathcal{E}$ , in  $S$  graduell  $(n, m)$ -PAT-verankerbar ist, und die instanziierten Objekte, die in der  $PAT_{S(\mathcal{L}, \mathcal{B}), \{E\}}$ -Relation zu dem konzeptualisierten Ereignis stehen, beobachterseits als gequantelt aufgefasst werden können.

*Accomplishments* legen für ihre filmische Repräsentation eine Ordnung nahe. Für die Reihenfolge der ein *Accomplishment* repräsentierenden Einstellungen ist für ein Defaultlayout ein Ablauf so festzulegen, dass im Default Abfolgen auf der Leinwand diegetischen Abfolgen homomorph entsprechen. Zudem muss die Telizität dieser graduellen Handlung durch Inszenierung eines natürlichen Endpunktes dem Beobachter vermittelt werden. Diese Randbedingungen werden für das Beispielfilmvideo untersucht und mit dem bisher entwickelten Begriffsapparat formal nachgezeichnet.

Handlungen können auf verschiedene Weise auf Einstellungen verteilt sein. Um dies zu modellieren, wird folgende Notation verwendet: Sei  $n$  die Anzahl der Einstellungen eines Segmentes und  $m$  die Anzahl der Handlungen, dann bedeutet

$\llbracket n \rrbracket_{\mathcal{B}}^m$  : in  $n$  Einstellungen werden für eine Beobachtermenge  $\mathcal{B}$   $m$  Handlungen repräsentiert,  
 $n \geq 1, m \geq 1$ .

Die eckigen Klammern » $\llbracket \cdot \rrbracket$ « werden hier gewählt, um in der Notation den Rahmen eines Filmbildes visuell anzudeuten. Ist die Beobachtermenge unerheblich, wird das  $\mathcal{B}$  weggelassen.

Bezüglich der Beziehung »Handlungen zu Einstellungen« werden nun drei mögliche Realisationstypen unterschieden:

- $\llbracket 1 \rrbracket^1$  Eine »ganze« Handlung wird beobachterseits in einer Einstellung identifiziert (wie in einer szenischen Einstellung);
- $\llbracket 1 \rrbracket^m$  mehrere (» $m$ «) (Teil-)Handlungen werden in einer (diese  $m > 1$  Handlungen fusionierenden) Einstellung repräsentiert;

$\llbracket n \rrbracket^1$  eine Handlung wird auf mehrere, wiederum mehr als eine ( $n > 1$ ) Einstellungen verteilt, wobei die Einstellungen syntagmatisch nicht gebunden sind oder verschiedenen großen Syntagmen zugeschlagen werden können.

### Realisationstyp $\llbracket 1 \rrbracket^1$

Im Falle des Realisationstyps  $\llbracket 1 \rrbracket^1$  wird angenommen, dass eine »ganze« Handlung vollständig in nur einer Einstellung präsentiert wird. Wie schon in Kapitel 2 erwähnt, beschreibt [Metz 1972, S. 172] dies wie folgt: Eine »ganze Szene wird in einer Einstellung behandelt, d. h. die Einheit der "Handlung" verleiht der Einstellung ihre Autonomie.« Er bezeichnet dieses Syntagma als Plansequenz. Diese Auffassung ist inhaltlich richtig, aber in der Anwendung voller Probleme, wie schon [Colin 1995, S. 55] feststellte. Mit den hier bereitgestellten Mitteln ist die Klassifikation einer Planszene als Realisation vom Typ  $\llbracket 1 \rrbracket^1$  aber kein Problem.

### Beispiel 3.5

Als Beispiel für eine Realisation vom Typ  $\llbracket 1 \rrbracket^1$  wird die Einstellung T17 im Beispielvideo »Snowboard« betrachtet. In dieser Planszene wird repräsentativ gezeigt, wie vier Metallteile am Brett befestigt werden.<sup>1</sup> Tab. 3.4 stellt diese Einstellung schematisch dar. In der linken Spalte der Tabelle werden die ShotRegions, die den Ausgangs- bzw. Endzustand einer (Teil-)Handlung repräsentieren, gezeigt, die beiden mittleren Spalten geben die einzelnen Handlungen wieder, in der rechten Spalte sind ggf. vorhandene Teilhandlungen vermerkt. Für Teilereignisse eines Teilergebnisses wird hier folgende Schreibweise verwendet:  $e^{(i,j)}$  für das  $j$ -te Teilereignis des Teilergebnisses  $e^{(i)}$ .

Die Einstellung beginnt mit einem Bild, auf dem der Beobachter ein bereits montiertes Metallteil erkennen kann. Die Befestigung der beiden folgenden Teile wird in der Einstellungsgröße »G« gezeigt, dann zoomt die Kamera weg und die Befestigung von zwei weiteren Metallteilen wird in der Einstellungsgröße »HN« gezeigt.

ShotRegion	Handlung		Teilhandlung
			
brett1 $s_{17,(1)}$			

<sup>1</sup> Aus dem Kontext (Sprechertext) weiß ein Beobachter, dass zweimal acht Löcher gebohrt werden.

		<p>1_metallteil_befestigen_1</p> <p><math>e^{(1)}</math></p>	<p>1_metallteil_in_Loch- legen_1</p> <p><math>e^{(1,1)}</math></p> <p>AG: akteur4</p>
		<p><math>e^{(1)}</math></p>	<p>1_metallteil_fest- drücken_1</p> <p><math>e^{(1,2)}</math></p> <p>AG: maschine4</p>
 <p>brett1 <math>s_{17,(2)}</math></p>			
		<p>1_metallteil_befestigen_2</p> <p><math>e^{(2)}</math></p>	<p>1_metallteil_in_Loch- legen_2</p> <p><math>e^{(2,1)}</math></p> <p>AG: akteur4</p>
		<p><math>e^{(2)}</math></p>	<p>1_metallteil_fest- drücken_2</p> <p><math>e^{(2,2)}</math></p> <p>AG: maschine4</p>
 <p>brett1 <math>s_{17,(3)}</math></p>			
<p>--- Zoom ---</p>			
 <p>brett1 <math>s_{17,(3)}</math></p>			

		1_metallteil_befestigen_3	1_metallteil_in_Loch- legen_3 $e^{(3,1)}$ AG: akteur4
		$e^{(3)}$	1_metallteil_fest- drücken_3 $e^{(3,2)}$ AG: maschine4
			
$s_{17,(4)}^{\text{brett1}}$			
		1_metallteil_befestigen_4	1_metallteil_in_Loch- legen_4 $e^{(4,1)}$ AG: akteur4
		$e^{(4)}$	1_metallteil_fest- drücken_4 $e^{(4,2)}$ AG: maschine4

Tab. 3.4: Handlungen und Teilhandlungen in T17

In der Einstellung  $T17$  sind vier Instanzen der Ereignisklasse  $1\_METALLTEIL\_BEFESTIGEN$  zu erkennen. Für die segment-thematische Relation  $PAT_{S(\mathcal{L},\mathcal{B}),\{1\_METALLTEIL\_BEFESTIGEN\}}$  der  $i$ -ten Instanz,  $1 \leq i \leq 4$ , gilt:

$$\begin{aligned}
 & PAT_{S(\mathcal{L},\mathcal{B}),\{1\_METALLTEIL\_BEFESTIGEN\}} \\
 &= \{(s, e) \mid s \in S(\mathcal{L}, \mathcal{B}), inst_{\mathcal{B}}(e, E), E \in \mathcal{E}\} \\
 &= \{(s, e) \mid s \in S(\{\text{BRETT}\}, \mathcal{B}), \\
 &\quad inst_{\mathcal{B}}(e, E), E \in \{1\_METALLTEIL\_BEFESTIGEN\}\} \\
 &= \{(s_{17,(i)}^{\text{brett1}}, e^{(i)}) \mid inst_{\mathcal{B}}(\text{brett1}, \text{BRETT}), \\
 &\quad inst_{\mathcal{B}}(e^{(i)}, 1\_METALLTEIL\_BEFESTIGEN)\}
 \end{aligned}$$

Da die PAT-Relation offensichtlich graduell ist und vier, sich nicht überlappende Instanzen konzeptionalisiert werden können, ist die Ereignisklasse  $1\_METALLTEIL\_BEFESTIGEN$  graduell (1,4)-PAT-verankerbar in  $T17$ . Da das betroffene Objekt, das Brett, gequantelt ist, ist das Ereignis  $1\_METALLTEIL\_BEFESTIGEN$  ein *Accomplishment*. Die Telizität dieses Ereignisses ist gerade bei den ersten beiden Instanzen sehr gut erkennbar, da das Ergebnis, das Brett mit den bereits montierten Metallteilen, in Großaufnahme gezeigt wird. Die beiden folgenden Instanzen werden in der Einstellungsgröße »HN« gezeigt, zudem wird das Brett durch den Akteur verdeckt, so dass es dem Beobachter schon schwerer fällt, das Ergebnis explizit zu erkennen. Das letzte Ergebnis ist sogar dem Schnitt zum Opfer gefallen. Hier kann man jedoch davon ausgehen, dass ein Beobachter kein Problem hat, die Telizität der dritten und vierten Instanz zu »interpolieren«.

Fokussiert ein Beobachter das Ereignis  $E = X\_METALLTEILE\_BEFESTIGEN$ , so sind für ihn für eine Instanz  $e$ ,  $inst_{\mathcal{B}}(e, X\_METALLTEILE\_BEFESTIGEN)$ ,  $X$  Teilereignisse  $e^{(i)} < e$ ,  $inst_{\mathcal{B}}(e^{(i)}, 1\_METALLTEIL\_BEFESTIGEN)$ ,  $1 \leq i \leq X$ ,  $X \in \mathbb{N}$ , und  $1\_METALLTEIL\_BEFESTIGEN < X\_METALLTEILE\_BEFESTIGEN$  erkennbar.

Im Beispielvideo »Snowboard« kann ein Beobachter die Befestigung von vier Metallteilen erkennen, es ist also  $X = 4$ . Für die segment-thematische Relation  $PAT_{S(\mathcal{L},\mathcal{B}),\{4\_METALLTEILE\_BEFESTIGEN\}}$  gilt:

$$\begin{aligned}
 & PAT_{S(\mathcal{L},\mathcal{B}),\{4\_METALLTEILE\_BEFESTIGEN\}} \\
 &= \{(s, e) \mid s \in S(\mathcal{L}, \mathcal{B}), inst_{\mathcal{B}}(e, E), E \in \mathcal{E}\} \\
 &= \{(s, e) \mid s \in S(\{\text{BRETT}\}, \mathcal{B}), \\
 &\quad inst_{\mathcal{B}}(e, E), E \in \{4\_METALLTEILE\_BEFESTIGEN\}\} \\
 &= \{(s_{17,(1)}^{\text{brett1}}, e^{(1)}), (s_{17,(2)}^{\text{brett1}}, e^{(2)}), (s_{17,(3)}^{\text{brett1}}, e^{(3)}), (s_{17,(4)}^{\text{brett1}}, e^{(4)}) \\
 &\quad \mid inst_{\mathcal{B}}(\text{brett1}, \text{BRETT}), \\
 &\quad inst_{\mathcal{B}}(e^{(i)}, 1\_METALLTEIL\_BEFESTIGEN) \text{ für } 1 \leq i \leq 4, \\
 &\quad 1\_METALLTEIL\_BEFESTIGEN < 4\_METALLTEILE\_BEFESTIGEN\}
 \end{aligned}$$

Jede Veränderung von  $s_{17,(i)}^{\text{brett1}}$  wird durch genau ein Teilereignis  $e^{(i)}$  bewirkt und umgekehrt ist zu jedem Teilereignis  $e^{(i)}$  eine Veränderung des Pixelsegmentes  $s_{17,(i)}^{\text{brett1}}$  des Brettes erkennbar, daher ist die Relation  $PAT_{S(\mathcal{L},\mathcal{B}),\{4\_METALLTEILE\_BEFESTIGEN\}}$  subereignis-abbildbar und subobjekt-abbildbar. Sie erfüllt zudem die Forderung nach Objekt-Eindeutigkeit, da die (Teil-)Segmente  $s_{17,(1)}^{\text{brett1}}, s_{17,(2)}^{\text{brett1}}, s_{17,(3)}^{\text{brett1}}, s_{17,(4)}^{\text{brett1}}$  mit dem gleichen Label »brett1« gelabelt wurden. Somit folgt, dass die Relation  $PAT_{S(\mathcal{L},\mathcal{B}),\{4\_METALLTEILE\_BEFESTIGEN\}}$  graduell und die Ereignisklasse  $4\_METALLTEILE\_BEFESTIGEN$  graduell (1,1)-PAT-verankerbar in der Planszene  $T17$  ist. Da das »betroffene Objekt«, in diesem Fall das Brett, gequantelt ist, liegt für das Ereignis  $4\_METALLTEILE\_BEFESTIGEN$  ein *Accomplishment* vor.

Eine weitere Beobachtermenge,  $\mathcal{B}'$ , betrachtet die Teilereignisse  $1\_METALLTEIL\_BEFESTIGEN$  noch genauer und erkennt, dass jedes (Teil-)Ereignis wiederum aus zwei Teilereignissen besteht: ein Akteur (hier:  $akteur4$ ) legt ein Metallteil in ein vorgebohrtes Loch, dann drückt eine Maschine (hier:  $maschine4$ ) das Metallteil fest (s. Tab. 3.4, rechte Spalte).

Es gilt für die segment-thematische Relation AG und PAT der  $i$ -ten Instanz,  $1 \leq i \leq 4$ , von  $1\_METALLTEIL\_BEFESTIGEN$ :

$$\begin{aligned}
& \text{AG}_{S(\mathcal{L}, \mathcal{B}'), \{1\_METALLTEIL\_BEFESTIGEN\}} \\
&= \{(s, e) \mid s \in S(\mathcal{L}, \mathcal{B}'), \text{inst}_{\mathcal{B}'}(e, E), E \in \mathcal{E} \\
&= \{(s, e) \mid s \in S(\{\text{PERSON}, \text{MASCHINE}\}, \mathcal{B}), \\
&\quad \text{inst}_{\mathcal{B}'}(e, E), E \in \{1\_METALLTEIL\_BEFESTIGEN\}\} \\
&= \{(s_{17,(i)}^{\text{akteur4}}, e^{(i,1)}), (s_{17,(i)}^{\text{maschine4}}, e^{(i,2)}) \\
&\quad \mid \text{inst}_{\mathcal{B}'}(\text{akteur4}, \text{PERSON}), \text{inst}_{\mathcal{B}'}(\text{maschine4}, \text{MASCHINE}), \\
&\quad \text{inst}_{\mathcal{B}'}(e^{(i,1)}, 1\_METALLTEIL\_IN\_LOCH\_LEGEN), \\
&\quad \text{inst}_{\mathcal{B}'}(e^{(i,2)}, 1\_METALLTEIL\_FESTDRÜCKEN), \\
&\quad 1\_METALLTEIL\_IN\_LOCH\_LEGEN < 1\_METALLTEIL\_BEFESTIGEN, \\
&\quad 1\_METALLTEIL\_FESTDRÜCKEN < 1\_METALLTEIL\_BEFESTIGEN\}
\end{aligned}$$

$$\begin{aligned}
& \text{PAT}_{S(\mathcal{L}, \mathcal{B}'), \{1\_METALLTEIL\_BEFESTIGEN\}} \\
&= \{(s, e) \mid s \in S(\mathcal{L}, \mathcal{B}'), \text{inst}_{\mathcal{B}'}(e, E), E \in \mathcal{E}\} \\
&= \{(s, e) \mid s \in S(\{\text{BRETT}\}, \mathcal{B}'), \\
&\quad \text{inst}_{\mathcal{B}'}(e, E), E \in \{1\_METALLTEIL\_BEFESTIGEN\}\} \\
&= \{(s_{17,(i*2-1)}^{\text{brett1}}, e^{(i,1)}), (s_{17,(i*2)}^{\text{brett1}}, e^{(i,2)}) \\
&\quad \mid \text{inst}_{\mathcal{B}'}(\text{brett1}, \text{BRETT}), \\
&\quad \text{inst}_{\mathcal{B}'}(e^{(i,1)}, 1\_METALLTEIL\_IN\_LOCH\_LEGEN), \\
&\quad \text{inst}_{\mathcal{B}'}(e^{(i,2)}, 1\_METALLTEIL\_FESTDRÜCKEN), \\
&\quad 1\_METALLTEIL\_IN\_LOCH\_LEGEN < 1\_METALLTEIL\_BEFESTIGEN, \\
&\quad 1\_METALLTEIL\_FESTDRÜCKEN < 1\_METALLTEIL\_BEFESTIGEN\}
\end{aligned}$$

**Realisationstyp  $\llbracket 1 \rrbracket^m$**

Im Falles des Realisationstyps  $\llbracket 1 \rrbracket^m$ ,  $m \geq 1$ , werden mehrere (Teil-)Handlungen in einer (diese Handlungen ggf. fusionierenden) Einstellung repräsentiert. Wenn in einem Segment  $S = (T_1, \dots, T_i, \dots, T_n)$  eine Handlung in  $T_i$  endet und in  $T_i$  eine weitere Handlung beginnt, wird eine solche Einstellung als Fusion bezeichnet, vgl. [Bateman, Schmidt 2011, S. 202]:

**Definition 3.8:** Eine Einstellung in einem cinematographischen Dokument ist für einen Layoutprozess und für eine Beobachtermenge eine **Fusion**, wenn

- (1) die zwei Segmente, die beide die gegebene Einstellung enthalten, von allen Beobachtern als Repräsentation zweier verschiedener Ereignisse erkannt werden, und
- (2) die gegebene Einstellung die einzige Schnittmenge dieser Einstellungsmengen in diesen Segmenten für den gegebenen Layoutprozess ist.

Die Handlungen, die in der fusionierenden Einstellung repräsentiert sind, können diegetisch zusammenhängen (Ursache-Wirkungs-Beziehung) oder willkürlich (im filmischen »telling of the story«) zusammengestellt sein. Ob und wie evtl. weitere zu diesen Handlungen gehörende Teilhandlungen in anderen Einstellungen des Dokumentes verteilt sind, ist für die konkrete Fusion unerheblich. Ist etwa eine der beiden Handlungen insgesamt szenisch oder sequentiell repräsentiert, ist die fusionierende Einstellung einfach in einer Szene oder Sequenz gebunden. Sie wird dann klassifikatorisch in einer syntagmatischen Analyse gar nicht mehr sichtbar – das ändert aber nichts an ihrem Fusionscharakter.

**Beispiel 3.6**

Als Beispiel wird nun die Einstellung  $T09$  des Videos »Snowboard« betrachtet<sup>1</sup> (s. Abb. 3.22).

<b>T09</b>					
					
1	bretter legen_1		bretter legen_2		kleber auftragen_1
2	bretter vom Stapel nehmen_1	bretter zurecht legen_1	bretter vom Stapel nehmen_2	bretter zurecht legen_2	kleber auftragen_1
3	bretter zusammenkleben_1				(→ T10)

Abb. 3.22: Handlungen in  $T09$

Die Handlung BRETTER\_LEGEN beginnt und endet in  $T09$  (Zeile 1 des Bildtextes in Abb. 3.22). Sie ist in  $T09$  (1,2)-PAT-verankert und, da die PAT-Relation offensichtlich graduell ist, graduell (1,2)-PAT-verankerbar. In  $T09$  beginnt aber auch die Handlung KLEBER\_AUFTRAGEN. Im Fall 1 des

<sup>1</sup> Beispiel 3.1 zeigt ein Labelling für diese Einstellung, Beispiel 3.2 die segment-thematischen Relationen zu den in dieser Einstellung konzeptualisierten Handlungen.

Beispiels endet sie auch hier und in  $T_{10}$  werden neue Instanzen zu KLEBER\_AUFTRAGEN konzeptualisiert; im Fall 2 und 3 des Beispiels wird sie dagegen in  $T_{10}$  fortgesetzt, so dass in diesem Fall eine Fusion für die Handlung KLEBER\_AUFTRAGEN vorliegt und der Relationstyp  $\llbracket 1 \rrbracket_{\mathcal{B}}^3$  vorliegt.

Wie bereits in Beispiel 3.1 und 3.2 erwähnt, können verschiedene Beobachtermengen sowohl das Filmbild als auch die in ihm repräsentierten Handlungen mit unterschiedlicher Granularität konzeptualisieren. Konzeptualisiert die Beobachtermenge  $\mathcal{B}'$  die Handlungen in einer detaillierteren Form (Zeile 2 des Bildtextes in Abb. 3.22), also in  $T_{09}$  die Handlungen BRETTER\_VOM\_STAPEL\_NEHMEN und BRETTER\_AUF\_DEN\_TISCH\_LEGEN mit je zwei Instanzen und eine Instanz von KLEBER\_AUFTRAGEN, so liegt in  $T_{09}$  sogar der Realisationstyp  $\llbracket 1 \rrbracket_{\mathcal{B}'}^5$  vor. Im umgekehrten Fall, wenn eine weitere Beobachtermenge,  $\mathcal{B}''$ , eine »grobe Sicht auf die Dinge« hat (Zeile 3 des Bildtextes in Abb. 3.22), dann konzeptualisiert diese in  $(T_{09}, T_{10})$  lediglich eine Instanz der Ereignisklasse BRETTER\_ZUSAMMENKLEBEN, die in  $T_{10}$  fortgesetzt wird. Es liegt in diesem Fall keine Fusion in  $T_{09}$  vor, sondern im Segment  $(T_{09}, T_{10})$  der Realisationstyp  $\llbracket 2 \rrbracket_{\mathcal{B}''}^1$ .

### Realisationstyp $\llbracket n \rrbracket^1$

Im Fall  $\llbracket n \rrbracket^1$  wird eine Handlung in  $n$  Einstellungen,  $n \geq 2$ , präsentiert, wobei wenigstens zwei Einstellungen syntagmatisch nicht gebunden sein müssen oder verschiedenen großen Syntagmen zugeschlagen werden können. Dies bedeutet, dass keine gemeinsame Szene oder Sequenz vorliegen muss, und in diesem Falle wenigstens zwei Einstellungen paarweise verschiedene Raumzeitgebiete zeigen, es also auf der Ebene der VideoSegmente wenigstens einen »Raumsprung« gibt.

Aus der Definition der Gradualität und der *Accomplishment*-Regel folgt, dass jedes Teilereignis eines *Accomplishments* wieder ein *Accomplishment* sein kann. Repräsentiert eine Handlung über  $n$  Einstellungen hinweg ein *Accomplishment*, dann werden in den verschiedenen Einstellungen verschiedene Teilereignisse dieses *Accomplishments* repräsentiert. Abb. 3.23 zeigt diese Situation o. B. d. A. für den Fall mit zwei Einstellungen  $(T_i, T_{i+1})$  und zwei Teilereignissen  $e^{(1)}, e^{(2)}$  mit  $e^{(1)} < e$  und  $e^{(2)} < e$ .



Abb. 3.23: Subereignisse eines Accomplishments in Einstellungen

In der Abb. 3.23 links dargestellten Situation werde in jeder Einstellung genau ein Teilereignis gezeigt. In jeder der Einstellungen liegt also eine Realisation vom Typ  $\llbracket 1 \rrbracket^1$  vor, allerdings bezogen auf eine einzige Instanziierung eines Teilereignisses. Um dies zu kennzeichnen, wird der

Exponent dieses einzelnen Teilereignisses auch im Exponenten des Realisationstyps notiert, so dass für das obige Segment gilt:  $\llbracket 1 \rrbracket^{(1)} + \llbracket 1 \rrbracket^{(2)} = \llbracket 2 \rrbracket^1$ .

In der in Abb. 3.23 rechts dargestellten Situation endet in der ersten Einstellung ein Teilereignis und beginnt ein weiteres desselben instanziierten Gesamtereignisses. Notationell wird eine solche Situation, in der ein einziges Ereignis in dieser Weise auf eine Einstellung verteilt ist, wie folgt notiert:  $\llbracket 1 \rrbracket^{(1),(2)}$ . In der ersten Einstellung liegt keine Fusion vom Typ  $\llbracket 1 \rrbracket^2$  vor, da eine Fusion definitionsgemäß immer die Verknüpfung von zwei verschiedenen Ereignissen betrifft (die natürlich dieselbe Ereignisklasse instanziiieren können). Insgesamt gilt somit für das obige Segment:  $\llbracket 1 \rrbracket^{(1),(2)} + \llbracket 1 \rrbracket^{(2)} = \llbracket 2 \rrbracket^1$ .

Die bisherigen Überlegungen geben Informationen über den »inneren« Aufbau eines *Accomplishments*, umgekehrt muss man fragen, welche Randbedingungen erfüllt sein müssen, damit ein Beobachter eine Handlung, die über  $n$  Einstellungen in paarweise verschiedenen Raumzeitgebieten verteilt ist, als EINE Handlung konzeptionalisieren kann. Oder anders gefragt: Welche Randbedingungen muss ein Layoutprozess beachten, der aus  $n$  Einstellungen, in denen jeweils eine (Teil-)Handlung repräsentiert wird, ein Segment zusammenstellt.

Als erste Bedingung gilt, dass nicht jede beliebige Permutation der Einstellungen möglich ist, sondern – wie bereits zu Beginn dieses Abschnitts erwähnt – ist für ein Defaultlayout ist ein Ablauf so festzulegen, die Abfolgen auf der Leinwand diegetischen Abfolgen homomorph entsprechen.

Als zweites muss im Fall  $\llbracket n \rrbracket^1$  für einen Beobachter die Bindung der  $n$  Einstellungen an die Handlung dadurch erkennbar sein, dass die Ereignisteile unter ggf. verschiedenen thematischen Relationen auf dasselbe Ereignis bezogen werden.

Eine solche Ereignisbindung ist insbesondere gegeben, wenn der Agent und/oder das Patiens in einer Handlung in der Einstellungsfolge auf die gleichen Ereignisteile bezogen werden. Wenn Einstellungen daher MovingRegions enthalten, die den Agenten und das Patiens derselben Handlung repräsentieren, werden die Einstellungen für den Beobachter (i. Allg.) leicht als kohärent verbunden konzeptionalisiert werden können.

Etwas anderes ist es, wenn die paarweise räumlich unzusammenhängenden Raumzeitgebiete vom Beobachter so konzeptionalisiert werden, dass er sie in ein gemeinsames Raumzeitgebiet einbettet. Dies ist dann nur eine Kohärenz, die sich auf die abgebildeten Raumzeitgebiete bezieht und nicht auf die Kohärenz der Handlungen.

**Beispiel 3.7**

Als Beispiel wird nun die Handlung »Unterseite des Snowboards herstellen« betrachtet. Diese wird im Beispielfideo in der Einstellungsfolge  $S = (T18, T19, T20)$  repräsentiert. Syntagmatisch handelt es sich hierbei um drei Planszenen. Diegetischer Ort von  $T18$  ist Arbeitsstation 6 ( $S6$ ), von  $T19$  Arbeitsstation 7 ( $S7$ ) und von  $T20$  Arbeitsstation 8 ( $S8$ ). Abb. 3.24 zeigt eine Übersicht.

	<b>T18 – Ort: S6</b>	<b>T19 – Ort: S7</b>	<b>T20 – Ort: S8</b>
			
1	schutzpapier- entfernen_1 $e^{(1)}$  AG: akteur5 PAT: brett1	glasfaserschicht_aufbringen_1  $e^{(2)}$  AG: akteur6 + akteur7 + maschine5 PAT: brett1	fräsen_1  $e^{(3)}$  AG: maschine6 PAT: brett1
2	schutzpapier- entfernen_1 $e^{(1)}$  AG: akteur5 PAT: brett1	glasfaserschicht- auflegen_1 $e^{(2)}$  AG: akteur6 + akteur7 PAT: brett1	zusammenpressen_1  $e^{(3)}$  AG: maschine5 PAT: brett1
			fräsen_1  $e^{(4)}$  AG: maschine6 PAT: brett1

Abb. 3.24: Handlungen in  $(T18, T19, T20)$ 

Ein Beobachtermenge  $\mathcal{B}$  konzeptionalisiere in jeder Einstellung eine Teilhandlung (Zeile 1 des Bildtextes in Abb. 3.24): jeweils eine Instanz von SCHUTZPAPIER\_ENTFERNEN in  $T18$ , GLASFASERSCHICHT\_AUFBRINGEN in  $T19$  und FRÄSEN in  $T20$ . Für die segment-thematischen Relationen AG und PAT gilt in diesem Fall:

$$\begin{aligned}
 & AG_{S(\mathcal{L}, \mathcal{B}), \{\text{UNTERSEITE\_HERSTELLEN}\}} \\
 &= \{(s, e) \mid s \in S(\mathcal{L}, \mathcal{B}), \text{inst}_{\mathcal{B}}(e, E), E \in \mathcal{E}\} \\
 &= \{(s, e) \mid s \in S(\{\text{PERSON}, \text{MASCHINE}\}, \mathcal{B}), \\
 &\quad \text{inst}_{\mathcal{B}}(e, E), E \in \{\text{UNTERSEITE\_HERSTELLEN}\}\} \\
 &= \{(s_{18,1}^{\text{akteur5}}, e^{(1)}), (s_{19,1}^{\text{akteur6}}, e^{(2)}), (s_{19,2}^{\text{akteur7}}, e^{(2)}), (s_{19,3}^{\text{maschine5}}, e^{(2)}), (s_{20,1}^{\text{maschine6}}, e^{(3)}) \\
 &\quad | \text{inst}_{\mathcal{B}}(\text{akteur5}, \text{PERSON}), \text{inst}_{\mathcal{B}}(\text{akteur6}, \text{PERSON}), \text{inst}_{\mathcal{B}}(\text{akteur7}, \text{PERSON}), \\
 &\quad \text{inst}_{\mathcal{B}}(\text{maschine5}, \text{MASCHINE}), \text{inst}_{\mathcal{B}}(\text{maschine6}, \text{MASCHINE}), \\
 &\quad \boxed{\text{inst}_{\mathcal{B}}(e^{(1)}, \text{SCHUTZPAPIER\_ENTFERNEN}),} \\
 &\quad \boxed{\text{inst}_{\mathcal{B}}(e^{(2)}, \text{GLASFASSERSCHICHT\_AUFBRINGEN}),} \\
 &\quad \boxed{\text{inst}_{\mathcal{B}}(e^{(3)}, \text{FRÄSEN}),} \\
 &\quad \text{SCHUTZPAPIER\_ENTFERNEN, GLASFASSERSCHICHT\_AUFBRINGEN,} \\
 &\quad \text{FRÄSEN} < \text{UNTERSEITE\_HERSTELLEN} \}
 \end{aligned}$$

$$\begin{aligned}
 & \text{PAT}_{S(\mathcal{L}, \mathcal{B}), \{\text{UNTERSEITE\_HERSTELLEN}\}} \\
 &= \{(s, e) \mid s \in S(\mathcal{L}, \mathcal{B}), \text{inst}_{\mathcal{B}}(e, E), E \in \mathcal{E}\} \\
 &= \{(s, e) \mid s \in S(\{\text{BRETT}\}, \mathcal{B}), \\
 &\quad \text{inst}_{\mathcal{B}}(e, E), E \in \{\text{UNTERSEITE\_HERSTELLEN}\}\} \\
 &= \{(s_{18,2}^{\text{brett1}}, e^{(1)}), (s_{19,4}^{\text{brett1}}, e^{(2)}), (s_{20,2}^{\text{brett1}}, e^{(3)}) \\
 &\quad \mid \text{inst}_{\mathcal{B}}(\text{brett1}, \text{BRETT}), \\
 &\quad \boxed{\text{inst}_{\mathcal{B}}(e^{(1)}, \text{SCHUTZPAPIER\_ENTFERNEN}),} \\
 &\quad \boxed{\text{inst}_{\mathcal{B}}(e^{(2)}, \text{GLASFASSERSCHICHT\_AUFBRINGEN}),} \\
 &\quad \boxed{\text{inst}_{\mathcal{B}}(e^{(3)}, \text{FRÄSEN}),} \\
 &\quad \text{SCHUTZPAPIER\_ENTFERNEN, GLASFASSERSCHICHT\_AUFBRINGEN,} \\
 &\quad \text{FRÄSEN} < \text{UNTERSEITE\_HERSTELLEN} \}
 \end{aligned}$$

Die in drei Einstellungen repräsentierte Handlung UNTERSEITE\_HERSTELLEN mit ihren drei Teilhandlungen (in der obigen Darstellung zur Verdeutlichung eingerahmt) wird also in einer Folge von drei Einstellungen mit je einer Teilhandlung unter AG und PAT ereignisgebunden repräsentiert. Mit der oben eingeführten Notation kann man schreiben:  $\llbracket 3 \rrbracket_{\mathcal{B}}^1 = \llbracket 1 \rrbracket_{\mathcal{B}}^{(1)} + \llbracket 1 \rrbracket_{\mathcal{B}}^{(2)} + \llbracket 1 \rrbracket_{\mathcal{B}}^{(3)}$ .

Eine andere Beobachtermenge,  $\mathcal{B}'$ , konzeptionalisiere die in dem Segment dargestellte Handlung in detaillierterer Form (Zeile 2 des Bildtextes in Abb. 3.24). Für sie werden in den drei Einstellungen vier (Teil-)Handlungen repräsentiert: jeweils eine Instanz von SCHUTZPAPIER\_ENTFERNEN in T18, GLASFASERSCHICHT\_AUFLEGEN UND ZUSAMMENPRESSEN in T19 und FRÄSEN in T20. Die feingranularere Konzeptionalisierung wird in der folgenden Darstellung der segment-thematischen Relationen AG und PAT durch den inneren, blauen Rahmen visualisiert. Für die Relationen gilt:

$$\begin{aligned}
 & \text{AG}_{S(\mathcal{L}, \mathcal{B}'), \{\text{UNTERSEITE\_HERSTELLEN}\}} \\
 &= \{(s, e) \mid s \in S(\mathcal{L}, \mathcal{B}'), \text{inst}_{\mathcal{B}'}(e, E), E \in \mathcal{E}\} \\
 &= \{(s, e) \mid s \in S(\{\text{PERSON}, \text{MASCHINE}\}, \mathcal{B}'), \\
 &\quad \text{inst}_{\mathcal{B}'}(e, E), E \in \{\text{UNTERSEITE\_HERSTELLEN}\}\} \\
 &= \{(s_{18,1}^{\text{akteur5}}, e^{(1)}), (s_{19,1}^{\text{akteur6}}, e^{(2)}), (s_{19,2}^{\text{akteur7}}, e^{(2)}), (s_{19,3}^{\text{maschine5}}, e^{(3)}), (s_{20,1}^{\text{maschine6}}, e^{(4)}) \\
 &\quad \mid \text{inst}_{\mathcal{B}'}(\text{akteur5}, \text{PERSON}), \text{inst}_{\mathcal{B}'}(\text{akteur6}, \text{PERSON}), \text{inst}_{\mathcal{B}'}(\text{akteur7}, \text{PERSON}), \\
 &\quad \text{inst}_{\mathcal{B}'}(\text{maschine5}, \text{MASCHINE}), \text{inst}_{\mathcal{B}'}(\text{maschine6}, \text{MASCHINE}), \\
 &\quad \boxed{\text{inst}_{\mathcal{B}'}(e^{(1)}, \text{SCHUTZPAPIER\_ENTFERNEN}),} \\
 &\quad \boxed{\text{inst}_{\mathcal{B}'}(e^{(2)}, \text{GLASFASSERSCHICHT\_AUFLEGEN}),} \\
 &\quad \boxed{\text{inst}_{\mathcal{B}'}(e^{(3)}, \text{ZUSAMMENPRESSEN}),} \\
 &\quad \boxed{\text{inst}_{\mathcal{B}'}(e^{(4)}, \text{FRÄSEN}),} \\
 &\quad \text{SCHUTZPAPIER\_ENTFERNEN, GLASFASSERSCHICHT\_AUFLEGEN,} \\
 &\quad \text{ZUSAMMENPRESSEN, FRÄSEN} < \text{UNTERSEITE\_HERSTELLEN} \}
 \end{aligned}$$

$$\begin{aligned}
& \text{PAT}_{S(\mathcal{L}, \mathcal{B}'), \{\text{UNTERSEITE\_HERSTELLEN}\}} \\
&= \{(s, e) \mid s \in S(\mathcal{L}, \mathcal{B}'), \text{inst}_{\mathcal{B}'}(e, E), E \in \mathcal{E}\} \\
&= \{(s, e) \mid s \in S(\{\text{BRETT}\}, \mathcal{B}'), \\
&\quad \text{inst}_{\mathcal{B}'}(e, E), E \in \{\text{UNTERSEITE\_HERSTELLEN}\}\} \\
&= \{(s_{18,2}^{\text{brett1}}, e^{(1)}), (s_{19,(1)}^{\text{brett1}}, e^{(2)}), (s_{19,(2)}^{\text{brett1}}, e^{(3)}), (s_{20,2}^{\text{brett1}}, e^{(4)}) \\
&\quad \mid \text{inst}_{\mathcal{B}'}(\text{brett1}, \text{BRETT}), \\
&\quad \boxed{\text{inst}_{\mathcal{B}'}(e^{(1)}, \text{SCHUTZPAPIER\_ENTFERNEN}),} \\
&\quad \boxed{\text{inst}_{\mathcal{B}'}(e^{(2)}, \text{GLASFASSERSCHICHT\_AUFLEGEN}),} \\
&\quad \boxed{\text{inst}_{\mathcal{B}'}(e^{(3)}, \text{ZUSAMMENPRESSEN}),} \\
&\quad \text{inst}_{\mathcal{B}'}(e^{(4)}, \text{FRÄSEN}),} \\
&\quad \text{SCHUTZPAPIER\_ENTFERNEN, GLASFASSERSCHICHT\_AUFLEGEN,} \\
&\quad \text{ZUSAMMENPRESSEN, FRÄSEN} < \text{UNTERSEITE\_HERSTELLEN}\}
\end{aligned}$$

Die Beobachtermenge  $\mathcal{B}'$  konzeptionalisiert also vier Teilhandlungen in den drei Einstellungen, was wie folgt notiert werden kann:  $\llbracket 3 \rrbracket_{\mathcal{B}'}^1 = \llbracket 1 \rrbracket_{\mathcal{B}'}^{(1)} + \llbracket 1 \rrbracket_{\mathcal{B}'}^{(2),(3)} + \llbracket 1 \rrbracket_{\mathcal{B}'}^{(4)}$ .

Für beide Beobachtermengen bleibt zu prüfen, ob im Segment  $S = (T18, T19, T20)$  ein *Accomplishment* vorliegt. Unabhängig von der Granularität der Konzeptionalisierung lässt sich als Erstes festhalten, dass die Einstellungen in ihrer diegetischen Reihenfolge auf ein Ereignis UNTERSEITE\_HERSTELLEN bezogen, gezeigt werden. Für diese Ereignisklasse gilt im Segment  $S = (T18, T19, T20)$  die Subereignis-Abbildbarkeit und die Subobjekt-Abbildbarkeit. In allen Einstellungen sind die zu dieser Ereignisklasse in der PAT-Relation stehenden ShotRegions mit dem gleichen Label »brett1« gelabelt.<sup>1</sup> Daher ist die PAT-Relation auch objekt-eindeutig. Das Ereignis UNTERSEITE\_HERSTELLEN ist im Segment  $S = (T18, T19, T20)$  graduell (1,1)-PAT-verankerbar und stellt ein *Accomplishment* dar, da das betroffene Objekt gequantelt ist.

<sup>1</sup> Dies ist für die Beobachter einfach möglich, da das Muster »Blümchen« gut erkennbar ist.

## 4 Experimente mit SMIL und HTML5

### 4.1 Einleitung

In diesem Kapitel werden – ausgehend von einer Repräsentation der filmischen Einheiten und der Relationen zwischen ihnen – verschiedene multimediale Genre-Experimente für Videos auf Basis des Videos »Snowboard« ausprobiert. »Genre« wird hier zunächst im Sinne der Definition von [Goguen 1999, S. 246] verstanden:

*»A genre is a collection of conventions for using a medium; these can be seen as further delimiting a sign system. For example the daily newspaper is a genre within the medium of multisection collections of large size pages. Soap operas are a genre for TV. Obviously genres have subgenres; e.g. soap operas about rich families.«*

Diese verschiedenen experimentellen Genres nutzen:

- (einfache) Text-Bild-Kombinationen (Bildergeschichten, s. 4.6),
- Hypervideos (s. 4.7, 4.8, 4.9).

Einfache Text-Bild-Kombinationen sind auf fast allen Ausgabegeräten möglich und zu diesem Genre liegen – zu unterschiedlichen Aspekten – viele Forschungsergebnisse vor, z. B. [Muckenhaupt 1986], [Straßner 2002], [Stöckl 2004], [Bateman 2014].

Dies ist bei Hypervideos nicht der Fall. Hypervideos setzen naturgemäß eine zeitabhängige Ausgabe mit Linksstrukturen voraus, mit denen genremäßig auch noch experimentiert wird. So definieren [Sawhney et al. 1996, S. 1]:

*»“Hypervideo” is digital video and hypertext, offering to its user and author the richness of multiple narratives, even multiple means of structuring narrative (or non-narrative), combining digital video with a polyvocal, linked text.«*

In dieser Arbeit wird der Begriff im Sinne der Definition von [Finke 2005] verwendet (s. auch Abb. 4.1, Quelle: [Finke 2005, S. 23]):

*»Hypervideo wird als ein videobasiertes Hypermedia-Format definiert, das eine nicht-lineare Informationsstruktur mit einer dynamischen audio-visuellen Informationspräsentation kombiniert.«*

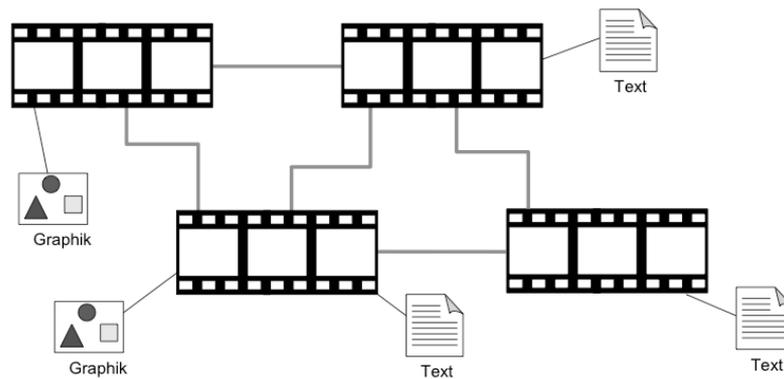


Abb. 4.1: Schematische Sicht eines Hypervideos

Die für die Experimente ursprünglich gegebene »physikalische« Ausgangsbasis (»Content Base«) ist in Abb. 4.2 dargestellt. Ein Video mit einer Audiospur liegt in einem digitalen Format vor. Zum Video wurde ein Transkriptionsprotokoll (in Word, s. 6.1) erstellt. Für jede Einstellung wurde ein Screenshot (statisches Bild) erzeugt, der im Transkriptionsprotokoll verwendet wird.

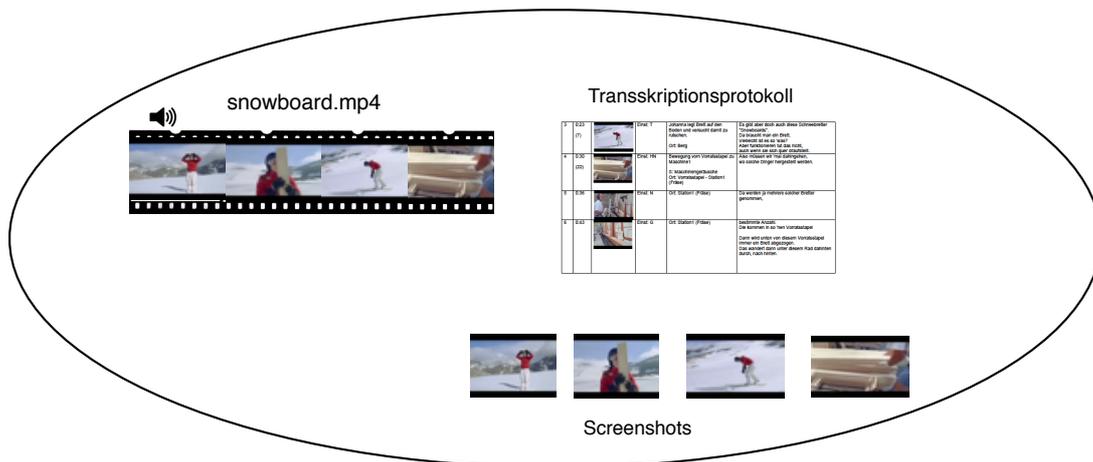


Abb. 4.2: Content Base (physikalisch)

In den verschiedenen Genre-Experimenten sollen Videoausschnitte, Bilder und Text frei kombiniert werden, wobei auf größtmögliche Wiederverwendbarkeit geachtet wird. Zur Ideenfindung und auch zur technischen Umsetzung wird von folgendem Modell ausgegangen: Für jede der 42 Einstellungen liegen *Content Portions* vor, je Einstellung eine Ressource vom Typ Video (incl. Audio), Image (Screenshot) und Text (des Sprechers) sowie notwendige Metadaten (s. Abb. 4.3). »Vorliegen« kann hier auf zwei Arten realisiert werden:

- (1) Die Ressource liegt physikalisch als abgeschlossene Einheit vor. Dies ist z. B. bei den Screenshots der Fall.
- (2) Die Ressource ist als »Teil« einer größeren Ressource direkt adressierbar. Dies wird in dieser Form beispielsweise für die filmischen Einheiten umgesetzt.

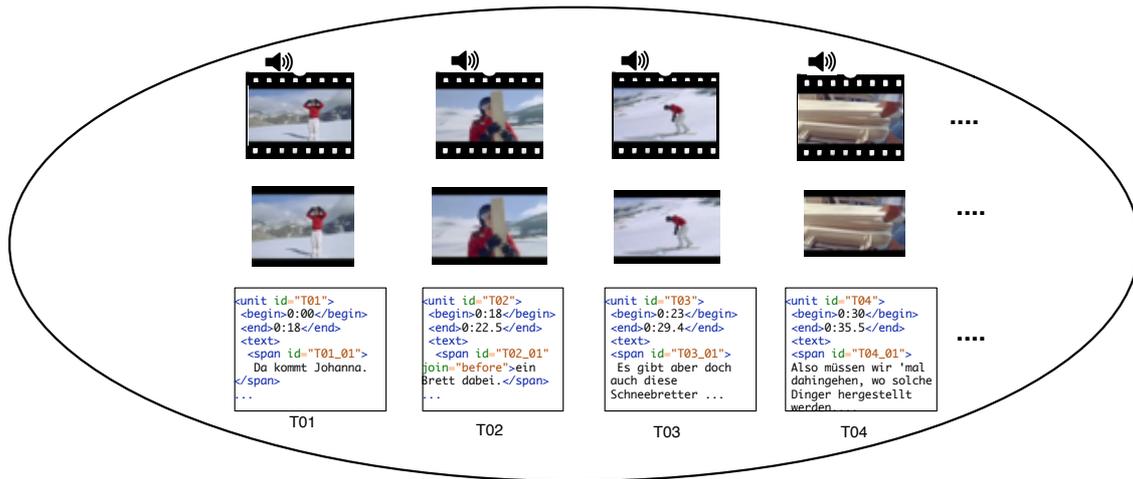


Abb. 4.3: Content Base bestehend aus »Content Portions«

Diese Materialbasis erlaubt es, kontinuierliche und seitenbasierte, sequentielle und parallele Ausgaben für unterschiedliche Ausgabekanäle zu erzeugen. Aus den reichhaltigen *Content Portions* des Tonfilms können so vielfältige »Genre-Experimente« erzeugt werden.

Weiterhin liegen (kurz zusammengefasst) auf Basis der Analysen der vorherigen Kapitel folgende Informationen zu den Einstellungen vor:

- Zeitliche (diegetische) Reihenfolge der Einstellungen,
- diegetischer Ort einer Einstellung, verteilt auf zwei diegetische (Haupt-)Handlungsorte: »Berg« und »Werkstatt«, wobei für den diegetischen Ort »Werkstatt« noch präzisere Informationen über die einzelnen Arbeitsstationen (Maschinen) vorliegen,
- vollständige syntagmatische Klassifikation:
  - Identifikation von Szenen und Sequenzen,
  - Identifikation eines Itinerars innerhalb des Raumes Werkstatt,
  - Identifikation einer Alternanz zwischen Berg- und Werkstatt-Segmenten mit den Kohärenzrelationen:
    - $R(\text{Johanna, Snowboard})$ : »Johanna gibt das Snowboard in die Werkstatt (zurück)«,
    - $R'(\text{Johanna, Snowboard})$ : »Johanna erhält das Snowboard zum Test«,
- Klassifikation der Text-Bild-Relationen zwischen Textsegmenten und filmischem Bild,
- Klassifikation der im Filmbild repräsentierten Handlungen.

Für die Aneinanderreihung von *Content Portions*, die unbeschränkt in vielen Fällen zu unverständlichen oder verwirrenden Dokumenten führen würde, werden auf Basis dieses Materials vorab Randbedingungen sowohl für ihre sequentielle als auch ihre parallele Anordnung hergeleitet (s. 4.2).

Technisch werden die verschiedenen Genre-Experimente, wie in Abb. 4.4 dargestellt, generiert.

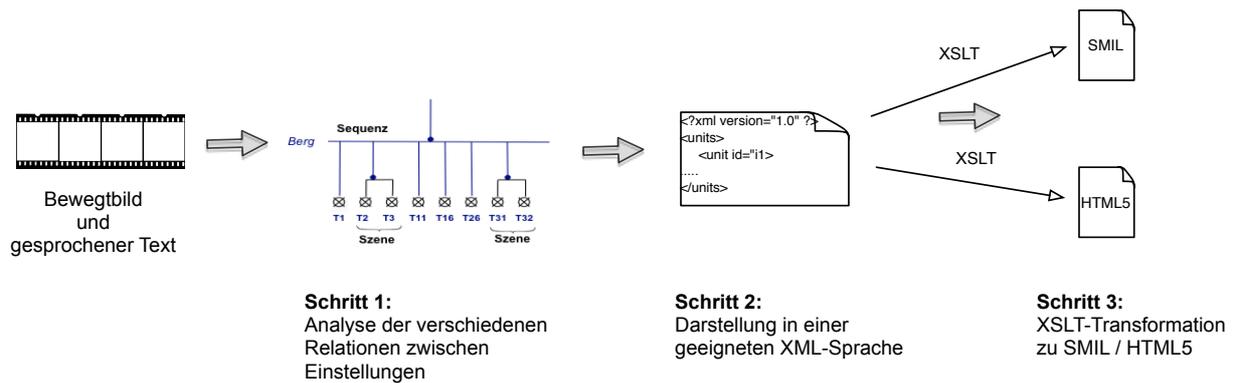


Abb. 4.4: Vorgehensweise zur Entwicklung weiterer multimedialer Genre-Experimente

- (1) Für die filmischen Einheiten und die Relationen zwischen ihnen wird eine XML-Darstellung entwickelt (Schritt 2 in Abb. 4.4).
- (2) Die Genre-Experimente werden mit Hilfe der XML-Anwendung SMIL (s. 4.4.1) und in HTML5 (s. 4.4.2) präsentiert. Die Zielformate SMIL und HTML5 werden ausgehend von (1) per XSLT (*extensible Stylesheet Language Transformations*)<sup>1</sup> generiert (Schritt 3 in Abb. 4.4). Zur Realisierung der Videosteuerung und diverser Interaktionen für das Zielformat HTML5 werden JavaScript-Funktionen genutzt.

## 4.2 Randbedingungen der Dokumentgenerierung

Wie in der Einleitung erläutert, kann zur Dokumentgenerierung von folgenden Ressourcen ausgegangen werden: Zu jeder Einstellung liegen *Content Portions* der folgenden Typen vor: Video (incl. Audio), Bild (Screenshot), Text (der Sprechertext, auch in Schriftform).

Die Erzeugung neuer Genres könnte experimentell ohne Randbedingungen erfolgen, indem eine zufällige Auswahl von *Content Portions* (eines Typs oder mehrerer Typen gleichzeitig) aus einem Pool gezogen und diese dann, z. B. entsprechend der gezogenen Reihenfolge, in einem Ausgabestrom sequentiell angeordnet werden. Geht man kombinatorisch in dieser Weise systematisch vor, könnten auch alle möglichen Permutationen und Kombinationen der *Content Portions* gebildet werden. Zum Beispiel lassen sich allein durch die Permutation der 42 *Content Portions* vom Typ Video  $42! - 1$  voneinander verschiedene »neue« Videos erzeugen. Dabei ist über die Art der Darstellung noch gar nichts gesagt. Jedoch wäre die Mehrzahl der Darstellungen für die Beobachter verwirrend oder unverständlich. Auch eine parallele Darstellung von *Content Portions* desselben oder verschiedenen Typs in mehreren Ausgabeströmen kann unverständlich sein, etwa wenn zu einem Bild ein Text präsentiert wird, der in keinerlei inhaltlichem Bezug zum Bild steht.

<sup>1</sup> XSLT ist eine XML-Sprache zur Transformation von XML-Dokumenten in andere (XML-)Dokumente. Verwendet wird hier die Version 2.0 (<http://www.w3.org/TR/xslt20> [letzter Aufruf: 04.03.2015]).

Es ist also notwendig, vorab Randbedingungen für neue, experimentelle Genres zu formulieren. Diese betreffen sowohl die sequentielle als auch die parallele Anordnung von *Content Portions* in einem Dokument.

#### 4.2.1 Randbedingungen der thematischen Entfaltung

Zur Entwicklung der Randbedingungen muss zunächst überlegt werden, welches kommunikative Ziel mit den neu erstellten experimentellen Genres erreicht werden soll. Insbesondere ist zu entscheiden, in welcher Art und Weise das Thema (im Beispiel immer »Bau/Test eines Snowboards«) thematisch entfaltet werden soll.

Nach [Brinker 2010, S. 56ff] sind die wichtigsten Grundformen der thematischen Entfaltung die deskriptive (beschreibende), explikative (erklärende), die narrative (erzählende) und die argumentative (begründende) Themenentfaltung.

Bei einer deskriptiven Themenentfaltung wird ein »*Thema in seinen Komponenten (Teilthemen) dargestellt und in Raum und Zeit eingeordnet*« [Brinker 2010, S. 56]. Typische Themen, bei denen eine deskriptive Themenentfaltung zu beobachten ist, sind die Darstellung eines einmaligen Vorgangs bzw. historischen Ereignisses, ein regelhaft (generalisierbar, wiederholbar) dargestellter Vorgang oder die Darstellung eines Lebewesens bzw. Gegenstandes. Eine deskriptive Themenentfaltung findet man daher bei informativen Texten (z. B. Nachricht, Bericht, Lexikonartikel), instruktiven Texten (z. B. Bedienungsanleitung, Kochrezept) und normativen Texten (z. B. Gesetz, Vertrag).

Die narrative Themenentfaltung ist charakteristisch für Alltagserzählungen. Nach [Brinker 2010, S. 62f] besteht ihre narrative Struktur aus den drei Grundkategorien Situierung, Repräsentation<sup>1</sup> und Resümee. Situierung meint die Einordnung des Themas. Situierende Elemente können an beliebigen Positionen einer Erzählung vorkommen oder sogar ganz entfallen. Im Mittelpunkt einer Erzählung steht »*die prozessual-aktionale Repräsentation des Ereignisses, die sich aus einer oder mehreren Ereignisphasen konstituiert. Jede Phase enthält als Kern eine Ereignis- bzw. Handlungssequenz, die sich aus einer Komplikation und einer Auflösung (Lösung) zusammensetzt; [...]*« [Brinker 2010, S. 62]. Unter Resümee versteht [Brinker 2010, S. 62] die »*zusammenfassende Einschätzung vom Erzählzeitpunkt (Gegenwartszeitpunkt) aus.*«

Eine explikative Themenentfaltung folgt i. Allg. dem Modell der wissenschaftlichen Erklärung von C. G. Hempel und P. Oppenheim (kurz: H-O-Schema). Nach diesem Modell erklärt ein Wissenschaftler einen Sachverhalt (Explanandum, das zu Erklärende) dadurch, dass er ihn aus bestimmten anderen Sachverhalten (Explanans, das Erklärende) logisch ableitet. Textsorten, in

<sup>1</sup> Der Begriff »Repräsentation« wird bei Brinker anderes verwendet als sonst in dieser Arbeit.

denen man die explikative Themenentfaltung findet, sind solche, die auf eine Erweiterung des Wissens zielen, z. B. in Lehrbüchern, populärwissenschaftlichen und wissenschaftlichen Texten.

Eine argumentative Themenentfaltung orientiert sich i. Allg. am Argumentationsmodell von St. Toulmin. Das Textthema wird durch eine These repräsentiert, welche durch Argumente begründet wird. Verwendet wird dazu eine (evtl. nur implizit vorhandene) Schlussregel, um zu zeigen, dass die angeführten Tatsachen Argumente für die These sein können. Die Zulässigkeit der Anwendung der Schlussregel wird durch Aussagen, wie z. B. Verweise auf Gesetze, Normen, Regeln, gestützt. Mit modalen Operatoren kann die Wahrscheinlichkeit der These angegeben werden. Eine Ausnahmebedingung kann die Umstände angeben, die den Geltungsbereich der Schlussregel einschränken. Textsorten, in denen diese Form der Themenentfaltung anzutreffen ist, sind z. B. appellative Texte, normative Texte (z. B. Gerichtsentscheidungen) oder informative Texte (z. B. Rezensionen).

Diese Grundformen sind jedoch nicht als disjunkte Kategorien zu verstehen.<sup>1</sup> So gibt es z. B. häufig eine Verbindung von explikativer und deskriptiver oder explikativer und argumentativer Themenentfaltung [Brinker 2010, S. 69]. Je nachdem, welche Grundform dominiert, wird ein Text dann primär dieser Kategorie zugeordnet [Brinker 2010, S. 56].

Wendet man gleichwohl diese Kategorisierung der Grundformen thematischer Entfaltung auf das Beispielvideo »Snowboard« an, gelangt man zu dem in Abb. 4.5 dargestellten Ergebnis.

Im Beispielvideo lassen sich leicht zwei Teildokumente identifizieren. Ein Teildokument zeigt die Herstellung des Snowboards in der Werkstatt und stellt die dazu notwendigen Handlungen dar. Es kann im obigen Sinne als deskriptiv klassifiziert werden. Das zweite Teildokument zeigt Johanna, die Snowboard fahren möchte. Es beginnt mit der Situierung, der Vorstellung von Johanna und ihrer Gebrauchssituation. Die Repräsentation besteht aus mehreren Phasen. In jeder versucht Johanna mit dem (teilweise fertiggestellten) Snowboard zu fahren, jedoch tritt eine Komplikation auf, das Fahren funktioniert nicht. Als (Auf-)Lösung wird das Snowboard in die Werkstatt gegeben, was allerdings im Video nicht explizit gezeigt wird. Dieser jeweilige Übergang von Komplikation zu Teillösung stellt eine Verknüpfung zwischen den beiden Teildokumenten her<sup>2</sup>. Das Teildokument mit Johanna und der ganze Film enden mit »nun endlich funktioniert es und Johanna kann fahren« – und der ganze Film kann resümiert werden.

---

<sup>1</sup> In der Diskurstheorie trifft man häufig auf nicht-disjunkte Kategorien.

<sup>2</sup> In Kapitel 2 wird diese Relation als  $R(\text{Johanna}, \text{Snowboard})$ : »Johanna gibt das Snowboard in die Werkstatt (zurück)« identifiziert.

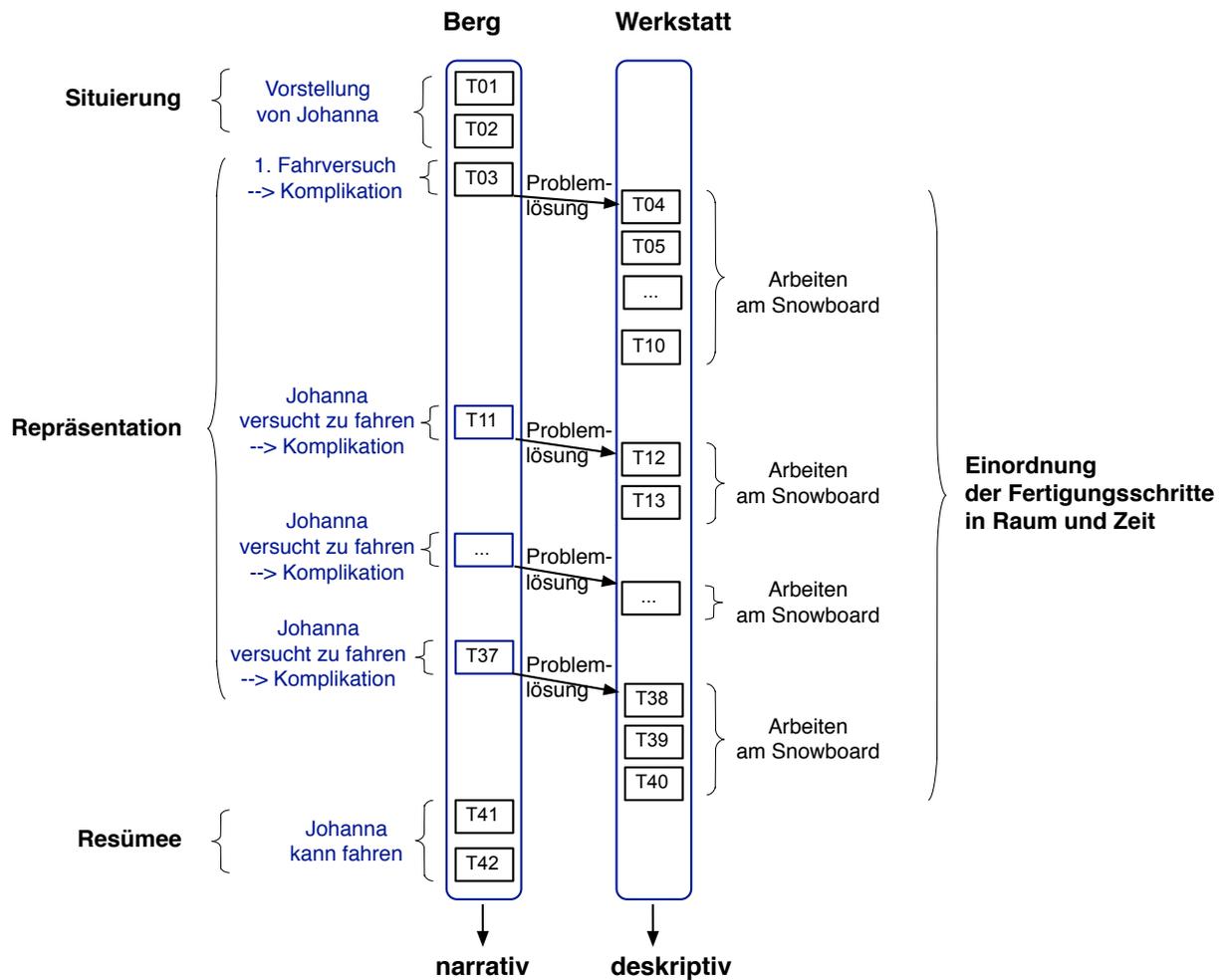


Abb. 4.5: Thematische Entfaltung im Video »Snowboard«

Insgesamt folgt, dass das Gesamtdokument sowohl Elemente einer deskriptiven als auch einer narrativen Themenentfaltung enthält, wobei das deskriptive Teildokument zeitlich umfangreicher ist. Die Frage, welche Grundform dominiert, wird von verschiedenen Beobachtern vermutlich unterschiedlich beantwortet. Es ist die deskriptive Themenentfaltung, wenn der Beobachter als Thema des Videos »Bau eines Snowboards« konzeptionalisiert. Steht für den Beobachter jedoch das Thema »Johanna möchte mit einem für sie gebauten Snowboard fahren« im Vordergrund, dominiert für ihn die narrative Themenentfaltung.

Mit dem vorhandenen Material ist also grundsätzlich eine deskriptive und/oder narrative Themenentfaltung möglich. Je nach ausgewählten Inhalten und deren Anordnung wird tendenziell die deskriptive oder die narrative Grundform verstärkt.

#### 4.2.2 Randbedingungen der Anordnung

Soll eine narrative oder deskriptive Darstellung für Beobachter plausibel sein, ist eine zufällige Reihenfolge, wie oben geschildert, im Film i. Allg. ungeeignet, da eine kausal-chronologische Abfolge von Ereignissen sich aus einer zufälligen Anordnung schwer ermitteln lässt. Für eine nicht-zufällige Anordnung von *Content Portions* werden zunächst die Randbedingungen ihrer sequentiellen, anschließend ihrer parallelen Anordnung behandelt.

##### Sequentielle Anordnung von *Content Portions*

Da eine kausal-chronologische Abfolge ein zentraler Punkt für das Verständnis einer narrativen Themenentfaltung – unabhängig von ihrer Darstellung – ist, ist ein Beobachter bei der Zerlegung in »narrativ sinnstiftende« Strukturen zu unterstützen. Dies wird erleichtert durch die Darstellung von »Beziehungen« insbesondere zwischen aufeinanderfolgenden Einstellungen, die ein Ereignis oder eine Folge von Ereignissen wiedergeben. Nach [Bateman 2013, S. 63] bietet das filmische Material potentiell sehr viele Elemente zu deren Identifikation an:

*»These cohesive<sup>1</sup> qualities of the filmic-material provide richly organized linkages across segments as a film unfolds [...].«*

Wie in 2.2.4 ausgeführt, gehören hierzu spezifisch filmisch u. a. einstellungsinterne Verweise, die innerhalb einer Einstellung oder über mehrere Einstellungen hinweg diese miteinander in Beziehung setzen, sowie einstellungsexterne Mittel, wie z. B. Musik, die Teile einer oder mehrerer Einstellungen beobachterseits als kohärent identifizierbar machen [Hickethier 2007, S. 113]. Die zu formulierenden Randbedingungen müssen also sicherstellen, dass für die Beobachter kohärente Elemente nicht ohne Grund aufgelöst werden.

Zur Identifikation und Klassifikation dieser kohärenten Elemente werden in dieser Arbeit die syntagmatische und die handlungstheoretische Analyse genutzt.

Die syntagmatische Analyse liefert Informationen über die logische Struktur eines cinematographischen Dokumentes ausgehend von den raum-zeitlichen Zusammenhängen zwischen den Einstellungen. Im Folgenden werden die drei narrativ basalen Syntagmen Szene, Sequenz, Planszene genutzt, wobei jedem Syntagma ein Default im Layout zugeordnet werden kann. Es gilt:

- Szene: Das Defaultlayout einer Szene ist das Hintereinanderabspielen der homomorph angeordneten Einstellungen.
- Sequenz: Das Defaultlayout einer Sequenz ist ebenfalls das Hintereinanderabspielen der Einstellungen. Von allen Layoutprozessen, die etwas einschieben, werden die bevorzugt, die es in den »zeitlichen Lücken« tun.

---

<sup>1</sup> In der Linguistik versteht man unter Kohäsion »die Verknüpfung der Oberflächenelemente des Textes durch bestimmte grammatische Mittel« [Brinker 2010, S. 17]. In dieser Arbeit wird, in Anlehnung an Brinker, nur der Begriff Kohärenz verwendet (s. auch 2.2.3).

- Planszene: Das Defaultlayout einer Planszene als einzelner Einstellung ist kanonisch.

Alternanzen werden erzeugt, indem zwischen Einstellungen dieser basalen Syntagmen bei Vorliegen einer Kohärenzrelation im Layout gewechselt wird.<sup>1</sup> Das Defaultlayout einer Alternanz sollte wenigstens externe Relationen respektieren, d. h. zwischen aufeinanderfolgenden Einstellungen sollte die paradigmatische Relation *paratactic – external | contrast* einen Wechsel verursachen.

Die bisherige Begriffsbildung gestattet es, Dokumente, deren zu layoutierende Strukturen sich als Planszenen, Szenen oder Sequenzen klassifizieren lassen, hinsichtlich ihres Layouts Randbedingungen zu unterwerfen. Dabei ist »Planszene« bisher noch eine Residualkategorie, die das aufnimmt, was nicht szenisch oder sequentiell gebunden ist.

Mit der Begriffsbildung aus Kapitel 3 ist es nun ergänzend möglich, z. B. Itinerare von Personen oder Gegenständen adäquat zu beschreiben, die an keiner Stelle eine Szene oder Sequenz enthalten. Auch der Begriff der Planszene ist nun algebraisch formalisiert, so dass ein handlungsorientierter Begriffsapparat für die Festlegung von Defaultlayouts zusätzlich zur Verfügung steht. Die in Kapitel 3 durchgeführte handlungstheoretische Analyse betrachtet die Darstellung einer Handlung in Einstellungen sowie ggf. den räumlichen Zusammenhang der Einstellungen. In diesem Kontext werden nur die Realisationstypen  $\llbracket 1 \rrbracket^1$ , d. h. eine ganze Handlung wird vollständig in einer Einstellung präsentiert, und  $\llbracket n \rrbracket^1$ , d. h. eine Handlung wird in  $n$  Einstellungen dargestellt, wobei auch Raumsprünge vorliegen müssen, verwendet.

Aus dem Handlungstyp der in den Einstellungen repräsentierten Handlung (*State*, *Activity*, *Achievement* oder *Accomplishment*) ergeben sich ggf. weitere Restriktionen für die Default-Reihenfolge der Einstellungen im Layout:

- Für *State* und *Activity* ist die Einstellungsreihenfolge, wenn keine speziellen gestalterischen Aspekte berücksichtigt werden müssen, für die konzeptionalisierte Handlung im Default nicht relevant.
- Einem *Achievement* können Prozesse vorausgehen, müssen aber nicht. In dieser Arbeit wird auf jeden Fall davon ausgegangen, dass ein *Achievement* als Realisationstyp  $\llbracket 1 \rrbracket^1$  realisiert wird.
- Ein *Accomplishment* ist eine Handlung, die auf ein Ziel gerichtet ist: Sie ist stufenweise fortschreitend und erreicht einen natürlichen Endpunkt. Ein *Accomplishment* kann in einer Einstellung (Realisationstyp  $\llbracket 1 \rrbracket^1$ ) oder in mehreren Einstellungen (Realisationstyp  $\llbracket n \rrbracket^1$ ) realisiert werden. Wird es in einer Einstellung realisiert, muss man sich über den homomorphen Ablauf keine Gedanken machen: »Eine Planszene bildet den zeitlichen Ab-

---

<sup>1</sup> Ein zufälliges Aneinandergeraten von Einstellungen aus basal narrativen Syntagmen ohne beobachterseitige Etablierung einer Kohärenzrelation ist keine Alternanz.

*lauf homomorph ab*« [Schmidt 2004, S. 262]. Wird es in mehreren Einstellungen realisiert, dann sollte die gezeigte Reihenfolge der einzelnen Handlungsschritte der Reihenfolge in der Diegese entsprechen (s. 3.8).

Darauf aufbauend werden nun die Randbedingungen für die sequentielle Anordnung von *Content Portions* entwickelt. Diese Randbedingungen gelten unter der Voraussetzung, dass die repräsentierten Handlungen vom Typ *Accomplishment* sind. Für die Genre-Experimente ist dies keine Einschränkung, da alle Handlungen im Video »Snowboard« von diesem Typ sind (s. 3.7).

Da *Accomplishments* zeit-diegetisch treu abgebildet werden sollen, wird als erste Regel notiert, dass Handlungen im Layout homomorph abgebildet werden sollen<sup>1</sup> (→ Randbedingung S1).

Für die syntagmatischen Kategorien lässt sich aus dem bisher Gesagten Folgendes ableiten: Um das Fortschreiten und Erreichen des natürlichen Endpunktes eines *Accomplishments* bestmöglich zu visualisieren, sollten Szenen, die ein *Accomplishment* enthalten, ohne Unterbrechung gezeigt werden (→ Randbedingung S2). Aus dem gleichen Grund sollten Sequenzen nur dort unterbrochen werden, wo eine zeit-diegetische Lücke für den Beobachter plausibel ist<sup>2</sup> (→ Randbedingung S3).

Ist die in einem Itinerar gezeigte Handlung ein *Accomplishment*, soll die Reihenfolge, in der die verschiedenen Orte durchlaufen werden, nicht permutiert werden (→ Randbedingung S4).

Eine Alternanz verknüpft i. Allg. zwei Handlungsstränge, wobei beobachterseits eine Kohärenzrelation vorliegen muss (s. ob.). Nicht in jedem Fall wird ein einzelner, isoliert gezeigter Handlungsstrang in sich verständlich sein. In diesem Falle kann das Hinzufügen einer Alternanz erfolgen, wenn eine Kohärenzrelation beobachterseits etabliert werden kann und die vorhandenen Segmente auf diese Art und Weise inhaltlich verknüpft werden. Wird eine Alternanz etabliert, müssen die Definitionskriterien eines alternanten Segmentes (s. Definition 2.13) natürlich erfüllt werden (→ Randbedingung S5).

Insgesamt folgen damit:

#### *Randbedingungen für die sequentielle Anordnung von Content Portions*

- (S1) Die Ordnung der Handlungen wird im Layout homomorph abgebildet.
- (S2) Szenen werden im Layout nicht auseinandergerissen.
- (S3) Sequenzen werden nur dort auseinandergerissen, wo eine zeit-diegetische Lücke beobachterseits angenommen werden kann.

---

<sup>1</sup> Würde (S1) und (S4) ganz streng interpretiert, dürfte in den Bildergeschichten und in den Hypervideos lediglich ein Vorwärtsblättern realisiert werden. Die Interaktionen des Rückwärtsblätterns bzw. des gezielten Zugriffs auf ein Kapitel über das Inhaltsverzeichnis stellen eine zusätzliche Interaktionsmöglichkeit dar.

<sup>2</sup> Bei einer Sequenz kann das Defaultlayout durch Einschübe in den zeitlichen Lücken unterbrochen werden, darauf wird jedoch aus o. a. Grund verzichtet.

- (S4) Die Orte in einem Itinerar werden in ihrer diegetischen Reihenfolge durchlaufen.
- (S5) Eine Alternanz darf etabliert werden, wenn eine Kohärenzrelation beobachterseits konzeptionalisiert werden soll und kann.

### **Parallele Anordnung von *Content Portions***

Parallele Anordnung von *Content Portions* bezieht sich hier auf das gleichzeitige Anzeigen von *Content Portions* beliebigen Typs in einem Viewport mit mehreren Ausgabeströmen und grundsätzlich beliebigen Bild-, Text- und Video-Kombinationen. Damit die parallele Anordnung keine reine Layoutanordnung ist, wird verlangt, dass alle *Content Portions*, die in einem parallelen Layout präsentiert werden, für den Beobachter inhaltlich keine Inkonsistenzen zulassen. Dabei ist im Folgenden die Anordnung der filmischen *Content Portions* leitend in dem Sinne, dass für eine Menge von parallel dargestellten *Content Portions* mit genau einer cinematographischen *Content Portion* beobachterseits keine Inkonsistenz zwischen der filmischen *Content Portion* und den anderen *Content Portions* etabliert werden kann.

#### *Text-Text-Kombinationen*

Bei einer zeitlich parallelen Anordnung von *Content Portions* vom Typ Text muss berücksichtigt werden, dass ein Leser einen Text sequentiell liest, wobei die Lesereihenfolge von deutschen Texten von links nach rechts und von oben nach unten ist. Daher sollte jede *Content Portion* vom Typ Text jeweils in einem Block angezeigt werden und die Blöcke entsprechend ihrer diegetischen Reihenfolge von links nach rechts und von oben nach unten angeordnet werden (→ Randbedingung P1).

#### *Text-Bild/Video-Kombinationen*

Die Analyse der Text-Bild-Relationen (s. 2.5.4) zeigt, dass eine enge Verknüpfung zwischen Sprechertext und Filmbild besteht. So liefert der Sprechertext häufig gleiche oder ergänzende Informationen zum Filmbild oder enthält direkte Verweise auf das Bild. Es macht daher im Kontext dieser Arbeit keinen Sinn, beliebige Text-Bild-Kombinationen zuzulassen, wie es mit der Vorgabe einer dramaturgischen Leitung durch *Content Portions* vom Typ Video schon festgelegt wurde. Zusätzlich sollten jeweils Text und Bild aus dem gleichen diegetischen Zeitintervall kombiniert werden (→ Randbedingung P2).

Beim Sprechertext gibt es einige Sätze, die über Einstellungsgrenzen hinweg gehen. Wird dies in einer Text-Bild-Anzeige 1:1 umgesetzt, werden z. B. im visuellen Layout abgeschnittene, halbierte Sätze angezeigt. Um den Beobachter nicht zu verwirren, sollte erkennbar sein, dass es sich in diesem Fall um ein einstellungsübergreifendes Textsegment handelt. (→ Randbedingung P3).

#### *Bild-Bild-, Bild-Video- und Video-Video-Kombinationen*

Bild-Bild-, Bild-Video- und Video-Video-Kombinationen, die *Content Portions* aus dem gleichen zeit-diegetischen Intervall enthalten, zeigen ggf. identische Information mehrfach an. Solche Kombinationen werden einen Beobachter nicht verwirren, aber evtl. langweilen. Interessanter sind Kombinationen mit *Content Portions* aus verschiedenen zeit-diegetischen Intervallen, da mit ihnen z. B. Visualisierungen von Zustandsänderungen oder des Arbeitsfortschrittes möglich sind. Jedoch können zu große zeitliche Lücken zwischen den angezeigten *Content Portions* wieder unverständlich werden, wenn etwa für den Beobachter »Handlungslücken« entstehen. Aus diesem Grund sollen nur Kombinationen angezeigt werden, die aus aufeinanderfolgenden<sup>1</sup> zeit-diegetischen Intervallen stammen (→ Randbedingung P4). Dies ist z. B. der Fall, wenn zwischen den *Content Portions* die Zeitrelation *meets* vorliegt (s. Tab. 2.2).

Insgesamt ergeben sich:

#### *Randbedingungen für die parallele Anordnung von Content Portions*

- (P1) Bei Text-Text-Kombinationen sollen die *Content Portions* in Blöcken entsprechend ihrer diegetischen Reihenfolge von links nach rechts und von oben nach unten angeordnet werden.
- (P2) Text-Bild-Kombinationen sollen aus dem gleichen diegetischen Zeitintervall stammen.
- (P3) Einstellungsübergreifende Textsegmente sollten erkennbar sein.
- (P4) Parallel gezeigte *Content Portions* vom Typ Bild und Video sollen aus zeit-diegetisch aufeinanderfolgenden Einstellungen stammen.

### 4.3 XML-Repräsentation der filmischen Einheiten und ihrer Relationen

Dieser Abschnitt erläutert die XML-Spezifikationen zur Repräsentation der filmischen Einheiten und der Analysen der in Kapitel 2 behandelten paradigmatischen, syntagmatischen sowie Text-Bild-Relationen. Zudem werden diverse XSLT-Stylesheets vorgestellt, die als Arbeitserleichterung für die Beobachter Vorlagen (*templates*) für die Analysen erzeugen.

Die Informationen über die Einstellungen und ihre Relationen werden jeweils getrennt gespeichert. Für die verschiedenen XML-Repräsentationen werden geeignete *Document Type Definitions* (DTDs) entwickelt. Tab. 4.1 zeigt alle DTDs in einer Übersicht.

Informationen über die filmischen Einheiten (Einstellungen)	units.dtd
Syntagmatische Analyse	syntagmatic_analysis.dtd
Paradigmatische Relationen	paradigmatic_relation.dtd
Paradigmatische Analyse	paradigmatic_analysis.dtd
Taxonomie der Text-Bild-Relationen	textImage_relation.dtd
Analyse der Text-Bild-Relationen	textImage_analysis.dtd

Tab. 4.1: Übersicht: DTDs für die XML-Repräsentation der Einstellungen und Relationen

<sup>1</sup> Hierbei wird angenommen, dass eine plausible Zeitlogik des Beobachters mit passender Granularität vorliegt.

Wünschenswert wäre es, wenn die XML-Repräsentationen der verschiedenen Relationstypen in einem »einheitlichen Stil« wären. Da die Relationstypen von ihren Grundkonzepten jedoch sehr verschieden sind (eine syntagmatische Analyse liefert eine Baumstruktur, paradigmatische Relationen bilden ein systemisches Netzwerk, Text-Bild-Relationen bilden eine Taxonomie), wird nicht versucht, »künstlich« eine Vereinheitlichung zu erreichen, sondern für die jeweilige Anwendung wird eine möglichst passende und auch für die Beobachter gut handhabbare Darstellungsform gewählt.

#### 4.3.1 XML-Repräsentation der filmischen Einheiten (Einstellungen)

Die XML-Darstellung der filmischen Einheiten basiert auf der DTD `units.dtd` (s. Code 4.1).

```

<!ELEMENT movie           (movie_info,units)>
<!ELEMENT movie_info     (name,file+)>
<!ELEMENT name           (#PCDATA)>
<!ELEMENT file           (#PCDATA)  <!-- Dateiname des Videos-->
<!ELEMENT units          (unit+)>
<!ELEMENT unit           (begin,end,location,text)>
<!ELEMENT begin          (#PCDATA)>
<!ELEMENT end            (#PCDATA)>
<!ELEMENT location       (loc_name,loc_detail*)>
<!ELEMENT loc_name       (#PCDATA)>
<!ELEMENT loc_detail     (#PCDATA)>
<!ELEMENT text           (span+)>  <!-- Text aus Protokoll -->
<!ELEMENT span           (#PCDATA)>

<!ATTLIST file           type    (mp4|ogg|webm)  #REQUIRED>
<!ATTLIST unit           id      ID             #REQUIRED>
<!ATTLIST span           id      ID             #REQUIRED
                        join    (next|before)  #IMPLIED >

```

Code 4.1: DTD zur Repräsentation der filmischen Einstellungen: `units.dtd`

Direkte Kindelemente des Rootelementes `<movie>` sind `<movie_info>` zur Erfassung von Informationen über das Video (Name und Dateiname der Videodatei<sup>1</sup>) und `<units>` mit den Kindelementen `<unit>` zur Repräsentation der Eigenschaften der Einstellungen.

Ein Element `<unit>` enthält Informationen einer Zeile des Transkriptionsprotokolls<sup>2</sup>:

- Für jede Einstellung werden ihre Beginn- und Endezeit in den Elementen `<begin>` bzw. `<end>` erfasst.
- Für den diegetischen Handlungsort ist das Element `<location>` vorgesehen. In seinem Kindelement `<loc_name>` wird der Name des Ortes erfasst. Das optionale Element `<loc_detail>` kann für weitere, präzisere Angaben zum Ort verwendet werden. Mit diesen Elementen kann beim Video »Snowboard« in `<loc_name>` der Raum »Berg« bzw. »Werk-

<sup>1</sup> Hier sind mehrere Dateinamen für verschiedene Videoformate möglich.

<sup>2</sup> Nicht erfasst werden Informationen wie Kameraposition und Einstellungsgröße. Diese sind filmwissenschaftlich zwar sehr wichtig, werden aber für die Generierung der Dokumente nicht benötigt.

statt« und in `<loc_detail>` die genaue Arbeitsstation (Maschine) innerhalb des Raumes »Werkstatt« erfasst werden.

- Zur Erfassung des transkribierten Textes ist das Element `<text>` vorgesehen. Da auch die Relationen zwischen Text und Bild analysiert und erfasst werden und in vielen Einstellungen verschiedene Text-Bild-Relationen vorkommen (s. 2.5.4), enthält das Element `<text>` 1 bis beliebig viele `<span>`-Elemente für die Textsegmente.
- Das Element `<unit>` verfügt über das Pflichtattribut `id` vom Typ ID zur Erfassung der Einstellungsnummer. Diese ist in der Form `Txx`, mit `xx` = zweistellige Zahl, anzugeben.
- Auch das Element `<span>` besitzt ein Pflichtattribut `id` vom Typ ID. Der Attributwert ist hier in der Form `Txx_yy` anzugeben, wobei `Txx` die `id` der Unit ist und über `yy` die Segmente innerhalb einer Unit eindeutig durchnummeriert werden. Zur Kennzeichnung einstellungsübergreifender Textsegmente hat `<span>` das optionale Attribut `join`. Der Attributwert `next` bzw. `before` bedeutet, dass dieses Textsegment mit dem nächsten bzw. vorigen Textsegment verbunden ist.

Der folgende Code 4.2 zeigt einen Ausschnitt aus der XML-Repräsentation des Videos »Snowboard«. Es ist erkennbar, dass von `T01` zu `T02` ein einstellungsübergreifender Sprechertext vorliegt.

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- XML-Repräsentation der filmischen Einheiten für das Video Snowboard-->
<!DOCTYPE movie SYSTEM "../dtd/units.dtd">
<movie>
  <movie-info>
    <name>Snowboard</name>
    <file type="mp4">video/snowboard.mp4</file>
    <file type="ogg">video/snowboard.ogg</file>
    <file type="webm">video/snowboard.webm</file>
  </movie-info>

  <units>
    <unit id="T01">
      <begin>0:00</begin>
      <end>0:18</end>
      <location>
        <loc_name>Berg</location_name>
      </location>
      <text>
        <span id="T01_01">Da kommt Johanna. </span>
        <span id="T01_02">Ausgerüstet wie der Teufel mit Sturzhelm und Schneebrille, </span>
        <span id="T01_03">und die will den Berg 'runterfahren. </span>
        <span id="T01_04">Aber womit? Vielleicht will sie skifahren? </span>
        <span id="T01_05" join="next">Und dazu hat sie jetzt erstmal </span>
      </text>
    </unit>

    <unit id="T02">
      <begin>0:18</begin>
      <end>0:22.5</end>
      <location>
        <loc_name>Berg</location_name>
      </location>
      <text>
```

```

    <span id="T02_01" join="before">ein Brett dabei. </span>
    <span id="T02_02">Aber zum Skifahren gehören doch zwei Bretter? </span>
  </text>
</unit>
...
</units>
</movie>

```

Code 4.2: Ausschnitt aus snowboard.xml

### 4.3.2 XML-Repräsentation der syntagmatischen Analyse

Die XML-Repräsentation einer syntagmatischen Analyse (s. 2.5.3) basiert auf der DTD syntagmatic\_analysis.dtd (s. Code 4.3).

```

<!ELEMENT syntagmaticAnalysis (info,analysis)>
<!ELEMENT info (file)>
<!ELEMENT file (#PCDATA)>
<!ELEMENT analysis (segment+)>
<!ELEMENT segment (ref+)>
<!ELEMENT ref EMPTY>

<!--
  segment id ID #REQUIRED
  segment class (unclassified|scene-shot|scene|sequence|itinerary
  |alternation) "unclassified">
-->
<!--
  ref xref NMTOKEN #REQUIRED
-->

```

Code 4.3: DTD zur Repräsentation der syntagmatischen Analyse: syntagmatic\_analysis.dtd

Innerhalb des Elementes `<syntagmaticAnalysis>` werden die Segmente erfasst (Elemente `<segment>`). Jedes Element `<segment>` verfügt über 1 bis beliebig viele Kindelemente `<ref>`. Ein Element `<ref>` verweist auf Einstellungen oder bereits erfasste Segmente, deren `id`-Wert im Attribut `xref` notiert wird. Somit ist die Abbildung des Dokumentbaums einschließlich der narrativen Hierarchisierung möglich. Die syntagmatische Klasse wird als Wert des Attributes `class` notiert.

Abb. 4.6 zeigt einen Ausschnitt aus der syntagmatischen Analyse des Videos »Snowboard« (s. 2.5.3):  $S_B = (T01, T02, T03, T11, T16, T36, T41, T42)$ , das Segment der Einstellungen auf dem Berg. Es enthält zwei Szenen, die Segmente  $S_{B_01} = (T02, T03)$  und  $S_{B_02} = (T41, T42)$ , als Unterelemente. Die entsprechende XML-Repräsentation zeigt Code 4.4.

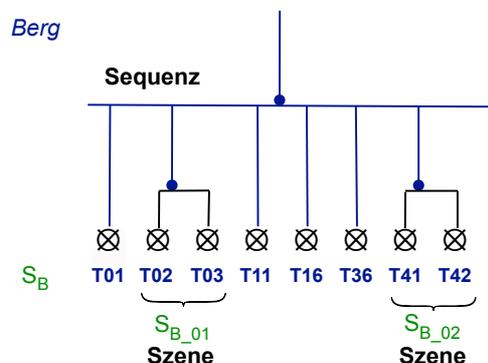


Abb. 4.6: Syntagmatische Analyse der Einstellungen auf dem Berg



Repräsentation zu den in 2.5.2 identifizierten paradigmatischen Relationen. Die Werte simultaner Auswahlmöglichkeiten werden durch einen »|« getrennt<sup>1</sup>.

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- Erfassung der vorkommenden paradigmatischen Relationen -->
<!DOCTYPE paraRelations SYSTEM "../dtd/paradigmatic_relation.dtd">
<paraRelations>
  <relation id="P1">
    <projection>non-projecting</projection>
    <taxis>hypotactic - extending</taxis>
    <plane>diegtic - event
      <temporal>ellipsis - forward | measurable</temporal>
      <spatial>contiguous - inclusion - narrowing</spatial>
      specific
    </plane>
  </relation>
  <relation id="P2">
    <projection>non-projecting</projection>
    <taxis>hypotactic - extending</taxis>
    <plane>diegtic - event
      <temporal>continuous</temporal>
      <spatial>contiguous - inclusion - broadening</spatial>
      specific
    </plane>
  </relation>
  ...
</paraRelations>
```

Code 4.6: Ausschnitt aus *paradigmatic\_relation.xml*

Die XML-Repräsentation einer paradigmatischen Analyse basiert auf der DTD *paradigmatic\_analysis.dtd* (s. Code 4.7).

```
<!ELEMENT paradigmaticAnalysis (info,analysis)>
<!ELEMENT info (file)>
<!ELEMENT file (#PCDATA)>
<!ELEMENT analysis (relation|heading)*>
<!ELEMENT relation EMPTY>
<!ELEMENT heading (#PCDATA)>

<!ATTLIST relation
  from NMTOKEN #REQUIRED
  to NMTOKEN #REQUIRED
  rel-id NMTOKEN #REQUIRED>
```

Code 4.7: DTD zur Repräsentation der paradigmatischen Analyse: *paradigmatic\_analysis.dtd*

`<relation>` ist ein leeres Element und besitzt die Attribute `from`, `to` und `rel-id`. Als Wert der Attribute `from` und `to` wird die Einstellungsnummer und als Wert des Attributes `rel-id` der `id`-Wert der paradigmatischen Relation aus dem Dokument *paradigmatic\_relation.xml* eingetragen. Das Element `<heading>` ist für Zwischenüberschriften vorgesehen, die dann in einem Ausgabedokument (z. B. in der PDF-Übersicht der paradigmatischen Relationen, s. 4.5.4) angezeigt werden können. Ein Ausschnitt der XML-Repräsentation für das Video »Snowboard« zeigt Code 4.8.

<sup>1</sup> In 2.5.2 wird dies durch einen Zeilenumbruch und eine passende Einrückung visualisiert.

```

<?xml version="1.0" encoding="UTF-8"?>
<!-- Paradigmatischen Analyse des Videos Snowboard -->
<!DOCTYPE paradigmaticAnalysis SYSTEM "../dtd/paradigmatic_analysis.dtd">
<paradigmaticAnalysis>
  <info>
    <file>snowboard.xml</file>
  </info>
  <analysis>
    <!--Relation-Elemente für benachbarte Einstellungen-->
    <heading>Benachbarte Einstellungen</heading>
    <relation from="T01" to="T02" rel-id="P1"/>
    <relation from="T02" to="T03" rel-id="P2"/>

    <relation from="T03" to="T04" rel-id="P3"/> <!-- Wechsel: Berg - Werkstatt -->
    <relation from="T04" to="T05" rel-id="P4"/>
    <relation from="T05" to="T06" rel-id="P4"/>
    <relation from="T06" to="T07" rel-id="P5"/>

    ...
  </analysis>
</paradigmaticAnalysis>

```

Code 4.8: Ausschnitt aus *paradigmatic\_analysis.xml*

Dieses XML-Dokument muss jedoch nicht komplett »per Hand« erstellt werden. Zur Unterstützung der Beobachter wird durch das XSLT-Stylesheet *paraRelation\_gen.xsl* (s. Code 4.9) aus dem Transkriptionsprotokoll (hier: *snowboard.xml*) eine leere »Vorlage« (*template*) für die Analyse erzeugt. Diese Vorlage enthält für alle zu analysierenden Einstellungspaare die notwendigen Elemente `<relation>` mit den Attributen `from` und `to`. Es muss lediglich der `id`-Wert der zugehörigen paradigmatischen Relation als Wert des Attributes `rel-id` eingetragen werden.

Das XSLT-Stylesheet arbeitet wie folgt: In der ersten `<xsl:for-each>`-Schleife werden die `<relation>`-Elemente für benachbarte Einstellungen erzeugt. Eine benachbarte Einstellung wird über den XPath-Ausdruck `./following-sibling::unit()[1]` adressiert.

Die zweite Schleife gibt die `<relation>`-Elemente für Übergänge zu/von weiter zurückliegenden (also nicht-benachbarten) Einstellungen des gleichen Ortes, sortiert nach den Orten, aus. Dazu iteriert diese Schleife über die Segmente der ortsübergreifenden Gruppierung, die vorab durch das benannte Template *segment* (s. 4.5.1) berechnet und als Baum (*temporary tree*) in der Variablen `segTree` gespeichert werden. Die nicht-benachbarte Einstellung des gleichen Ortes findet man über den XPath-Ausdruck `following-sibling::segment()[@location=$loc][1]`, d. h. von den nachfolgenden Geschwisterelementen wird das erste Element ausgewählt, dessen Attributwert von `location` identisch zu dem des betrachteten Elementes ist.

Die dritte Schleife erzeugt `<relation>`-Elemente für nicht-benachbarte Einstellungen, bei denen die detaillierten Ortsangaben verschieden sind.<sup>1</sup> Verwendet wird dazu das benannte Template *segment\_loc\_detail* (s. 4.5.1), welches einstellungsübergreifende Ortsgruppen basierend auf den detaillierten Ortsinformationen (gespeichert im Element `<loc_detail>`) berechnet. »Nicht-benachbart« bedeutet, dass sich die Einstellungsnummern um mehr als 1 unterscheiden, daher

<sup>1</sup> Im Snowboard-Video sind dies die Arbeitsstationen in der Werkstatt.

wird der XPath-Ausdruck `number(substring(./following-sibling::segment[1]/@begin,2)) - number(substring(./@begin,2)) gt 1` zu `<xsl:if>` als Bedingung notiert.

```
<xsl:template match="/">
  <xsl:comment>Template zur Erfassung der paradigmatischen Relationen</xsl:comment>
  <paradigmaticAnalysis>
    <info>
      <file><xsl:value-of select="/movie/movie_info/name"/>.xml</file>
    </info>

    <!-- Benachbarte Einstellungen -->
    <xsl:comment>Relation-Elemente für benachbarte Einstellungen</xsl:comment>
    <heading>Benachbarte Einstellungen</heading>
    <xsl:for-each select="movie/units/unit[position()<!=last()]">
      <relation from="{./@id}" to="{./following-sibling::unit[1]/@id}" rel-id=".."/>
    </xsl:for-each>

    <!-- Nichtbenachbarte Einstellungen - gleicher Ort (loc_name) -->
    <heading>Nichtbenachbarte Einstellungen</heading>
    <xsl:comment>Relation-Elemente für nichtbenachbarte Einstellungen</xsl:comment>

    <!-- in der Variable segTree wird der temporäre Baum der Segmentierung gespeichert
    Gruppierung über location/loc_name -->
    <xsl:variable name="segTree">
      <xsl:call-template name="segment"/>
    </xsl:variable>

    <xsl:for-each select="$segTree/segments/segment">
      <xsl:sort select="@location"/>
      <xsl:variable name="loc" select="@location"/>
      <xsl:if test="exists(./following-sibling::segment[@location = $loc][1])">
        <relation from="{./@end}"
          to="{./following-sibling::segment[@location = $loc][1]/@begin}"
          rel-id=".."/>
      </xsl:if>
    </xsl:for-each>

    <!-- Nichtbenachbarte Einstellungen Wechsel bei loc_detail -->
    <xsl:comment>Relation-Elemente für nichtbenachbarte Einstellungen -
    Wechsel bei loc_detail</xsl:comment>
    <heading>Nichtbenachbarte Einstellungen - Wechsel bei loc_detail</heading>

    <!-- in der Variable segTree_loc_detail wird der temporäre Baum der Segmentierung
    gespeichert
    Gruppierung über location/loc_detail, falls vorhanden -->
    <xsl:variable name="segTree_loc_detail">
      <xsl:call-template name="segment_loc_detail"/>
    </xsl:variable>

    <xsl:for-each select="$segTree_loc_detail/segments/segment">
      <xsl:variable name="loc" select="@location"/>
      <xsl:if
        test="exists(./following-sibling::segment[1])
        and <!-- Einstellungsnummern unterscheiden sich um mehr als 1 -->
        (number(substring(./following-sibling::segment[1]/@begin,2))
        -number(substring(./@begin,2)) gt 1)">
        <relation from="{./@begin}"
          to="{./following-sibling::segment[1]/@begin}" rel-id=".."/>
      </xsl:if>
    </xsl:for-each>
  </paradigmaticAnalysis>
</xsl:template>
```

Code 4.9: Ausschnitt aus `paraRelation_gen.xsl`

#### 4.3.4 XML-Repräsentation der Analyse der Text-Bild-Relationen

Die XML-Repräsentation der Taxonomie der Text-Bild-Relationen (s. 2.4) basiert auf der DTD `textImage_relation.dtd` (s. Code 4.10).

```
<!ELEMENT textImageRelations (relation+)>
<!ELEMENT relation (relation*)>

<!-- ATTLIST relation
      id ID #REQUIRED
      name CDATA #REQUIRED -->
```

Code 4.10: DTD zur Repräsentation der Taxonomie der Text-Bild-Relationen: `textImage_relation.dtd`

Die Struktur einer darauf basierenden XML-Instanz gibt die Baumstruktur der Taxonomie wieder, wobei beliebig tiefe Baumstrukturen erfasst werden können. Für den Wert eines Attributes `id` gilt folgendes Benennungsschema: Die oberste Ebene erhält als Wert `rx`, wobei `x` eine zweistellige Zahl ist. Die `id` einer Subrelation der folgenden Ebene wird jeweils um `_yy` ergänzt. Innerhalb einer Ebene wird durchnummeriert. Auf diese Weise ist anhand der `id` immer die Position einer Relation in der Taxonomie erkennbar. Code 4.11 zeigt einen Ausschnitt aus `textImage_relation.xml`. Die mit `»*«` markierten Relationen (im unteren Bereich) sind die im Rahmen der Analyse des Beispielvideos »Snowboard« neu hinzugefügten Relationen (s. 2.5.4).

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- Text-Bild Relationen nach Kong und eigene Erweiterungen aus der Analyse (Kapitel 2) -->
<!DOCTYPE relations SYSTEM "../dtd/textImage_relation.dtd">
<textImageRelations>
  <!-- Relationen nach Kong -->
  <relation id="r01" name="expansion">
    <relation id="r01_01" name="elaboration">
      <relation id="r01_01_01" name="explanation"/>
      <relation id="r01_01_02" name="exemplification"/>
      <relation id="r01_01_03" name="specification"/>
      <relation id="r01_01_04" name="identification"/>
    </relation>
    ...
  </relation>
  ...
  <!-- neu hinzugefügte Relationen -->
  <relation id="r03" name="decoration">
    <relation id="r03_01" name="hypophora*"/>
    <relation id="r03_02" name="comment*"/>
    <relation id="r03_03" name="invitation*"/>
    <relation id="r03_04" name="preparation*"/>
  </relation>
</textImageRelations>
```

Code 4.11: Ausschnitt aus `textImage_relation.xml`

Die Darstellung der Analyse der tatsächlich vorliegenden Text-Bild-Relationen basiert auf der DTD `textImage_analysis.dtd` (s. Code 4.12).

```

<!ELEMENT textImageAnalysis (info,analysis)>
<!ELEMENT info (file)>
<!ELEMENT file (#PCDATA)>
<!ELEMENT analysis (unit+)>
<!ELEMENT unit (relation+)>
<!ELEMENT relation EMPTY>

<!ATTLIST unit id ID #REQUIRED
location NMTOKEN #REQUIRED>
<!ATTLIST relation span-id ID #REQUIRED
rel-id NMTOKEN #REQUIRED
join (next|before) #IMPLIED >

```

Code 4.12: DTD zur Repräsentation der Analyse der Text-Bild-Relationen: *textImage\_analysis.dtd*

Für jede Einstellung wird im Analysedokument ein Element `<unit>` notiert, als Wert seines Attributes `id` wird der `id`-Wert der Einstellung<sup>1</sup> eingetragen. Ein weiteres Attribut von `<unit>` ist `location` zur Speicherung des Handlungsortes.<sup>2</sup> Für jedes Textsegment einer Einstellung wird innerhalb von `<unit>` ein Kindelement `<relation>` notiert. `<relation>` ist ein leeres Element, das Attribut `span-id` dient zur Erfassung des `id`-Wertes des Textsegmentes, das Attribut `rel-id` zur Erfassung des `id`-Wertes der identifizierten Relation aus dem Dokument `textImage_relation.xml`.

Durch das optionale Attribut `join` kann auch (analog zu `<span>` in `units.dtd`) bei einem `<relation>`-Element vermerkt werden, dass hier ein einstellungsübergreifendes Textsegment vorliegt<sup>3</sup>.

Das XSLT-Stylesheet `textImage_relation_gen.xsl` (s. Code 4.13) erzeugt aus einem Transkriptionsprotokoll eine (leere) Vorlage des XML-Dokumentes. In dieses müssen die Beobachter nun lediglich die `id`-Werte der vorliegenden Text-Bild-Relationen aus dem Dokument `textImage_relation.xml` eingetragen.

```

<xsl:template match="/">
  <xsl:comment>Template zur Erfassung der Text-Bild-Relationen</xsl:comment>
  <textImageAnalysis>
    <info>
      <file><xsl:value-of select="/movie/movie_info/name"/>.xml</file>
    </info>
    <analysis>
      <xsl:for-each select="//unit">
        <unit id="{@id}" location="{location/loc_name}">
          <xsl:for-each select="text/span">
            <relation span-id="{@id}" rel-id="..">
              <xsl:if test="@join">
                <xsl:attribute name="join">
                  <xsl:value-of select="@join"/>
                </xsl:attribute>
              </xsl:if>
            </relation>
          </xsl:for-each>
        </unit>
      </xsl:for-each>
    </analysis>
  </textImageAnalysis>

```

<sup>1</sup> Genauer gesagt der Wert des Attributes `id` des Elementes `<unit>` in der XML-Repräsentation des Transkriptionsprotokolls.

<sup>2</sup> Auch für die Erfassung der Text-Bild-Relationen wird mit Hilfe eines XSLT-Stylesheets ein Template generiert. In diesem sind die Handlungsorte bereits eingetragen, so dass sie nicht erneut erfasst werden müssen. Die Speicherung an dieser Stelle erleichtert das XSLT-Stylesheet zur Auswertung der Text-Bild-Relationen (`textImage_report.xsl`, s. 4.5.5) erheblich.

<sup>3</sup> Das Attribut `join` und sein Attributwert werden ebenfalls vom XSLT-Stylesheet in das Template eingetragen.

```

        </xsl:for-each>
    </unit>
</xsl:for-each>

</analysis>
</textImageAnalysis>
</xsl:template>

```

Code 4.13: Ausschnitt aus `textImage_relation_gen.xsl`

Einen Ausschnitt aus der XML-Repräsentation der Text-Bild-Analyse des Videos »Snowboard« (`textImage_Analysis.xml`) zeigt Code 4.14.

```

<?xml version="1.0" encoding="UTF-8"?>
<!-- Analyse der Text-Bild-Relationen des Videos Snowboard -->
<!DOCTYPE textImageAnalysis SYSTEM "../dtd/textImage_analysis.dtd">
<textImageAnalysis>
  <info>
    <file>snowboard.xml</file>
  </info>
  <analysis>
    <unit id="T01" location="Berg">
      <relation span-id="T01_01" rel-id="r01_02_01"/>
      <relation span-id="T01_02" rel-id="r01_01_03"/>
      <relation span-id="T01_03" rel-id="r01_03_10"/>
      <relation span-id="T01_04" rel-id="r03_01"/>
      <relation span-id="T01_05" rel-id="r01_01_01" join="next"/>
    </unit>

    <unit id="T02" location="Berg">
      <relation span-id="T02_01" rel-id="r01_01_01" join="before"/>
      <relation span-id="T02_02" rel-id="r03_01"/>
    </unit>
    ...
  </analysis>
</textImageAnalysis>

```

Code 4.14: Ausschnitt aus `textImage_analysis.xml`

## 4.4 Verwendete Layoutsprachen

Für Bewegbilddaten gibt es eine bis heute nicht abgeschlossene Geschichte der Layoutsprachen. Als deklarative Layoutsprache wurde über viele Jahre vom W3C SMIL entwickelt. Leider hat sich diese Sprache nicht durchgesetzt. Innerhalb der HTML Welt gibt es erst mit HTML5 eine dedizierte Möglichkeit, Bewegbilddaten und Audiodaten in eine Layoutumgebung einzubetten. Allerdings erfordert die Videosteuerung und Interaktion in vielen Fällen zusätzliche JavaScript Programmierung.

### 4.4.1 SMIL

Die *Synchronized Multimedia Integration Language* (SMIL) ist eine XML-Anwendung zur Spezifikation von räumlichen und zeitlichen Beziehungen zwischen Medienobjekten (Text, Bild, Audio, Video, Animation, Textstrom<sup>1</sup>) und Interaktionen zwischen ihnen. Die Medienobjekte werden nicht in das SMIL-Dokument eingebettet, sondern referenziert und erst bei Bedarf geladen. Somit realisiert SMIL eine Trennung zwischen Struktur, Inhalt und Verhalten. Im Juni 1998 wurde die Version 1.0 als W3C-Recommendation veröffentlicht. Die aktuellste Version ist die Version 3.0, die seit Dezember 2008 als W3C-Recommendation vorliegt<sup>2</sup>.

Problematisch bei der Arbeit mit SMIL ist zum einen das Fehlen von Programmen, mit denen Autoren Layout und Ablauf einer Präsentation über einen visuellen Editor, wie dies z. B. mit Adobe Flash möglich ist, gestalten können; zum anderen fehlen auch geeignete Player. Der einzige Player, der SMIL 3.0 unterstützt, ist das Open-Source-Programm »Ambulant Player«, das seit Februar 2015 in der Version 2.6 vorliegt<sup>3</sup>.

#### Aufbau eines SMIL-Dokumentes

Das Wurzelement jedes SMIL-Dokumentes ist `<smil>`. Direkte Kindelemente von `<smil>` sind `<head>` und `<body>`.

Im `<head>` können Metadaten (z. B. Name, Autor, Datum) mit Hilfe eines oder mehrerer Elemente `<meta>` notiert werden. Es folgt das Element `<layout>`. Innerhalb dieses Elementes wird das Layoutschema für die gesamte Präsentation festgelegt. Das leere Element `<root-layout>` dient zur Definition des Hauptfensters (Höhe, Breite, Hintergrundfarbe usw.). Dann folgen ein oder mehrere `<region>`-Elemente, mit denen die einzelnen Bereiche (Regionen) innerhalb des Hauptfensters festgelegt werden. Jede Region benötigt zur eindeutigen Referenzierung ein Attribut `xml:id`. Die Position und Größe einer Region wird durch die Attribute `top`, `left` bzw. `bottom`, `right` sowie

<sup>1</sup> Textdokument, das Zeitinformationen enthält, so dass ein zeitabhängiges Rendering von Textabschnitten möglich ist.

<sup>2</sup> <http://www.w3.org/TR/SMIL3/> [letzter Aufruf: 04.03.2015]

<sup>3</sup> <http://www.ambulantplayer.org/> [letzter Aufruf: 04.03.2015]

`width`, `height` angegeben. Zudem gibt es weitere Attribute, z. B. `backgroundColor`, zur Formatierung. Weiterhin können mit Hilfe des Elementes `<transition>` auch Überblendungsklassen definiert werden, die dann von den Medienobjekten referenziert werden (s. u.).

Im `<body>`-Bereich werden die Medien referenziert und ihr Ablauf gesteuert. Zur Referenzierung eines Medienobjektes definiert SMIL das generische Element `<ref src="URI">`. Alternativ zum `<ref>`-Element können auch die folgenden »sprechenden« Elemente genutzt werden:

- Text: `<text src="text.txt">`
- Bild: ``
- Video: `<video src="video.mp4">`
- Audio: `<audio src="audio.wav">`
- Animation: `<animation src="anim.swf">`
- Textstrom: `<textstream src="text.txt">`

Per Default werden die Medien im Hauptfenster platziert. Die Platzierung in einer bestimmten Region ist durch das Attribut `region` möglich. Als Attributwert wird der Wert der `xml:id` der gewünschten Region angegeben.

### **Zeitliche Steuerung und Synchronisation**

Die Steuerung des Präsentationsbeginns, der Präsentationsdauer und des Präsentationsendes eines Medienobjektes ist durch Angabe der Attribute `begin`, `dur` und `end` möglich. Soll nur ein Ausschnitt einer Video- oder Audiodatei abgespielt werden, werden die Attribute `clipBegin` bzw. `clipEnd` dem Element `<video>` bzw. `<audio>` hinzugefügt.

Zur zeitlichen Synchronisation der Medienanzeige stellt SMIL sogenannte Zeit-Container zur Verfügung:

- `<par>`: Zeitliche Parallelität: Die innerhalb des Containers referenzierten Medien werden parallel angezeigt.
- `<seq>`: Zeitliche Sequenz: Die innerhalb des Containers referenzierten Medien werden nacheinander angezeigt.
- `<excl>`: Zeitliche Exklusivität: Die innerhalb des Containers referenzierten Medien werden unabhängig von der notierten Reihenfolge zu unterschiedlichen Zeiten angezeigt.

Neben der Angabe absoluter Werte für die Attribute `begin` und `end` ist auch eine sogenannte ereignisbasierte Synchronisation möglich, d. h. der Zeitpunkt ist vom Eintreten eines bestimmten Ereignisses abhängig. Mögliche Attributwerte sind z. B. `id.beginEvent`, `id.endEvent`, `id.activateEvent`, also der Zeitpunkt, bei dem das Objekt mit der notierten `id` beginnt bzw. endet oder z. B. durch einen Mausklick aktiviert wird.

### Inline-Text

Das Element `<text>` kann nur auf eine externe Ressource verweisen. Um auch renderbaren Textinhalt direkt in ein SMIL-Dokument einzubetten, wurde in der Version 3.0 das Element `<smilText>` definiert [Bulterman, Rutledge 2008, S. 182]. Innerhalb von `<smilText>` können u. a. die Elemente `<div>`, `<p>`, `<span>` zur Textsegmentierung und die Attribute `textColor`, `textFontStyle` zur Formatierung verwendet werden. Zudem bietet SMIL auch die Möglichkeit, Text zeit- oder eventabhängig darzustellen (`<tev>`-Element) sowie Textbewegungen, z. B. scrollen, zu realisieren.

### Links

Wie in HTML kann durch `<a href="ziel">objekt</a>` jedes Objekt Quellanker eines Links werden. Mehr Möglichkeiten bietet jedoch das Element `<area>`, das als Kindelement jedem Medienobjekt hinzugefügt werden kann. Im Gegensatz zu `<a>` definiert `<area>` einen Anker, der nur einen bestimmten Bereich des Medienobjektes umfasst (»Hotspot«). Dieser Bereich kann räumlich (spatial, wie in HTML) oder zeitlich (temporal) sein. Ein temporaler Anker ist ein Anker, der nur für eine bestimmte Zeitspanne der Objektlaufzeit aktiv ist. Zur Angabe räumlicher Bereiche dienen die Attribute `shape` und `coords`. Mit `shape` wird die Form des Bereichs angegeben, erlaubte Werte sind `rect`, `circle` und `poly`. Der Wert von `coords` gibt die Position und Größe des Bereichs an. Zeitliche Bereiche werden durch die Attribute `begin`, `end`, `dur` festgelegt.

### Überblendungen

Zwischen den Präsentationen einzelner Medienobjekte können Übergänge, wie z. B. ein Auf- oder Abblenden, angezeigt werden. Dazu wird im `<head>`-Bereich durch das Element `<transition>` eine Überblendungsklasse definiert. Notwendig ist hier das Attribut `xml:id` zur eindeutigen Identifikation. Der Wert des Attributes `type` gibt die Effektart an. SMIL bietet eine große Auswahl an Effekten an, z. B. `fade` für ein einfaches Einblenden oder auch Effekte wie z. B. `slideWipe`, `spiralWipe`, `windShieldWipe`. Mit weiteren Attributen, wie `begin`, `end`, `dur`, wird die genaue Ausführung spezifiziert. Um einem Medienobjekt einen Überblendungseffekt hinzuzufügen, erhält es das Attribut `transIn`, wenn der Effekt beim Präsentationsbeginn, bzw. `transOut`, wenn er zum Präsentationsende ausgeführt werden soll. Wert des notierten Attributes ist die `xml:id` der Überblendungsklasse.

#### 4.4.2 Videointegration in HTML5

Zur Integration von Videos in Webseiten muss bis zur HTML Version 4 ein PlugIn verwendet werden. Am weitesten verbreitet ist das proprietäre Flash-PlugIn. Aufgrund der Sicherheitslücken und dem großen Speicherverbrauch wurde dies jedoch häufig kritisiert. Diverse mobile Endgeräte, z. B. iPad und iPhone, unterstützen es aus diesen Gründen nicht.

HTML5 sieht die native Unterstützung von Multimedia vor. Zu diesem Zweck wurden spezifische Elemente zur Einbindung von Audio- und Videodateien, die Elemente `<audio>` und `<video>`, definiert, die von neueren Browserversionen des Internet Explorers, Firefox, Opera, Chrome und Safari unterstützt werden. Zur Integration eines Videos wird das Element `<video>` in den HTML-Quelltext eingefügt. Dessen Kindelement `<source>` besitzt die Attribute `src` und `type` zur Angabe der Videoquelle und des Videoformates. Zur Zeit werden die drei Videoformate MP4, WebM<sup>1</sup> und Ogg Theora<sup>2</sup> unterstützt. Da jedoch keines der Videoformate von allen Browsern unterstützt wird<sup>3</sup>, sollte das Video in den verschiedenen Formaten bereitgestellt werden. Im Quellcode sind hierzu alternative `<source>`-Elemente mit dem entsprechenden `type`-Attribut anzugeben (s. Code 4.15). Der Browser lädt dann das Format, das er anzeigen kann.

```
<video id="video1" width="420">
  <source src="snowboard.mp4" type="video/mp4" />
  <source src="snowboard.ogg" type="video/ogg" />
  <source src="snowboard.webm" type="video/webm" />
  Your browser does not support HTML5video.
</video>
```

Code 4.15: Integration verschiedener Videoformate in HTML5

Zur Steuerung der Wiedergabe eines Videos können folgende Attribute einem `<video>`-Element hinzugefügt werden:

- `autoplay`: spielt das Video automatisch ab, sobald es geladen ist,
- `controls`: zeigt vom Browser vorgegebene Bedienelemente (play, pause, Lautstärke),
- `muted`: das Video wird ohne Ton abgespielt,
- `loop`: das Video wird endlos wiederholt,
- `poster`: statt des ersten Video-Frames wird ein benutzerdefiniertes Vorschaubild angezeigt.

<sup>1</sup> Offenes Videoformat, speziell für das Internet entwickelt, <http://www.webmproject.org/> [letzter Aufruf: 04.03.2015]

<sup>2</sup> Offenes Videokompressionsformat der Xiph.org Foundation, <http://www.xiph.org/ogg/> [letzter Aufruf: 04.03.2015]

<sup>3</sup> FireFox: WebM, Ogg; Opera: WebM, Ogg; Chrome: WebM, MP4; Safari: MP4; Internet Explorer: MP4

Da das Element `<video>` ein natives HTML-Element ist, gibt es im DOM-Modell<sup>1</sup> (*Document Object Model*) JavaScript-Eigenschaften und -Methoden zur Kontrolle darüber, wie ein Video abgespielt werden soll.<sup>2</sup> Die wichtigsten sind:

- `currentTime`: die aktuelle Wiedergabeposition des Videos (in sec),
- `volume`: die Lautstärke des Videos,
- `play()`: das Video wird ab der aktuellen Wiedergabeposition gestartet,
- `pause()`: das Video pausiert.

### Grenzen

Es existiert keine `stop()`-Methode. Wenn das Video gestoppt und »zurückgespult« werden soll, muss man die Methode `pause()` verwenden und die Wiedergabeposition `currentTime` auf 0 setzen. Weiterhin existiert keine Methode, mit der man angeben kann, dass ein bestimmtes Teilsegment eines Videos abgespielt wird. Fehlende Funktionen müssen daher mit JavaScript programmiert werden. Entwickelt wurden die folgenden Java-Script Funktionen (s. auch 6.3.2):

- `videoPlayLoop(id,begin,dur,loop)`  
Das Video wird ab einem bestimmten Zeitpunkt und für eine bestimmte Dauer abgespielt. Hat der Parameter `loop` den Wert `true`, wird das Video in einer Endlosschleife abgespielt.
- `videoPlayStop(id1,id2,begin,dur)`  
Das Video mit der Id `id1` wird ab einem bestimmten Zeitpunkt für eine bestimmte Dauer abgespielt. Ist die Dauer erreicht, wird das Video mit der Id `id2` gestoppt.
- `videoSeq(x)`  
Eine Folge von Videosegmenten wird abgespielt. Übergabeparameter ist ein Array, das Instanzen des (selbstdefinierten) JavaScript-Objektyps `video` enthält. In einer solchen Instanz werden Informationen über ein Videosegment (Beginn, Dauer, mit oder ohne Ton, beschreibender Text) gespeichert.
- `videoDelay(id,delay,begin,dur)`  
Das Video mit der Id `id` wird nach der Verzögerungszeit `delay` ab einem bestimmten Zeitpunkt für eine bestimmte Dauer abgespielt.

---

<sup>1</sup> Spezifikation einer Schnittstelle für den Zugriff auf ein HTML- oder XML-Dokument

<sup>2</sup> siehe: <http://www.w3.org/TR/html5/embedded-content-0.html#media-elements> [letzter Aufruf: 04.03.2015]

## 4.5 Analytische Hilfsmittel

Dieser Abschnitt fasst »Hilfsmittel« zusammen, die die Programmierung bzw. die verschiedenen Analysen unterstützen. Die ersten Hilfsmittel (s. 4.5.1) dienen der Modularisierung und Wiederverwendung von Programmcode. Da die Berechnungen einstellungsübergreifender Ortsgruppen mehrmals benötigt werden, werden sie in benannte Templates ausgelagert, welche dann bei Bedarf inkludiert werden. Ein wertvolles Hilfsmittel für die Analysen ist die Möglichkeit, jede Einstellung direkt anzusteuern und abzuspielen (s. 4.5.2). Zur paradigmatischen Analyse müssen je zwei Einstellungen betrachtet werden. Ein XSLT-Stylesheet erzeugt ein PDF-Dokument, ein weiteres HTML5- und SMIL-Dokumente, in welchen je zwei Screenshots bzw. zwei Einstellungen in direkter Abfolge zusammen mit den bereits vorliegenden Analyseergebnissen betrachtet werden können (s. 4.5.4). Zwei XSLT-Stylesheets (s. 4.5.3 bzw. 4.5.5) erzeugen Übersichten, die auch schon in den Analysen des Kapitels 2 verwendet wurden: Dies sind die Wechseltabelle (Tab. 2.3) und die Auswertung der Text-Bild-Relationen (Abb. 2.34).

### 4.5.1 Berechnung einstellungsübergreifender Ortsgruppen

Das XSLT-Template `segment` (s. Code 4.16) fasst die aufeinanderfolgenden Einstellungen eines diegetischen Ortes zu einer Gruppe zusammen und stellt Informationen über diese Gruppe zur Verfügung, die für weitere Publikationsprozesse genutzt werden können. Im Template wird die Instruktion `<xsl:for-each-group>` verwendet, die die im `select` Attribut angegebene Eingabesequenz anhand eines Gruppierungskriteriums sortiert. Das Gruppierungskriterium wird als Wert des Attributes `group-adjacent` notiert, so dass nur aufeinanderfolgende gleiche Werte einer Gruppe zugeordnet werden.<sup>1</sup> Hier ist der Attributwert `location/loc_name`, daher werden alle aufeinanderfolgenden Einstellungen mit gleichem Ortsnamen zu einer Gruppe zusammengefasst. Mit der XPath-Funktion `current-grouping-key()` wird das aktuell gültige Gruppierungskriterium ausgegeben. Die Funktion `current-group()` gibt die Sequenz der Items der aktuell verarbeitenden Gruppe zurück. In diesem Fall werden Informationen des ersten und letzten Items gespeichert. Hierbei wird für jedes Segment notiert, mit welcher Einstellung und zu welchem Zeitpunkt das Segment beginnt und endet.

```
<!-- Gruppierung über location/loc_name -->
<xsl:template name="segment">
  <segments>
    <xsl:for-each-group select="/movie/units/unit" group-adjacent="location/loc_name">
      <segment begin="{current-group()[1]/@id}"
              end="{current-group()[last()]/@id}"
              location="{current-grouping-key()}"
              clipBegin="{current-group()[1]/begin}"
              clipEnd="{current-group()[last()]/end}"/>
    </xsl:for-each-group>
  </segments>
</template>
```

<sup>1</sup> Wird das Gruppierungskriterium als Wert des Attributes `group-by` notiert, werden alle Items der Eingabesequenz, die den gleichen Wert haben, einer Gruppe zugeordnet.

```
</segments>
</xsl:template>
```

Code 4.16: Template segment zur Berechnung der Segmente der verschiedenen Orte

Für das Video »Snowboard« ergibt sich die in Code 4.17 gezeigte XML-Darstellung.<sup>1</sup>

```
<segments>
  <segment begin="T01" end="T03" location="Berg" clipBegin="0:00" clipEnd="0:29.6"/>
  <segment begin="T04" end="T10" location="Werkstatt" clipBegin="0:30" clipEnd="1:57"/>
  <segment begin="T11" end="T11" location="Berg" clipBegin="1:58" clipEnd="2:18.6"/>
  <segment begin="T12" end="T15" location="Werkstatt" clipBegin="2:19.3" clipEnd="3:05"/>
  <segment begin="T16" end="T16" location="Berg" clipBegin="3:06" clipEnd="3:17.5"/>
  <segment begin="T17" end="T35" location="Werkstatt" clipBegin="3:18" clipEnd="6:14.2"/>
  <segment begin="T36" end="T36" location="Berg" clipBegin="6:15" clipEnd="6:26"/>
  <segment begin="T37" end="T40" location="Werkstatt" clipBegin="6:27" clipEnd="6:53"/>
  <segment begin="T41" end="T42" location="Berg" clipBegin="6:54" clipEnd="7:39"/>
</segments>
```

Code 4.17: Segmente für das Video »Snowboard« berechnet mit dem Template segment

Das benannte Template segment ist in einer eigenen Datei (segment.xml) gespeichert und kann so bei Bedarf in ein XSLT-Stylesheet inkludiert werden. Verwendet wird es in dieser Arbeit zur Generierung des Inhaltsverzeichnisses der Bildergeschichten (s. 4.6), den Modifikationen der Alternanz unter Erhalt der Kohärenzrelation (s. 4.7) sowie dem Wegfall der Kohärenzrelation (s. 4.9). In den inkludierenden Stylesheets wird das Template innerhalb einer Variablendefinition (hier: segTree) aufgerufen (s. Code 4.18). In dieser Variablen wird dann das Resultat des Templates als temporärer Baum (*temporary tree*) gespeichert.

```
<!-- in der Variable segTree wird der temporäre Baum der Segmentierung gespeichert -->
<xsl:variable name="segTree">
  <xsl:call-template name="segment"/>
</xsl:variable>
```

Code 4.18: Aufruf des Templates segment

In gleicher Weise arbeitet das XSLT-Template segment\_loc\_detail (s. Code 4.19, gespeichert in der Datei segment\_loc\_detail.xml). Im Unterschied zum Template segment wird hier die Gruppierung über die detaillierten Ortsangaben, gespeichert im Element <loc\_detail>, durchgeführt.<sup>2</sup> Die <xsl:for-each-group>-Schleife iteriert nur über die <unit>-Elemente, für die eine solche Detailangabe vorliegt. Gewährleistet wird dies durch den Existenztest im Prädikat des XPath-Ausdrucks /movie/units/unit[location/loc\_detail].

```
<!-- Gruppierung nach location/loc_detail -->
<xsl:template name="segment_loc_detail">
  <segments>
    <xsl:for-each-group select="/movie/units/unit[location/loc_detail]"
                      group-adjacent="location/loc_detail">
      <segment begin="{current-group()[1]/@id}"
              end="{current-group()[last()]/@id}"
              location="{current-grouping-key()}">
```

<sup>1</sup> Dies entspricht der Darstellung von Tab. 2.5 in Kapitel 2.5.3, ergänzt durch die Zeitangaben.

<sup>2</sup> Im Video »Snowboard« sind dies die einzelnen Arbeitsstationen in der Werkstatt.

```

        clipBegin="{current-group()[1]/begin}"
        clipEnd="{current-group()[last()]/end}"/>
    </xsl:for-each-group>
</segments>
</xsl:template>

```

Code 4.19: Template `segment_loc_detail` zur Berechnung der Segmente an verschiedenen Orten

Verwendet wird das Template `segment_loc_detail` im XSLT-Stylesheet zur Generierung des Templates zur Erfassung der paradigmatischen Relationen `paraRelation_gen.xml` (s. 4.3.3). Es dient dort zur Berechnung der nicht-benachbarten Einstellungen, bei denen die detaillierten Ortsangaben verschieden sind.<sup>1</sup>

```

<segments>
<segment begin="T04" end="T08" location="S1" clipBegin="0:30" clipEnd="1:26.9"/>
<segment begin="T09" end="T10" location="S2" clipBegin="1:27.1" clipEnd="1:57"/>
<segment begin="T12" end="T13" location="S3" clipBegin="2:19.3" clipEnd="2:41.6"/>
<segment begin="T14" end="T15" location="S4" clipBegin="2:42" clipEnd="3:05"/>
<segment begin="T17" end="T17" location="S5" clipBegin="3:18" clipEnd="3:31"/>
<segment begin="T18" end="T18" location="S6" clipBegin="3:32" clipEnd="3:38"/>
<segment begin="T19" end="T19" location="S7" clipBegin="3:39" clipEnd="3:54"/>
<segment begin="T20" end="T20" location="S8" clipBegin="3:55" clipEnd="4:07.4"/>
<segment begin="T21" end="T31" location="S9" clipBegin="4:08" clipEnd="5:53"/>
<segment begin="T32" end="T35" location="S10" clipBegin="5:54" clipEnd="6:14.2"/>
<segment begin="T37" end="T37" location="S11" clipBegin="6:27" clipEnd="6:32"/>
<segment begin="T38" end="T38" location="S12" clipBegin="6:33" clipEnd="6:38.3"/>
<segment begin="T39" end="T39" location="S13" clipBegin="6:39" clipEnd="6:47.9"/>
<segment begin="T40" end="T40" location="S14" clipBegin="6:48" clipEnd="6:53"/>
</segments>

```

Code 4.20: Segmente für das Video »Snowboard« berechnet mit dem Template `segment_loc_detail`

<sup>1</sup> Im Snowboard-Video entspricht dies dem Wechsel von einer Arbeitsstation zur nächsten.



```

        <xsl:value-of select="@id"></xsl:value-of>
      </xsl:attribute>
      <xsl:attribute name="class">black</xsl:attribute>
      <xsl:attribute name="onClick">
        toggleClass('<xsl:value-of select="@id"/>', 'black', 'blue');
        videoPlayLoop('video1', <xsl:value-of select="fu:inSec(begin)"/>,
          <xsl:value-of select="fu:dauer(begin,end)*1000"/>, false)
      </xsl:attribute>
      <xsl:value-of select="@id"/>
    </button> &#160;
  </xsl:for-each>
</td>
</tr>
</xsl:for-each-group>

```

Code 4.21: Ausschnitt aus show\_units.xsl

### 4.5.3 Erzeugung der Wechseltabelle

Ausgehend von der XML-Repräsentation eines Transkriptionsprotokolls (hier: snowboard.xml) wird die Wechseltabelle (s. Tab. 2.3) mit Hilfe des XSLT-Stylesheet transition\_table.xsl erzeugt. Hierzu wird zuerst ein FO-Dokument<sup>1</sup> erzeugt, das ein FO-Renderer (hier: Apache FOP<sup>2</sup>) anschließend in das PDF-Format umwandelt. Das XSLT-Stylesheet (s. Code 4.22) arbeitet wie folgt: Als Tabellenkopf werden zwei Zeilen ausgegeben. Die erste Zeile enthält die verschiedenen (Haupt-)Orte (gespeichert im Element <loc\_name>), die zweite Zeile die Detaillierungen der Ortsangaben (Element <loc\_detail>), falls vorhanden, wenn nicht ein »-«. Zur Ausgabe der ersten Tabellenkopfzeile muss daher berechnet werden, wie viele verschiedene Detaillierungen es für den jeweiligen Ort gibt (Variable *anz*), damit die Anzahl der Spalten, die eine Kopfzeile überspannen soll, als Attributwert vom *number-columns-spanned* gesetzt werden kann.<sup>3</sup>

Für jedes Element <unit> wird durch <xsl:for-each select="unit"> eine Tabellenzeile erzeugt. Innerhalb dieser Schleife werden durch <xsl:for-each select="\$dist\_loc\_detail"> die verschiedenen Orte durchlaufen. Stimmt der Inhalt von *location* mit dem betrachteten Ort überein (XPath-Ausdruck: *index-of(\$dist\_loc\_detail,\$loc)=position()*), wird die Einstellungsnummer in die Tabellenzelle eingetragen, falls nicht ein Leerzeichen.

```

<!-- die verschiedenen Orte - im Element loc_name -->
<xsl:variable name="dist_loc" select="distinct-values(//location/loc_name)"/>

<!-- die verschiedenen Orte - loc_detail, wenn nicht vorhanden loc_name -->
<xsl:variable name="dist_loc_detail"
  select="distinct-values(//location/loc_detail union
    //location[not(loc_detail)]/loc_name)"/>
...
<xsl:template match="units">
  <fo:table border-style="solid">
    <fo:table-body>
      <!-- Tabellenkopf 1. Zeile-->

```

<sup>1</sup> FO = *Formatting Objects*

<sup>2</sup> <http://xmlgraphics.apache.org/fop/> [letzter Aufruf: 04.03.2015]

<sup>3</sup> Sind keine Detaillierungen vorhanden, wird der Wert der Variable auf 1 gesetzt.

```

<fo:table-row>
  <xsl:for-each select="$dist_loc">
    <xsl:variable name="kn" select="."/>
    <xsl:variable name="loc_detail"
      select="distinct-values(
        $root/movie/units/unit/location[loc_name=$kn]/loc_detail)"/>

    <!-- Auch wenn keine Detaillierungen vorhanden, muss eine Zelle erzeugt werden.
      In diesem Fall muss number-columns-spanned=1 sein. -->
    <xsl:variable name="anz"
      select="if (count($loc_detail)=0) then 1 else count($loc_detail)"/>

    <fo:table-cell xsl:use-attribute-sets="tableCellHead"
      number-columns-spanned="{ $anz}" text-align="center">
      <fo:block> <xsl:value-of select="."/> /fo:block>
    </fo:table-cell>
  </xsl:for-each>
</fo:table-row>

<!-- Tabellenkopf 2. Zeile -->
<fo:table-row>
  <xsl:for-each select="$dist_loc">
    <xsl:variable name="kn" select="."/>
    <xsl:variable name="loc_detail"
      select="distinct-values(
        $root/movie/units/unit/location[loc_name=$kn]/loc_detail)"/>

    <xsl:choose>
      <xsl:when test="count($loc_detail)=0">
        <fo:table-cell xsl:use-attribute-sets="tableCellHead">
          <fo:block> - </fo:block>
        </fo:table-cell>
      </xsl:when>
      <xsl:otherwise>
        <xsl:for-each select="$loc_detail">
          <fo:table-cell xsl:use-attribute-sets="tableCellHead">
            <fo:block> <xsl:value-of select="."/> </fo:block>
          </fo:table-cell>
        </xsl:for-each>
      </xsl:otherwise>
    </xsl:choose>

  </xsl:for-each>
</fo:table-row>

<!-- Ausgabe Tabellenzeilen -->
<xsl:for-each select="unit">

  <!-- loc_detail, wenn nicht vorhanden loc_name, der aktuellen unit-->
  <xsl:variable name="loc"
    select="if (exists(location/loc_detail))
      then location/loc_detail
      else location/loc_name"/>

  <xsl:variable name="uid" select="@id"/>

  <fo:table-row>
    <xsl:for-each select="$dist_loc_detail"> <!-- die verschiedenen Orte -->
      <fo:table-cell xsl:use-attribute-sets="tableCell">
        <fo:block>
          <xsl:choose>
            <!-- stimmt der Inhalt von location mit dem betrachteten Ort überein -->
            <xsl:when test="index-of($dist_loc_detail,$loc)=position()">
              <xsl:value-of select="$uid"/>
            </xsl:when>
          </xsl:choose>
        </fo:block>
      </fo:table-cell>
    </xsl:for-each>
  </fo:table-row>
</xsl:for-each>

```

```

        <xsl:otherwise>#160;</xsl:otherwise>
    </xsl:choose>
</fo:block>
</fo:table-cell>
</xsl:for-each>
</fo:table-row>

</xsl:for-each>

</fo:table-body>
</fo:table>

</xsl:template>

```

Code 4.22: Ausschnitt aus *transition\_table.xsl*

#### 4.5.4 Unterstützung der Analyse der paradigmatischen Relationen

Zur paradigmatischen Analyse müssen die Beziehungen von genau zwei Einstellungen zueinander klassifiziert werden. Dazu ist es hilfreich, wenn nur die zu analysierenden Einstellungen isoliert betrachtet werden können. XSLT-Stylesheets erzeugen ein PDF-Dokument, SMIL- und HTML5-Dokumente.

##### Technische Umsetzung des PDF-Dokumentes

Das XSLT-Stylesheet *paraRelation\_show\_fo.xsl* erzeugt ein PDF-Dokument (s. Abb. 4.8). In diesem werden die zu analysierenden Einstellungspaare jeweils in einer Zeile durch das entsprechende Bild aus dem Transkriptionsprotokoll dargestellt. Da das erzeugte Dokument auch zum Korrekturlesen während bzw. nach der Analyse dienen soll, wird für jede bereits erfasste *rel-id* die ausführliche textuelle Darstellung der zugeordneten paradigmatischen Relation in der rechten Spalte angezeigt. Eingabedokument der Transformation ist das Dokument *paradigmatic\_analysis.xml*, in dem alle zu analysierenden Übergänge durch die Elemente `<relation from="Txx" to="Tyy" rel-id =".."/>` eingetragen sind (s. Code 4.8). Das Stylesheet enthält daher im Wesentlichen eine `<xsl:for-each>`-Schleife, die über die `<relation>`-Elemente iteriert. Zusätzlich wird das Dokument *paradigmatic\_relation.xml* eingelesen, das die textuelle Darstellung der paradigmatischen Relationen enthält, sowie das Transkriptionsprotokoll (hier: *snowboard.xml*), um Einstellungsdetails, wie Orts- und Zeitangaben, darstellen zu können.

Übersicht der paradigmatischen Relationen im Video "Snowboard"		
Benachbarte Einstellungen		
<b>T01 Ort: Berg (0:00 - 0:18)</b> 	<b>T02 Ort: Berg (0:18 - 0:22.5)</b> 	[P1] [non-projecting; hypotactic - extending; diegetic - event ellipsis - forward   measurable contiguous - inclusion - narrowing specific]
<b>T02 Ort: Berg (0:18 - 0:22.5)</b> 	<b>T03 Ort: Berg (0:23 - 0:29.6)</b> 	[P2] [non-projecting; hypotactic - extending; diegetic - event continuous contiguous - inclusion - broadening specific]
<b>T03 Ort: Berg (0:23 - 0:29.6)</b> 	<b>T04 Ort: Werkstatt /S1 (0:30 - 0:35.5)</b> 	[P3] [non-projecting; paratactic - external   contrast; diegetic - event ellipsis - forward   indefinite non-contiguous - distant specific]
<b>T04 Ort: Werkstatt /S1 (0:30 - 0:35.5)</b> 	<b>T05 Ort: Werkstatt /S1 (0:36 - 0:42.5)</b> 	[P4] [non-projecting; hypotactic - extending; diegetic - event continuous contiguous - inclusion - narrowing specific]

Abb. 4.8: Übersicht: Paradigmatische Relationen – Darstellung im PDF-Dokument

### Technische Umsetzung in HTML5 und SMIL

In der HTML- und SMIL-Lösung werden die entsprechenden Einstellungen direkt aufeinanderfolgend (als Sequenz) abgespielt. Erzeugt werden die Dokumente durch das XSLT-Stylesheet `paraRelation_show.xsl`. Für jedes Einstellungspaar wird eine HTML- und eine SMIL-Datei erzeugt, so dass der Zugriff auf ein bestimmtes Einstellungspaar direkt über das Dateisystem möglich ist. Die erzeugten Dateien werden daher passend zu ihrem »Inhalt« benannt: `para_Txx_Tyy.htm` bzw. `para_Txx_Tyy.smil`. Analog zur PDF-Generierung iteriert eine `<xsl:for-each>`-Schleife über die `<relation>`-Elemente. Zur Erzeugung der verschiedenen Dateien wird die XSLT-Instruktion `<xsl:result-document href="URL">` verwendet.

In der HTML-Version werden zur Generierung der Videosequenz die Informationen über die beiden Einstellungen im Array `source` gespeichert (s. Code 4.23, oben). Dazu enthält jedes Element des Arrays eine Instanz des (selbstdefinierten) JavaScript-Objektyps `Video` mit den Eigenschaften `id`, `begin`, `dauer`, `sound` und `txt`. Beim Laden des HTML-Dokumentes wird die Funktion `videoSeq(source)` aufgerufen, die die Einstellungsfolge abgespielt. Das `<video>`-Element besitzt das Attribut `poster`, um vor dem Start des Videos das zur Einstellung gehörende Bild des Transkriptionsprotokolls als Vorschaubild zu zeigen.

In SMIL werden die als Sequenz abzuspielenden Medien innerhalb des Zeitcontainers `<seq>` angeordnet (s. Code 4.23, unten). In diesem Fall befinden sich innerhalb von `<seq>` zwei `<par>`-Container, in denen jeweils das Video und das Vorschaubild referenziert werden.

```

<!-- Videosequenz in HTML -->
<head>
...
<script>
var source = new Array();
source[source.length]=new Video("video1",
    <xsl:value-of select="fu:inSec($unit1/begin)"/>,
    <xsl:value-of select="fu:dauer($unit1/begin,$unit1/end)*1000"/>,1,'');
source[source.length]=new Video("video2",
    <xsl:value-of select="fu:inSec($unit2/begin)"/>,
    <xsl:value-of select="fu:dauer($unit2/begin,$unit2/end)*1000"/>,1,'');
</script>
</head>
<body onload="videoSeq(source);">
<h1>Paradigmatische Relationen</h1>
...
<tr>
<td>
<video id="video1" poster="../smil/img/snowboard_g/{fu:bname($u1Id)}">
<source src="{ $path }{$vname1}" type="video/mp4"/>
...
</video>
</td>
<td>
<video id="video2" poster="../smil/img/snowboard_g/{fu:bname($u2Id)}">
<source src="{ $path }{$vname1}" type="video/mp4"/>
...
</td>
</tr>
...
</body>

<!-- Videosequenz in SMIL -->
<seq>
<par>
<video region="video1" src="{ $vname1 }"
clipBegin="{ $unit1/begin }" clipEnd="{ $unit1/end }" fill="freeze"/>
<!-- Bild in Region video2 -->

</par>
<par dur="media">
<video region="video2" src="{ $vname1 }"
clipBegin="{ $unit2/begin }" clipEnd="{ $unit2/end }" fill="freeze"/>
<!-- Bild in Region video1 -->

</par>
</seq>

```

Code 4.23: Ausschnitt aus paraRelation\_show.xsl

#### 4.5.5 Auswertung der Text-Bild-Relationen

Die in Abb. 2.34 dargestellte Auswertung der Häufigkeiten der verschiedenen Text-Bild-Relationen wird durch das XSLT-Stylesheet `textImage_report.xsl` generiert. Eingabedokument der Transformation ist die Analyse der Text-Bild-Relationen (hier: `textImage_analysis.xml`). Das XSLT-Stylesheet liest die in `textImage_relation.xml` gespeicherte Taxonomie ein. Diese wird rekursiv durchlaufen, wobei gezählt wird, wie viele Relationen des gerade betrachteten Typs insgesamt und an den einzelnen Orten im Analysedokument vorhanden sind. Zur Visualisierung der Baumstruktur der Taxonomie wird das Ergebnis passend eingerückt. Diese Einrückung berechnet das benannte Template einzug anhand der Länge der `id` eines `<relation>`-Elementes. Erzeugt wird ein FO-Dokument, welches anschließend in das PDF-Format gerendert wird.

```
<!-- Relationen aus textImage_analysis.xml -->
<xsl:variable name="rel" select="//relation"/>

<!-- die verschiedenen Orte -->
<xsl:variable name="loc" select="distinct-values(//unit/@location)"/>

<!-- Taxonomie der TextBild-Relationen einlesen -->
<xsl:variable name="taxonomie"
  select
    = "document('../xml/textimage_relation.xml')//textImageRelations/relation"/>
  <!-- Variable enthält eine Sequenz von relation-Elementen -->

<xsl:template match="/">
  <!-- fo Grundgerüst -->
  <fo:root xmlns:fo="http://www.w3.org/1999/XSL/Format">
    <fo:layout-master-set>
      ...
    </fo:layout-master-set>

    <fo:page-sequence master-reference="my-page">
      <fo:flow flow-name="xsl-region-body">
        <fo:block font-size="12pt" font-weight="bold" padding-after="6pt">
          Auswertung Text-Bild-Relationen</fo:block>
        <fo:block font-size="10pt" padding-after="6pt">
          Anzahl gesamt: <xsl:value-of select="count($rel)"/>
          <!-- Anzahl für die einzelnen Orte -->
          (<xsl:for-each select="$loc">
            <xsl:value-of select="."/>: <xsl:variable name="l" select="."/>
            <xsl:value-of select="count($rel[../@location=$l])"/>
            <xsl:choose>
              <xsl:when test="position()=last()">,&#160;</xsl:when>
              <xsl:otherwise></xsl:otherwise>
            </xsl:choose>
          </xsl:for-each>
        </fo:block>

        <xsl:apply-templates select="$taxonomie"/>

      </fo:flow>
    </fo:page-sequence>
  </fo:root>
</xsl:template>

<!-- Hier wird die Taxonomie durchlaufen -->
<xsl:template match="relation">
  <xsl:variable name="v" select="@id"/>
```

```

<!-- Länge der id = Tiefe im Baum -->
<xsl:variable name="leng" select="string-length(@id)+1"/>
<fo:block font-size="9pt">
  <!-- Einzug entsprechend der Tiefe im Baum -->
  <xsl:call-template name="einzug">
    <xsl:with-param name="anzahl" select="$leng*2"/>
  </xsl:call-template>
  <xsl:variable name="anz" select="count($rel[substring(@r,0,$leng)=$v])"/> &#160;
  <xsl:choose>
    <xsl:when test="$anz > 0">
      <fo:inline font-weight="bold">
        <xsl:value-of select="@name"/>:&#160; <xsl:value-of select="$anz"/>
      </fo:inline>
      <!-- Anzahl für die einzelnen Orte -->
      <xsl:for-each select="$loc">
        <xsl:value-of select="."/>: <xsl:variable name="l" select="."/>
        <xsl:value-of
          select="count($rel[substring(@r,0,$leng)=$v][. /@location=$l])"/>
        <xsl:choose>
          <xsl:when test="position()=last()">&#160;</xsl:when>
          <xsl:otherwise></xsl:otherwise>
        </xsl:choose>
      </xsl:for-each>
    </xsl:when>
    <xsl:otherwise>
      <xsl:value-of select="@name"/>:&#160; <xsl:value-of select="$anz"/>
    </xsl:otherwise>
  </xsl:choose>
</fo:block>
<xsl:apply-templates/>
</xsl:template>

<!-- Einrückung: Zahl der Zeichen wird als Parameter übergeben -->
<xsl:template name="einzug">
  <xsl:param name="anzahl"/>
  <xsl:param name="zaehler" select="1"/>
  <xsl:value-of select="'&#8211;'" />
  <xsl:if test="not($anzahl=$zaehler)">
    <xsl:call-template name="einzug"> <!-- Rekursion -->
      <xsl:with-param name="anzahl" select="$anzahl"/>
      <xsl:with-param name="zaehler" select="$zaehler+1"/>
    </xsl:call-template>
  </xsl:if>
</xsl:template>

```

Code 4.24: Ausschnitt aus textImage\_report.xsl

## 4.6 Bildergeschichten

Sowohl in einer Bildergeschichte als auch im Film steht eine ikonisch bildliche Repräsentation kombiniert mit einer bildlich dargestellten Entfaltung über die Zeit, die für narrative Zwecke nutzbar ist, im Vordergrund, wie [Bateman 2013, S. 57] ausführt:

*»Both comic and film typically share a reliance on iconic pictorial representations combined with a depicted unfolding over time employed for narrative purposes.«*

Aufgrund dieser Gemeinsamkeit von Film und Bildergeschichte macht es Sinn, das Beispielvideo als Genre »Bildergeschichte« zur repräsentieren. Abb. 4.9 stellt die »Transformation« vom Film zur Bildergeschichte dar: Statt eines »bewegten Bildes« wird ein »statisches Bild« angezeigt; aus dem gesprochenen Text wird nun sichtbarer geschriebener Text. Für den Beobachter bedeutet dies, dass er nicht über einen auditiven Kanal verfügen muss. Eine Bildergeschichte erlaubt eine rein visuelle Präsentation auf einem geeigneten Device.



Abb. 4.9: »Transformation« vom Film zur Bildergeschichte

Wesentliche Unterschiede zwischen dem Betrachten des Ausgangsvideos und dem Betrachten der dazu »analogen« Bildergeschichte liefern Interaktionen, die in die Bildergeschichte integriert werden und die der Situation »ein Buch lesen« nachempfunden sind. Während ein Video am Beobachter »vorbeirauscht«, sollte eine Bildergeschichte seitenbasiert ausgabefähig sein, so dass ein Beobachter selbst vor- oder auch zurückblättern kann und hier auch muss, was den Vorteil hat, dass er Bild und Text so lange betrachten kann, wie er möchte. Mit der Umstellung auf ein seitenbasiertes Genre »erbt« man auch deren Navigationsmöglichkeiten. So kann durch die Anzeige eines Inhaltsverzeichnisses nicht nur ein direkter Sprung zu einem bestimmten »Kapitel« leicht ermöglicht werden, sondern auch die Größe des Kapitels über die Paginierung geschätzt werden.

### Erfüllung der Randbedingungen

Beim Blättern durch eine Bildergeschichte werden Bild und ggf. Text in gleicher Reihenfolge wie im Filmbild gezeigt, daher sind die Randbedingungen (S1) bis (S5) »per se« erfüllt. Da auf einer »Seite« nur der Text einer einzelnen Einstellung präsentiert wird, ist Randbedingung (P1) nicht von Bedeutung. Randbedingung (P2) ist erfüllt, da auf einer »Seite« Bild und Text der gleichen Einstellung gezeigt werden. Die Anzeige eines Inhaltsverzeichnisses verletzt diese Randbedingung nicht, da das Inhaltsverzeichnis keine *Content Portion* darstellt, sondern lediglich zur Information und Navigation dient. Geht der Sprechertext über Einstellungsgrenzen hinweg, verlangt Randbedingung (P3), dass dies erkennbar sein soll. Hier bietet sich folgende Lösung an: In

der Einstellung, in der der einstellungsübergreifende Text beginnt, werden dem Text »...« angefügt und in der folgenden Einstellung, in der der einstellungsübergreifende Text ja fortgesetzt wird, werden »...« dem Text vorangestellt. Randbedingung (P4) ist nicht relevant, da keine Text-Video-Kombination vorliegt.

Basierend auf diesen Vorüberlegungen werden zwei Formen einer Bildergeschichte umgesetzt: eine »einfache« und eine »detaillierte« Bildergeschichte.

#### 4.6.1 Einfache Bildergeschichte

Abb. 4.10 zeigt einen Screenshot der »einfachen« Bildergeschichte (picStory\_simple.htm). »Einfach« bedeutet hier, dass für jede Einstellung des Videos genau ein Bild gezeigt wird. Verwendet werden die Bilder, die für das Transkriptionsprotokoll (s. 6.1) erzeugt wurden und sich bereits in der *Content Base* (s. Abb. 4.2, Abb. 4.3) befinden.



Abb. 4.10: Screenshot der einfachen Bildergeschichte in HTML

In der HTML-Ansicht im Browser kann der Betrachter selbst entscheiden, ob er die transkribierten Texte parallel zum Bild sehen möchte oder nicht. Zu dieser Auswahl dienen die Radio-Buttons im oberen Bereich. Blättern kann der Betrachter in der Geschichte mit Hilfe des »weiter« bzw. »zurück«-Links, wobei beim Übergang von einem Bild zum nächsten bzw. vorigen ein FadeOut/FadeIn-Effekt angezeigt wird. Beim FadeOut/FadeIn-Effekt verschwindet die angezeigte Seite langsam und die nächste wird langsam eingeblendet. Dies ist eine reine Designentscheidung, um das fehlende physikalische Umblättern zu ersetzen.

Im linken Bereich wird ein »Inhaltsverzeichnis« angezeigt. Hier werden die verschiedenen Orte und in eckigen Klammern die zugehörigen Einstellungen angezeigt, so dass ein direkter Sprung zur ersten Einstellung des entsprechenden Ortes möglich ist. Das jeweils angezeigte »Kapitel« ist farblich markiert.

Designentscheidungen (wie FadeOut/FadeIn oder eine farbliche Markierung) sind für einen Beobachter natürlich sehr wichtig, müssen aber die geforderten Randbedingungen einhalten.

### Technische Umsetzung in HTML

Die Bildergeschichte im HTML-Format wird mit Hilfe des XSLT-Stylesheet `picStory_simple.html.xsl` aus der XML-Repräsentation des Transkriptionsprotokolls (hier: `snowboard.xml`) generiert. Das XSLT-Stylesheet (s. Code 4.25) enthält im Wesentlichen eine `<xsl:for-each>`-Schleife, die über die Elemente `<unit>` iteriert und für jedes Element ein `<div>`-Element mit Bild und Text erzeugt. Die `<div>`-Elemente erhalten ein Attribut `id`, dessen Wert eine fortlaufende Zahl mit vorangestelltem »T« ist, so dass die `id`-Werte den Einstellungsnummern entsprechen. Durch die Schleife wird also für jede »Seite« der Bildergeschichte ein `<div>`-Element erzeugt. Eine JavaScript-Funktion (in Datei `picstory.js`, s. 6.3.2) sorgt dafür, dass nur die vom Betrachter ausgewählte Division sichtbar ist. Die `<xsl:if>`-Anweisungen vor bzw. nach der Textausgabe geben »...« aus, falls einstellungsübergreifender Text vorhanden ist.

```
<body>
  <h1>Einfache Bildergeschichte: <xsl:value-of select="movie/movie_info/name"/></h1>
  <div id="auswahl">
    <form>
      <input type="radio" name="r1" checked="checked"/> keine Texte
      <input type="radio" name="r1"/> alle Texte
    </form>
  </div>
  <div id="brueck">zurück</div>
  <div id="bweiter">weiter</div>
  <div id="container">
    <div id="inner">
      <xsl:for-each select="movie/units/unit">
        <xsl:variable name="pos" select="position()"/>
        <div class="slide" id="T{$pos}">
          <table width="400px">
            <tr>
              <td align="left"> <xsl:value-of select="@id"/> </td>
              <td align="right"> Ort: <xsl:value-of select="location/loc_name"/>
                <xsl:if test="location[loc_detail]"> - <xsl:value-of
                  select="location/loc_detail"/>
                </xsl:if>
            </td>
            </tr>
          </table>
          <div class="image"></div>
          <div <xsl:value-of select="@id"/></div>
          <div class="text">
            <!-- falls einstellungsübergreifender Text zur vorigen Einstellung vorliegt,
              werden ... vorangestellt-->
            <xsl:if test="./text/span[1]/@join='before'">
              <span>... </span>
            </xsl:if>
          </div>
        </div>
      </xsl:for-each>
    </div>
  </div>
```

```

        <xsl:value-of select="text" />

        <!-- falls einstellungsübergreifender Text zur nächsten Einstellung vorliegt,
        wird ... angezeigt-->
        <xsl:if test="./text/span[position()=last()]/@join='next'">
            <span> ...</span>
        </xsl:if>
    </div>
</div>
</xsl:for-each>

</div>
</div>
<!-- Inhaltsverzeichnis erzeugen -->
<xsl:call-template name="content" />
</body>

```

Code 4.25: Ausschnitt aus *picStory\_simple\_html.xsl*

Das Inhaltsverzeichnis wird durch das benannte Template `content` (s. Code 4.26) generiert. Darin wird das benannte Template `segment` aufgerufen, das eine ortsübergreifende Gruppierung berechnet (s. 4.5.1). Die `<p>`-Elemente des Inhaltsverzeichnisses erhalten als Wert des Attributes `id`, den `id`-Wert der anzuzeigenden Division. Dies ist in diesem Fall gerade die `id` der Einstellung des ersten Elementes der jeweiligen Gruppe, gespeichert als Wert des Attributes `begin`. Die darauf angewandte Funktion `substring(@begin,2)` entfernt das erste Zeichen (hier den Buchstaben »T«), somit wird sichergestellt, dass eindeutige `id`-Werte vergeben werden.

```

<!-- Benanntes Template zur Generierung des Inhaltsverzeichnisses -->
<xsl:template name="content">
    <xsl:variable name="segTree">
        <xsl:call-template name="segment" />
    </xsl:variable>
    <div id="inhalt">
        <fieldset style="width:160px">
            <legend>Inhalt</legend>
            <br/>
            <xsl:for-each select="$segTree/segments/segment" >
                <!-- id des ersten Elementes der Gruppe -->
                <xsl:variable name="vid1" select="substring(@begin,2)" />
                <p class="entry" id="{\$vid1}">
                    <xsl:value-of select="@location" />
                    [<xsl:value-of select="@begin" /> - <xsl:value-of select="@end" />]
                </p>
            </xsl:for-each>
        </fieldset>
    </div>
</xsl:template>

```

Code 4.26: Benanntes Template `content` zur Generierung des Inhaltsverzeichnisses in HTML

Die Interaktionen (Sprung zum Kapitel, FadeIn/FadeOut-Effekt beim Blättern, farbige Markierung des angezeigten Segmentes) sind mit JavaScript unter Zuhilfenahme des Frameworks

*jQuery*<sup>1</sup> in der Datei `picstory.js` (s. 6.3.2) realisiert.<sup>2</sup> Das Layout ist in der CSS-Datei `picstory.css` beschrieben.

### Technische Umsetzung in SMIL

Das Bildschirmlayout für die Präsentation der Bildergeschichte in SMIL ist identisch zum HTML-Layout jedoch ohne Radio-Buttons, da dieses Interaktionselement in SMIL fehlt. Daher werden in dieser Bildergeschichte grundsätzlich Bild und Text gemeinsam angezeigt.

Das XSLT-Stylesheet `picStory_simple_smil.xsl` erzeugt aus der XML-Repräsentation des Transkriptionsprotokolls (hier: `snowboard.xml`) die Bildergeschichte im SMIL-Format. Eine `<xsl:for-each>`-Schleife iteriert über alle `<unit>`-Elemente und erzeugt für jedes Element ein SMIL-Element `<par>`, das alle Informationen einer »Seite« der Bildergeschichte enthält. Ein »Umblättern« entspricht einem Sprung zum nächsten bzw. vorigen `<par>`-Element. Daher ist es notwendig, dass jedes `<par>`-Element auch den Code des Inhaltsverzeichnisses enthält. Zur Erzeugung des Inhaltsverzeichnisses wird das benannte Template `content` aufgerufen (s. Code 4.27). Im Unterschied zum entsprechenden Template für die HTML-Version wird diesem Template die `id` der Einstellung als Parameter übergeben. Dies ist notwendig, da die farbigen Positionsmarkierungen nicht zur Laufzeit durch eine Skriptsprache erzeugt werden können, und daher der Quellcode für den entsprechenden Eintrag im Inhaltsverzeichnis bereits die farbliche Markierung (Attribut `textColor="#ffff00"`) enthalten muss. Zur Realisierung des Überblendungseffektes beim »Umblättern« wird im `<head>`-Bereich die Überblendungsklasse `<transition xml:id="fade" type="fade" dur="3s"/>` definiert. Den Medienobjekten wird durch das Attribut `transIn="fade"` dieser Effekt hinzugefügt.

```
<!-- Benanntes Template zur Generierung des Inhaltsverzeichnisses -->
<xsl:template name="content">
  <xsl:param name="tid"/>

  <xsl:variable name="segTree">
    <xsl:call-template name="segment"/>
  </xsl:variable>

  <xsl:for-each select="$segTree/segments/segment">
    <xsl:variable name="vid1" select="number(substring(@begin,2))"/>
    <!-- id des ersten Elementes der Gruppe -->
    <xsl:variable name="p" select="position()"/>

    <xsl:variable name="color">
      <xsl:choose>
        <xsl:when test="$tid >=@begin and $tid &lt;=@end">#ffff00</xsl:when>
        <xsl:otherwise>#ffffff</xsl:otherwise>
      </xsl:choose>
    </xsl:variable>

    <a href="#par{${vid1]}" xmlns="http://www.w3.org/ns/SMIL">
```

<sup>1</sup> <http://jquery.com> [letzter Aufruf: 04.03.2015]

<sup>2</sup> Die Eventhandler sind an dieser Stelle im Code nicht sichtbar, da sie durch *jQuery* beim Laden der HTML-Seite hinzugefügt werden.

```

    <smilText region="c{$p}" textColor="{ $color}">
      <xsl:value-of select="@location"/>
      [<xsl:value-of select="@begin"/> - <xsl:value-of select="@end"/>]
    </smilText>
  </a>
</xsl:for-each>
</xsl:template>

```

Code 4.27: Benanntes Template content zur Generierung des Inhaltsverzeichnisses in SMIL

Die einzelnen Textlinks des Inhaltsverzeichnisses werden in verschiedenen Regionen platziert. Diese werden vorab innerhalb des <layout>-Bereiches durch die in Code 4.28 dargestellte <xsl:for-each-group>-Schleife generiert.

```

<layout>
...
  <!-- Regionen für content -->
  <xsl:for-each-group select="//unit" group-adjacent="location/loc_name">
    <xsl:variable name="p" select="position()"/>
    <region xml:id="c{$p}" width="140" height="20" left="30" top="{50+22*$p}"
      backgroundColor="#000033" z-index="2"/>
  </xsl:for-each-group>
...
</layout>

```

Code 4.28: for-each-group-Schleife zur Generierung der Regionen des Inhaltsverzeichnisses

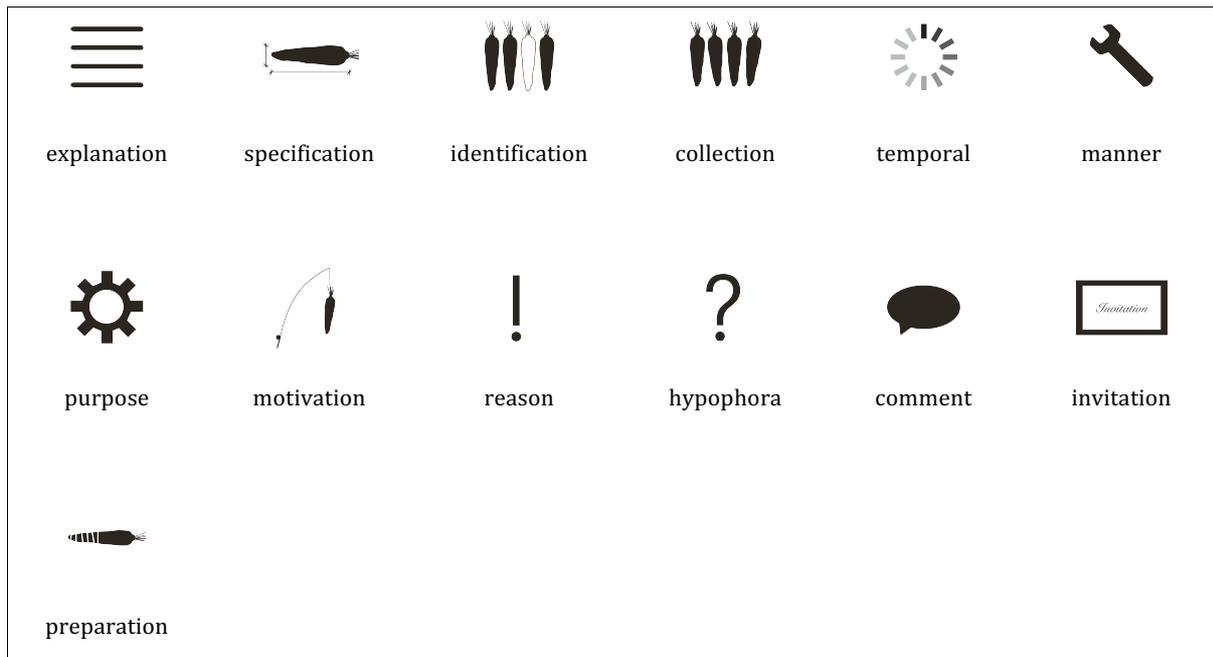
#### 4.6.2 Detaillierte Bildergeschichte

Insbesondere in längeren Einstellungen fällt es schwer, ein »typisches« Bild für diese Einstellung zu identifizieren. Daher wurde auch eine Bildergeschichte entwickelt, die den Beobachtern eine größere Detaillierung bietet. Im Unterschied zur einfachen Bildergeschichte, die die Einteilung des Videos übernimmt, ist in der detaillierten Bildergeschichte der Text zusätzlich dramaturgisch leitend. Die detaillierte Bildergeschichte zeigt für jedes Textsegment<sup>1</sup> jeder Einstellung ein Bild bzw. Bild und Text. Da die 42 Einstellungen insgesamt 114 Textsegmente beinhalten<sup>2</sup>, besteht die detaillierte Bildergeschichte aus 114 »Seiten«. Der Ressourcenpool (s. Abb. 4.3) muss hierfür um die fehlenden Bilder ergänzt werden.

Zusätzlich wird bei der Darstellung die Analyse der Text-Bild-Relationen (s. 2.5.4) berücksichtigt. Die zwischen Text und Bild vorliegende Relation wird durch ein passendes Icon (s. Tab. 4.2) visualisiert, um dem Leser die Möglichkeit zu geben, abhängig von seinem Informationsbedürfnis bei der Lektüre selektiv zu sein. Abb. 4.11 zeigt einen Screenshot der detaillierten Bildergeschichte zum Video »Snowboard«.

<sup>1</sup> In der XML-Darstellung sind dies die <span>-Elemente.

<sup>2</sup> Nur 11 Einstellungen beinhalten ein einziges Textsegment.



Tab. 4.2: Übersicht: Icons zur Visualisierung der Text-Bild-Relationen



Abb. 4.11: Screenshot der detaillierten Bildergeschichte in HTML

### Technische Umsetzung in HTML

Das XSLT-Stylesheet `picStory_detail_html.xsl` generiert die detaillierte Bildergeschichte im HTML-Format. Eingabedatei ist die XML-Repräsentation der Einstellungen (hier: `snowboard.xml`). Die einzelnen »Seiten« der Bildergeschichte werden durch eine `<xsl:for-each>`-Schleife, die über die Elemente `<span>` iteriert, erzeugt. Die `id`-Werte der `<span>`-Elemente haben die Form `Txx_yy`. Es ist jedoch nicht sinnvoll, den erzeugten `<div>`-Elementen in HTML bzw. `<par>`-Elementen in SMIL diese `id`-Werte zu geben, da sich das Vor- bzw. Zurückblättern einfacher realisieren lässt,

wenn deren `id`-Werte fortlaufend sind. Die `id`-Werte im Inhaltsverzeichnis benötigen daher die entsprechende Positionsnummer des anzuzeigenden `<span>`-Elementes als Wert (s. Code 4.29).

```
<div id="inhalt" style="color:#ffffff">
  <fieldset style="width:90px">
    <legend>Inhalt</legend><br>
    <p class="entry" id="1">Berg</p>
    <p class="entry" id="11">Werkstatt</p>
    <p class="entry" id="37">Berg</p>
    .....
  </fieldset>
</div>
```

Code 4.29: Erzeugter HTML-Code für das Inhaltsverzeichnis der detaillierten Bildergeschichte

Zur Berechnung dieser Werte wird zu Beginn des XSLT-Stylesheets die Hilfsvariable `span` definiert. Sie enthält für jedes `<span>`-Element der Eingabedatei ein leeres `<span>`-Element mit den Attributen `id`, Attributwert ist der `id`-Wert des `<span>`-Elementes der Eingabedatei, `id_upp`, Attributwert ist der `id`-Wert der Großelternelementes `<unit>`, sowie `pos`, Attributwert ist die Positionsnummer des `<span>`-Elementes, die durch die XPath-Funktion `position()` berechnet wird.

```
<!-- "Zwischenspeicher" für span-Elemente
      wird bei der Generierung des Inhaltsverzeichnisses zur Berechnung der Position
      im Dokument anhand der id verwendet -->
<xsl:variable name="span">
  <spans>
    <xsl:for-each select="//span">
      <span id="{@id}" id_upp="{../../@id}" pos="{position()}" />
    </xsl:for-each>
  </spans>
</xsl:variable>
```

Code 4.30: Definition der Hilfsvariablen `span`

Zur Erzeugung des Inhaltsverzeichnisses gruppiert eine `<xsl:for-each-group>`-Schleife die `<span>`-Elemente des gleichen Ortes. Der `id`-Wert eines Eintrags des Inhaltsverzeichnisses (HTML: `<p>`) ist die Positionsnummer des ersten Elementes einer Gruppe. Die `id` des ersten Elementes der Gruppe wird in der Hilfsvariablen `vid1` gespeichert. Die Positionsnummer des `<span>`-Elementes mit dieser `id` liefert der XPath-Ausdruck `$span/spans/span[@id=$vid1]/@pos`.

```
<!-- Benanntes Template zur Erzeugung des Inhaltsverzeichnisses -->
<xsl:template name="content">
  <xsl:variable name="segTree">
    <xsl:call-template name="segment" />
  </xsl:variable>

  <div id="inhalt">
    <fieldset style="width:160px">
      <legend>Inhalt</legend>
      <br/>
      <xsl:for-each select="$segTree/segments/segment" >
        <!-- id des ersten Elementes der Gruppe -->
        <xsl:variable name="vid1" select="concat(@begin,'_01')"/>
        <!-- Positionsnummer des ersten Elementes der Gruppe -->
        <xsl:variable name="vid2" select="$span/spans/span[@id=$vid1]/@pos"/>
```

```

    <p class="entry" id="{\$vid2}">
      <xsl:value-of select="@location"/>
      [<xsl:value-of select="@begin" /> - <xsl:value-of select="@end"/>]
    </p>
  </xsl:for-each>
</fieldset>
</div>
</xsl:template>

```

Code 4.31: Erzeugung des Inhaltsverzeichnisses für die detaillierte Bildergeschichte

### Technische Umsetzung in SMIL

Analog zur einfachen Bildergeschichte in SMIL werden auch in der detaillierten Bildergeschichte immer Bild und Text gemeinsam angezeigt. Das XSLT-Stylesheet `picStory_detail_smil.xml`, das die Bildergeschichte erzeugt, ist analog zu dem der einfachen Bildergeschichte aufgebaut, allerdings iteriert die `<xsl:for-each>`-Schleife zur Erzeugung der `<par>`-Elemente für die »Seiten« über die `<span>`-Elemente, wie dies auch im XSLT-Stylesheet zur Erzeugung der detaillierten Bildergeschichte im HTML-Format der Fall ist. Auch die Linkgenerierung im Inhaltsverzeichnis funktioniert analog zur HTML-Lösung.

## 4.7 Modifikation der Alternanz unter Erhalt der Kohärenzrelation

Wie bereits mehrfach betont, sind für das Verständnis eines Films die Kohärenzbeziehungen zwischen Einstellungen bzw. Einstellungsmengen von entscheidender Bedeutung. Eine Änderung oder sogar Zerstörung dieser Relationen, z. B. durch eine Änderung in der Layoutierung, kann den Beobachter verwirren und sogar dazu führen, dass der Film unverständlich wird oder eine andere als die intendierte Bedeutung erlangt. Die in 4.2 formulierten Randbedingungen (S2), Szenen sollen im Layout nicht, und (S3), Sequenzen nur in ihren zeitdiegetischen Lücken auseinandergerissen werden, sorgen dafür, dass wesentliche Kohärenzrelationen sich auch im Layout widerspiegeln. In den folgenden Abschnitten wird die Kohärenzrelation zwischen den alternanten Segmenten, die paradigmatisch als *paratactic – external | contrast* klassifiziert wurde, betrachtet. Im Teil dieses Abschnitts (s. 4.7.1) wird ein Genre-Experiment entwickelt, das die Kohärenzrelation selbst stärker betont, im zweiten Teil (s. 4.7.2) ein Genre-Experiment, das die durch die Kohärenzrelation verbundenen Objekte stärker betont.

Im darauf folgenden Abschnitt (s. 4.8) wird aus dem bestehenden Material eine alternative Alternanz entwickelt, wodurch dann auch eine alternative Kohärenzrelation etabliert werden kann. Weglassen der Kohärenzrelation (s. 4.9) führt dazu, dass das Dokument in Teildokumente »zerfällt«.

#### 4.7.1 Betonung der Kohärenzrelation

Eine Alternanz ist eine Folge von *Content Portions* vom Typ Video mit der abwechselnden Darstellung von i. Allg. zwei Handlungssträngen, einer im Raumzeitgebiet  $SpaceTime_1$ , der andere im Raumzeitgebiet  $SpaceTime_2$ . Entscheidend für das Verständnis der Alternanz ist die Etablierung einer Kohärenzrelation zwischen (mindestens) einem Objekt aus  $SpaceTime_1$  und (mindestens) einem Objekt aus  $SpaceTime_2$ . Diese Kohärenzrelation kann stärker betont werden, wenn die *Content Portions*, die diese Objekte enthalten, räumlich parallel und zeitlich gleichzeitig angezeigt werden.

Nach der syntagmatischen Analyse des Beispielvideos »Snowboard« (s. 2.5.3) liegen zwischen den Berg- und Werkstatt-Segmenten folgende Kohärenzrelationen vor:

$R(\text{Johanna}, \text{Snowboard})$ : »Johanna gibt das Snowboard in die Werkstatt (zurück).«

$R'(\text{Johanna}, \text{Snowboard})$ : »Johanna erhält das Snowboard zum Test.«

Bei der Untersuchung der thematischen Entfaltung (s. 4.2) wurde der Übergang von Komplikation (ein Problem ist aufgetreten) zu (Teil-)Lösung als »Verknüpfungsrelation« zwischen den Teildokumenten »Johanna möchte Snowboard fahren« und »Herstellung eines Snowboards« identifiziert.

Um diese Kohärenzrelationen stärker zu betonen, bietet es sich an, *Content Portions* mit folgenden Inhalten gleichzeitig zu präsentieren:

- Darstellung des Problems: Johanna auf dem Berg mit dem Snowboard vor der Weiterbearbeitung,
- Handlungen zur Lösung des Problems: Weiterbearbeitung des Snowboards in der Werkstatt,
- Lösung des Problems: Johanna auf dem Berg mit dem Snowboard nach der Weiterbearbeitung.

Wurden bisher *Content Portions* aus zwei verschiedenen Raumzeitgebieten im Wechsel gezeigt, erfolgt nun die gleichzeitige Anzeige von drei *Content Portions*. Abb. 4.12 zeigt das geplante Screendesign zur Darstellung des Genre-Experimentes »Arbeitsfortschritt«.

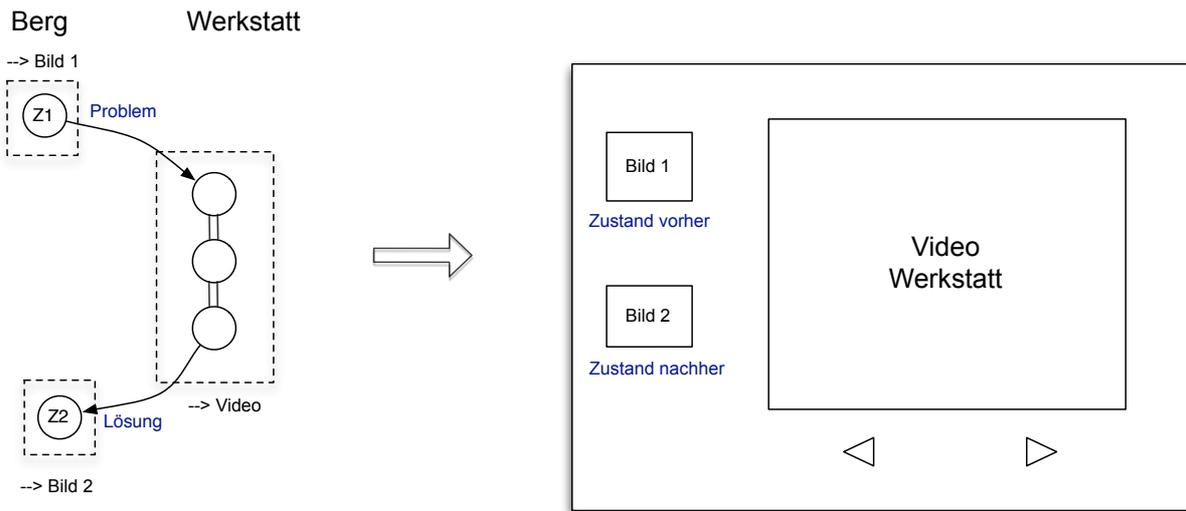


Abb. 4.12: Screendesign für das Genre-Experiment »Arbeitsfortschritt«

Da die gleichzeitige Anzeige von drei Videosegmenten für die Beobachter verwirrend sein kann, wird für die Anzeige der *Content Portions* »Darstellung des Problems« (= »Zustand vorher«) und »Lösung des Problems« (= »Zustand nachher«) jeweils eine *Content Portion* vom Typ Bild verwendet. Im Videobereich werden die Handlungen in der Werkstatt gezeigt, die das Snowboard vom »Zustand vorher« zum »Zustand nachher« weiterverarbeiten. Diese Videoausschnitte zeigen jeweils eine deskriptive Themenentfaltung, welche durch die Anzeige der »Zustandsbilder« noch verstärkt wird und stellen – inhaltlich gesehen – jeweils auch den Arbeitsfortschritt in einem Werkstatt-Segment dar. Aus diesem Grund wird ein entsprechender Text auch als Überschrift für die Zuschauer angezeigt.

### Erfüllung der Randbedingungen

Die einzelnen Videosequenzen zeigen jeweils ein »vollständiges« Segment in der Werkstatt zwischen den jeweiligen Wechseln vom Handlungsort Berg zur Werkstatt und zurück, daher ist Randbedingung (S1) erfüllt. Da Szenen und Sequenzen hierbei nicht auseinandergerissen werden, sind die Randbedingungen (S2) und (S3) erfüllt. Ist ein Segment vollständig abgespielt, muss der Beobachter auf den »Weiter«-Button klicken, um so das Abspielen des nächsten Werkstatt-Segmentes und die parallele Anzeige der Zustandsbilder zu starten. Auf diese Weise können die Werkstatt-Segmente in ihrer diegetischen Reihenfolge durchlaufen werden, so dass auch (S4) erfüllt ist. Eine Alternanz im Sinne von Definition 2.14 wird nicht etabliert, daher ist (S5) nicht relevant.

Da hier eine Bild-Video-Kombination vorliegt, ist bezüglich der parallelen Anordnung von *Content Portions* nur Randbedingung (P4) relevant. Diese ist erfüllt, da die Bilder zu einer Einstellung gehören, die dem Werkstatt-Segment zeit-diegetisch vorangeht bzw. folgt.

### Technische Umsetzung in SMIL und HTML5

Die entsprechenden SMIL und HTML5-Dokumente (work\_progress.smil und work\_progress1.htm bis work\_progress4.htm) werden durch das XSLT-Stylesheet work\_progress.xsl gleichzeitig erzeugt. Das Stylesheet verwendet die durch das Template segment berechnete Segmentierung (s. 4.5.1). In einer `<xsl:for-each>`-Schleife werden die Segmente, bei denen das Attribut `location` den Wert »Werkstatt« hat, durchlaufen. Anhand der Einstellungsnummer der ersten bzw. letzten Einstellung des Segmentes kann die Beginn- bzw. Endezeit des zu zeigenden Videoausschnittes bestimmt werden. Die anzuzeigenden Bilder stammen aus der letzten Einstellung des vorangehenden Geschwisterelementes bzw. der ersten Einstellung des nachfolgenden Geschwisterelementes, bei dem das Attribut `location` den Wert »Berg« besitzt. Zur Berechnung des Dateinamens eines Bildes wird die XSLT-Funktion `fu:bname` (s. 6.3.1) verwendet. Code 4.1 zeigt einen Ausschnitt des Stylesheets zur Generierung des SMIL-Dokumentes.

```
<xsl:for-each select="$segTree/segments/segment[@location='Werkstatt']">
  <xsl:variable name="pos" select="position()"/>
  <par xml:id="par{$pos}">
    <smilText region="h1" textColor="#000099" textFontSize="18px" textFontWeight="bold">
      <xsl:value-of select="$moviename"/>: Arbeitsfortschritt
      <xsl:value-of select="$pos"/> / <xsl:value-of select="last()"/>
    </smilText>
    
    
    <video src="{vname1}" clipBegin="{@clipBegin}" clipEnd="{@clipEnd}" region="vid1" />
    ...
  </par>
</xsl:for-each>
```

Code 4.32: Ausschnitt aus work\_progress.xsl

Abb. 4.13 zeigt einen Screenshot der Darstellung in SMIL.

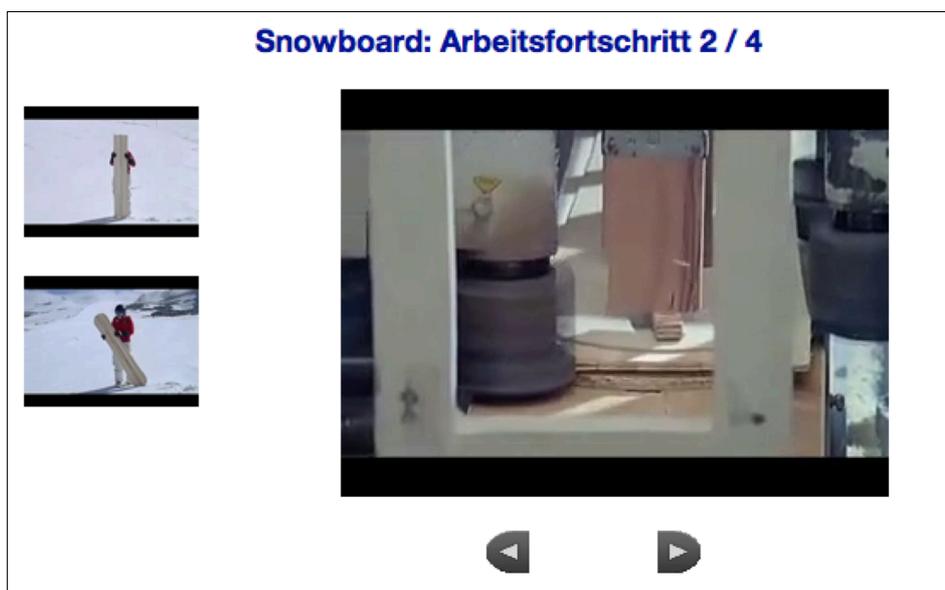


Abb. 4.13: Screenshot des Genre-Experimentes »Arbeitsfortschritt«

#### 4.7.2 Betonung des kohärent Relationierten

In einer Alternanz werden die kohärent verbundenen Videosegmente im Wechsel gezeigt. Soll betont werden, welche Segmente in einer Kohärenzrelation stehen, bietet sich eine Darstellung an, die diese Videosegmente räumlich parallel anordnet und zeitlich gleichzeitig präsentiert. In diesem Fall liegt dann – abweichend zu der in 2.2.1 vorausgesetzten Situation – ein Layoutprozess vor, der mehr als einen Ausgabestrom erzeugt, was in *Split-Screen-Technik* realisiert werden kann. Nach [Bateman 2013, S. 58] wird die *Split-Screen-Technik* vorwiegend eingesetzt, um simultan stattfindende Handlungen parallel darzustellen, etwa um Vergleiche oder Gegensätze zu visualisieren:

*»Split-screens most commonly involve simultaneity plus a strong sense of comparison or contrast.«*

Abb. 4.14 zeigt zwei mögliche Varianten eines Screendesigns für das Beispielfideo »Snowboard«.

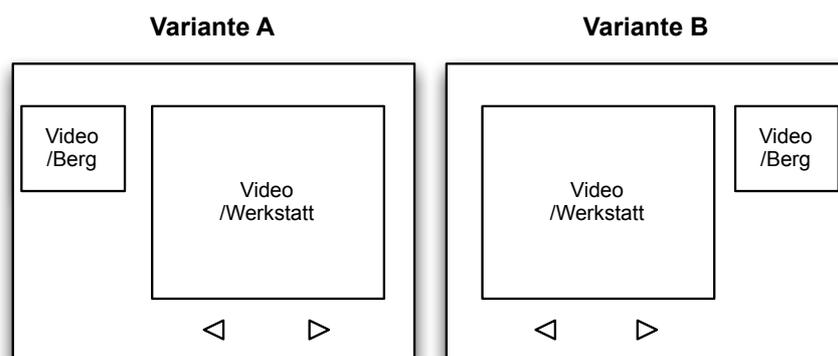


Abb. 4.14: Varianten des Screendesigns für das Genre »Produktion und Test«

In den in Abb. 4.14 parallel angeordneten Videosegmenten werden jedoch keine gleichzeitig stattfindenden Handlungen dargestellt, sondern kohärent verbundene Segmente. Für das Beispielfideo »Snowboard« gibt es – wie Abb. 4.15 zeigt – grundsätzlich zwei Varianten der Kombination von Segmenten entlang der unterstellten Kohärenzrelation.

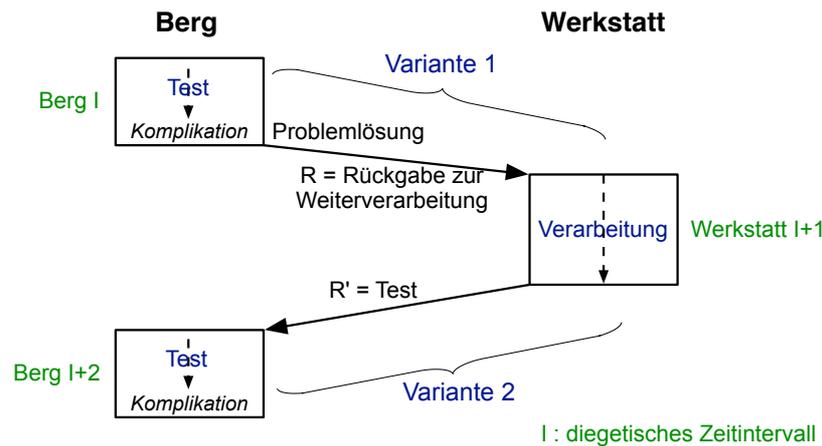


Abb. 4.15: Varianten der Kombination von Content Portions

In Variante 1 werden die Segmente, die durch die Kohärenzrelation »Rückgabe zur Weiterverarbeitung« verbunden sind, gemeinsam präsentiert. Ein Beobachter sieht also das zeitlich vorangehende Berg-Segment gleichzeitig mit dem nachfolgenden Werkstatt-Segment, d. h. er sieht das Testen des Snowboard durch Johanna, welches mit einer Komplikation endet, und gleichzeitig den daran anschließenden Produktionsschritt. Variante 2 zeigt die durch die Kohärenzrelation »Test des Snowboards« verbundenen Segmente, also ein Werkstatt-Segment gemeinsam mit dem zeitlich nachfolgenden Berg-Segment. Ein Beobachter sieht also während der Darstellung eines Produktionsschrittes bereits das »Ziel«, den »Erfolg« des Produktionsschrittes: »So kann Johanna nach diesem Produktionsschritt fahren.« Da Variante 2 für die Beobachter vermutlich attraktiver ist, wird diese als Ausgangsbasis für die weiteren Überlegungen zum Genre-Experiment »Produktion und Test« verwendet.

Bei beiden in Abb. 4.14 vorgeschlagenen Layoutvarianten ist der Bereich für das Videosegment »Werkstatt« flächenmäßig größer als der Bereich für das Videosegment »Berg«. Dies ist sinnvoll, da der eigentliche Zweck des Videos »Snowboard« die Darstellung des Produktionsprozesses ist. Es bleibt noch zu diskutieren, ob der Bereich »Berg« links vom Bereich »Werkstatt« (Variante A) platziert werden soll, oder die umgekehrte Anordnung (Variante B) geeigneter ist. Hierbei ist zu beachten, dass die Lesereihenfolge im deutschsprachigen Raum von links nach rechts ist. Für Variante A spricht, dass ein Beobachter, der das Originalvideo vorab gesehen hat, sich an die ersten Einstellungen mit Johanna auf dem Berg erinnert und daher evtl. deren Darstellung als »Erstes« – im Layout also links – erwartet. Bei Variante B dagegen werden die zeitlich vorangehenden Segmente links von den zeitlich nachfolgenden gezeigt, d. h. die Anordnung der *Content Portions* spiegelt für den Beobachter die zeitliche Ordnung wieder. Da die Abbildung der zeitlichen Ordnung als sinnvoller erachtet wird, wird hier Variante B umgesetzt.

Problematisch bei der geplanten Umsetzung ist, dass das Video vollständig von einem Sprecher kommentiert wird. Bei einer parallelen Präsentation der Videosegmente würden sich also zwei Audioströme überlagern. Da der eigentliche Zweck des Videos »Snowboard«, wie bereits er-

wähnt, die Darstellung des Produktionsprozesses ist, ist es sinnvoll, die Videosegmente mit Johanna auf dem Berg ohne Ton abzuspielen.

Ein zweites Problem stellen die unterschiedlichen Längen der parallel zu zeigenden Segmente dar. Tab. 4.3 zeigt eine Übersicht.

Werkstatt			Berg		
Einstellungen	von - bis	Dauer	Einstellungen	von - bis	Dauer
T04 – T10	0:30 – 1:57	87 s	T11	1:58 – 2:18	20 s
T12 – T15	2:19 – 3:05	46 s	T16	3:06 – 3:17	11 s
T17 – T35	3:18 – 6:14	176 s	T36	6:15 – 6:26	11 s
T37 – T40	6:27 – 6:53	26 s	T41 – T42	6:54 – 7:39	45 s

Tab. 4.3: Zeitliche Dauer von Berg- und Werkstatt-Segmenten

Man erkennt: Die ersten drei Werkstatt-Segmente sind länger als die Berg-Segmente, zwei und drei sogar deutlich, dagegen ist das vierte Werkstatt-Segment etwas kürzer als das Berg-Segment. Es ergeben sich daraus die folgenden möglichen Darstellungen:

- (1) Beide Segmente starten gleichzeitig und bleiben, wenn sie beendet sind, auf dem letzten Bild stehen.
- (2) Ist das Berg-Segment kürzer als das Werkstatt-Segment, wird es solange wiederholt, bis das Werkstatt-Segment beendet ist.
- (3) Das kürzere Video beginnt mit einer festen Verzögerungszeit.
- (4) Beide Videos enden zum gleichen Zeitpunkt (Endesynchronisation).

### Erfüllung der Randbedingungen

Da jeweils ein komplettes Segment zwischen den Ortswechseln gezeigt wird, sind die Randbedingungen (S1) bis (S3) erfüllt. Randbedingung (S4) ist erfüllt, da die Reihenfolge, in der die Videosegmente durchlaufen werden können, der diegetischen Reihenfolge entspricht. Da hier eine Video-Video-Kombination vorliegt, sind die Randbedingungen zur parallelen Anordnung (P1) bis (P3) nicht relevant. Randbedingung (P4) ist erfüllt, da die parallel gezeigten Videosegmente zeit-diegetisch aufeinanderfolgend sind.

### Technische Umsetzung in HTML und SMIL

Das XSLT-Stylesheet `video_synchron.xsl` erzeugt die gewünschten HTML5- und SMIL-Dateien. Es verwendet die durch das Template `segment` berechnete Segmentierung (s. 4.5.1). In einer `<xsl:for-each>`-Schleife werden die Segmente, bei denen das Attribut `location` den Wert »Werkstatt« hat, durchlaufen und die Videos platziert.

Um alle o. a. Darstellungsvarianten mittels eines XSLT-Stylesheets zu ermöglichen, besitzt das Stylesheet den globalen Parameter `variante`, so dass der Beobachter sich die Variante seiner Wahl erzeugen kann (Ausgabedateien: `video_synchron_v.smil` bzw. `video_synchron_v_1.htm` bis

video\_synchron\_v\_4.htm, v steht für die Nummer der Variante). Dementsprechend ist im Stylesheet eine Fallunterscheidung programmiert (s. Code 4.33). Für die HTML-Lösung wird für Variante 1 das Abspielen beider Videos durch die JavaScript-Funktion `videoPlayLoop` realisiert. Für Variante 2 wird, wenn das Berg-Segment kürzer als das Werkstatt-Segment ist, die Wiederholung des Berg-Segmentes durch die JavaScript-Funktion `videoPlay(id, begin,dur,loop)` gesteuert, wobei der vierte Parameter auf `true` gesetzt wird. Die Funktion `videoPlayStop(id1,id2,begin,dur)` steuert das Abspielen des Werkstatt-Segmentes und beendet auch das (wiederholte) Abspielen des Berg-Segmentes. Für Variante 3 wird das Video, das verzögert abgespielt werden soll, durch die JavaScript-Funktion `videoDelay(id,delay,begin,dur)` gesteuert, wobei für den Parameter `delay` der feste Wert 3000 (Millisekunden) übergeben wird. Variante 4 ähnelt der Variante 3, allerdings wird als Wert für den Parameter `delay`, die Differenz der Abspieldauern der beiden Segmente übergeben, um so eine Endesynchronisation zu realisieren. Um die Lautstärke des Berg-Segmentes auf 0 zu setzen, wird das Attribut `muted="true"` dem entsprechenden `<video>`-Element hinzugefügt.

Für die SMIL-Lösung werden für Variante 1 lediglich die `<video>`-Elemente für Berg und Werkstatt in einem `<par>`-Container platziert. Für Variante 2 wird die evtl. Wiederholung des Berg-Segmentes durch das Attribut `repeatDur` gesteuert, Attributwert ist die Dauer des gleichzeitig zu zeigenden Werkstatt-Segmentes, die durch die XSLT-Funktion `fu:dauer` (s. 6.3.1) berechnet wird. Für Variante 3 erhält das `<video>`-Element des kürzeren Segmentes das Attribut `begin="3s"` und für Variante 4 erhält das Attribut `begin` den Differenzwert der Abspieldauern der beiden Segmente. Um die Lautstärke des Berg-Segmentes auf 0 zu setzen, wird in SMIL dem entsprechenden `<video>`-Element das Attribut `soundLevel="0"` hinzugefügt.

```

<!-- Globaler Parameter für die Variante -->
  <xsl:param name="variante"/>
  ...
<body>
  <xsl:attribute name="onload">
    <xsl:choose>
      <xsl:when test="$variante='1'">
        videoPlayLoop('video1',<xsl:value-of select="$begin1"/>,
          <xsl:value-of select="$dauer1"/>,false)
        videoPlayLoop('video2',<xsl:value-of select="$begin2"/>,
          <xsl:value-of select="$dauer2"/>,false)
      </xsl:when>

      <xsl:when test="$variante='2'">
        <xsl:choose>
          <xsl:when test="$dauer1 > $dauer2">
            videoPlayStop('video1','video2',
              <xsl:value-of select="$begin1"/>,
              <xsl:value-of select="$dauer1"/>);
            videoPlayLoop('video2',<xsl:value-of select="$begin2"/>,
              <xsl:value-of select="$dauer2"/>,true)
          </xsl:when>
          <xsl:otherwise>
            videoPlayLoop('video1',<xsl:value-of select="$begin1"/>,
              <xsl:value-of select="$dauer1"/>,false);
          </xsl:otherwise>
        </xsl:choose>
      </xsl:when>
    </xsl:choose>
  </xsl:attribute>

```

```

        videoPlayLoop('video2', <xsl:value-of select="$begin2"/>,
                    <xsl:value-of select="$dauer2"/>, false)
    </xsl:otherwise>
</xsl:choose>

<xsl:when test="$variante='3'">
    <xsl:choose>
        <xsl:when test="$dauer1 > $dauer2">
            videoPlayLoop('video1', <xsl:value-of select="$begin1"/>,
                        <xsl:value-of select="$dauer1"/>, false)
            videoDelay('video2', 3000, <xsl:value-of select="$begin2"/>,
                    <xsl:value-of select="$dauer2"/>)
        </xsl:when>
        <xsl:otherwise>
            videoDelay('video1', 3000, <xsl:value-of select="$begin1"/>,
                    <xsl:value-of select="$dauer1"/>)
            videoPlayLoop('video2', <xsl:value-of select="$begin2"/>,
                    <xsl:value-of select="$dauer2"/>, false)
        </xsl:otherwise>
    </xsl:choose>
</xsl:when>

<xsl:when test="$variante='4'">
    <xsl:choose>
        <xsl:when test="$dauer1 > $dauer2">
            videoPlayLoop('video1', <xsl:value-of select="$begin1"/>,
                        <xsl:value-of select="$dauer1"/>, false)
            videoDelay('video2', <xsl:value-of select="$diff"/>,
                    <xsl:value-of select="$begin2"/>,
                    <xsl:value-of select="$dauer2"/>)
        </xsl:when>
        <xsl:otherwise>
            videoDelay('video1', <xsl:value-of select="abs($diff)"/>,
                    <xsl:value-of select="$begin1"/>,
                    <xsl:value-of select="$dauer1"/>)
            videoPlayLoop('video2', <xsl:value-of select="$begin2"/>,
                    <xsl:value-of select="$dauer2"/>, false)
        </xsl:otherwise>
    </xsl:choose>
</xsl:when>
</xsl:choose>
</xsl:attribute>

```

Code 4.33: Ausschnitt aus video\_synchron.xsl

Abb. 4.16 zeigt einen Screenshot der Realisierung in HTML.



Abb. 4.16: Screenshot des Genre-Experimentes »Produktion und Test«

#### 4.8 Modifikation der Alternanz und der Kohärenzrelation

Schon geringfügige Modifikationen des Bildmaterials oder die Bereitstellung nur eines Teils desselben können zur Verfälschung der ursprünglichen Aussage oder sogar zu einer gänzlich anderen Aussage führen. Für einzelne *Content Portions* ist das trivial. In diesem Genre-Experiment soll auf Basis des vorhandenen Materials nicht die Aussage einer einzelnen *Content Portion*, sondern die Kohärenzrelation der Alternanz so modifiziert werden, dass zwischen den alternanten Segmenten eine tatsächlich andere Kohärenzrelation als die des Originalvideos beobachterseits etabliert werden muss. Theoretisch kann man sich nun beliebig viele Kohärenzrelationen denken, jedoch werden auf Basis des vorhandenen Materials nur einige davon wirklich realisierbar sein. Für die Planung muss daher der Inhalt des Beispielvideos nochmals genauer betrachtet werden.

Im Kern geht es im Beispielvideo ja um die Darstellung des Produktionsprozesses eines Snowboards. Werden die Werkstatt-Segmente in ihrer chronologischen Reihenfolge betrachtet, sind diese auch ohne die Berg-Segmente verständlich (s. dazu auch 4.9). Johanna ist im Originalvideo eine Identifikationsfigur für die kindlichen und jugendlichen Beobachter. Sie testet das teilweise fertiggestellte Snowboard, allerdings hat ihre »Tätigkeit« keinen Einfluss auf die Fertigstellung des Snowboards. Ohne Sachkompetenz sollte man die Inhalte der Werkstatt-Segmente unverändert beibehalten und in diese nicht eingreifen. Modifikationen an den Berg-Segmenten und der Rolle von Johanna sind dagegen leichter vorzunehmen, wodurch sich dann auch die Kohärenzre-

lation zwischen den Berg- und Werkstatt-Segmenten ändern kann. Nach Durchsicht des Materials ergibt sich folgende Idee für eine »neue« Rolle von Johanna:

*»Johanna beobachtet die Snowboardproduktion und wartet ungeduldig auf die Fertigstellung des Snowboards.«*

Für die wartende Johanna kann das Videosegment von 0:09 bis 0:17 Minuten verwendet werden. Das ungeduldige Warten kann dann dadurch signalisiert werden, dass dieses Segment in mehr oder weniger regelmäßigen Abständen eingefügt wird.

Diese Idee wird in Form einer narrativen Themenentfaltung umgesetzt. Abb. 4.17 zeigt eine Skizze des geplanten »Drehbuchs«. Das Video beginnt mit einer Situierung, die Johanna als Person und Wartende einführt. Wie im Originalvideo soll hierzu das erste Segment mit Johanna vollständig oder zumindest teilweise gezeigt werden. Ein erstes Auftreten von Johanna zwischen Arbeitsstation 1 und Arbeitsstation 2 wäre dagegen verwirrend, der Beobachter würde sich fragen: »Wer ist das? Was soll das?« Die Repräsentation besteht in der Darstellung des Produktionsprozesses mit den Einschüben der »wartenden Johanna« und zwar jeweils zwischen den Wechseln von einer Arbeitsstation zur nächsten. Als Resümee soll am Ende des Videos dem Beobachter auch klar werden, dass das Snowboard nun gebrauchsfertig ist. [Branigan 1992, S. 20] führt dazu aus:

*»A narrative ends when its cause and effects chains are judged to be totally delineated. There is a reversibility in that the ending situation can be traced back to the beginning; or, to state it another way, the ending is seemingly entailed by the beginning.«*

In diesem Genre-Experiment bietet es sich somit an, die Schlusszene des Originalvideos zu zeigen: Johanna fährt mit dem Snowboard den Berg hinab.

Zur Information der Zuschauer über den Fortschritt des Herstellungsprozesses wird während des Abspielens eines Werkstatt-Segmentes die aktuelle und die letzte Station angezeigt.

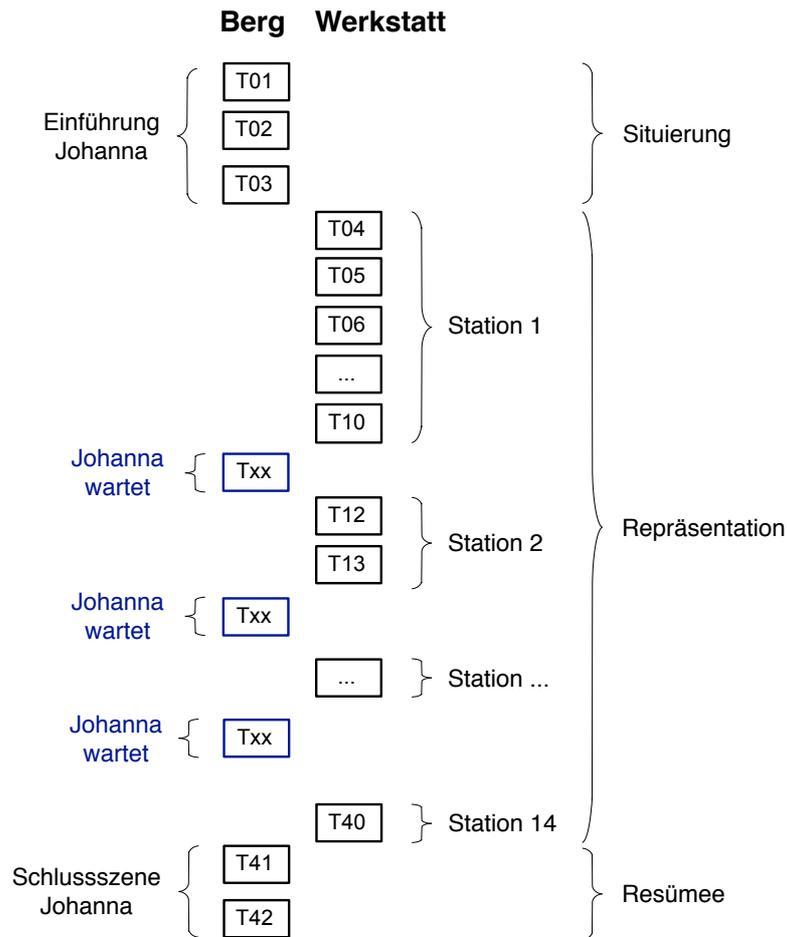


Abb. 4.17: »Drehbuch« zu Modifikation der Alternanz und der Kohärenzrelation

### Erfüllung der Randbedingungen

Die Werkstatt-Segmente im erzeugten Video werden in gleicher Reihenfolge wie im Originalvideo gezeigt, daher sind (S1) und (S4) erfüllt. Die »wartende Johanna« wird jeweils zwischen den Stationswechseln eingeblendet, daher werden die Arbeitsschritte an einer Station (syntagmatisch handelt es sich hierbei um Planszenen, Szenen oder Sequenzen) nicht unterbrochen und (S2) und (S3) sind erfüllt. Von den Beobachtern können die Kohärenzrelationen

$R(\text{Johanna, Snowboard})$ : »Johanna wartet auf das Snowboard.«

$R'(\text{Johanna, Snowboard})$ : »Das Snowboard wird von Johanna erwartet.«

konzeptionalisiert werden, so dass nach Randbedingung (S5) eine Alternanz etabliert werden darf.

Die Randbedingungen (P1) – (P4) sind nicht relevant, da keine *Content Portions* parallel angeordnet werden.

### Technische Umsetzung

Da keine »starre« Lösung, die nur zum Video »Snowboard« passt, realisiert werden soll, werden alle Informationen, die spezifisch für das jeweilige Beispielvideo sind, in einer Konfigurationsda-

tei notiert. Es wird hierbei davon ausgegangen, dass für die Repräsentation ein Itinerar vorliegt und die Segmente für Situierung, alternantes Segment und Resümee jeweils durch die Angabe eines Zeitintervalls des Originalvideos identifiziert werden. Die Konfigurationsdatei basiert auf der DTD `config_alternation.dtd` (s. Code 4.34).

```

<!ENTITY % attr.Clip "clipBegin CDATA #REQUIRED
                    clipEnd CDATA #REQUIRED
                    sound (0|1) #REQUIRED">

<!ELEMENT presentation (transcription,syntagmaticAnalysis,introduction,
                        itinerary,alternanteSegment,resume)>
<!ELEMENT transcription EMPTY>
<!ELEMENT syntagmaticAnalysis EMPTY>
<!ELEMENT introduction EMPTY>
<!ELEMENT itinerary EMPTY>
<!ELEMENT alternanteSegment EMPTY>
<!ELEMENT resume EMPTY>

<!ATTLIST transcription file NMTOKEN #REQUIRED>
<!ATTLIST syntagmaticAnalysis file NMTOKEN #REQUIRED>
<!ATTLIST introduction %attr.Clip;>
<!ATTLIST itinerary ref-id NMTOKEN #REQUIRED>
<!ATTLIST alternanteSegment %attr.Clip;>
<!ATTLIST resume %attr.Clip;>

```

Code 4.34: DTD der Konfigurationsdatei zu »Modifikation der Alternanz und der Kohärenzrelation«

Die Elemente `<transcription>` und `<syntagmaticAnalysis>` enthalten, jeweils als Wert des Attributes `file`, den Dateinamen des Transkriptionsprotokolls bzw. der syntagmatischen Analyse. Die Beginn- und Endezeit des Videosegmentes zur Situierung, des alternanten Segmentes bzw. des Resümees werden als Werte der Attribute `clipBegin` und `clipEnd` zu den Elementen `<introduction>`, `<alternanteSegment>` bzw. `<resume>` hinzugefügt. Diese Elemente besitzen zudem das Attribut `sound` für die Angabe, ob das jeweilige Videosegment mit oder ohne Ton (Attributwert 1 oder 0) abgespielt werden soll. Das Element `<itinerary>` enthält als Wert des Attributes `ref-id` die `id` des Itinerars aus der syntagmatischen Analyse, welches die Basis der Repräsentation darstellt. Code 4.35 zeigt die Konfigurationsdatei zum Video »Snowboard« (Datei `config_snowboard.xml`).

```

<?xml version="1.0" encoding="UTF-8"?>
<!-- Konfigurationsdatei zu 'Modifikation der Alternanz und der Kohärenzrelation'
      für das Video Snowboard -->
<!DOCTYPE presentation SYSTEM "../dtd/config_alternation.dtd">
<presentation>
  <transcription file="snowboard.xml"/>
  <syntagmaticAnalysis file="syntagmatic_analysis.xml"/>
  <introduction clipBegin="0:00" clipEnd="0:29.6" sound="1"/>
  <itinerary ref-id="SW"/>
  <alternanteSegment clipBegin="0:09" clipEnd="0:17" sound="0"/>
  <resume clipBegin="6:54" clipEnd="7:39" sound="1"/>
</presentation>

```

Code 4.35 Konfigurationsdatei zu »Modifikation der Alternanz und der Kohärenzrelation«

Diese Konfigurationsdatei ist Eingabedokument der XSLT-Transformation mittels des Stylesheets `changeAlternation_html.xsl` für die HTML-Ausgabe bzw. `changeAlternation_smil.xsl` (Code 4.37) für die SMIL-Ausgabe.

Im Stylesheet werden zuerst die notwendigen Informationen des Itinerars (erste und letzte Einstellung) berechnet. Dazu dient das benannte Template `itinerar` (s. Code 4.36, abgespeichert in der Datei `itinerar_gen.xsl`). Das Template erhält als Übergabeparameter die Referenz auf den Dokumentbaum der syntagmatischen Analyse und die `id` eines Elementes. Beginnend mit diesem Element wird (durch weitere Template-Regeln) der syntagmatische Baum rekursiv durchlaufen. Da nach Randbedingung (S2) bzw. (S3) Szenen bzw. Sequenzen im Layout nicht auseinandergerissen werden sollen, wird für diese syntagmatischen Klassen ein leeres Element `<segment>` notiert und die erste und letzte Einstellung als Wert des Attributes `begin` bzw. `end` gespeichert. Da auch Planszenen in analoger Weise in das Itinear übernommen werden sollen, enthält das Prädikat der Template-Regel auch die oder-Verknüpfung `»or @class='scene-shot'«`.<sup>1</sup> Bei Planszenen ist natürlich die »erste« und »letzte« Einstellung identisch.

```
<xsl:template name="itinerar">
  <xsl:param name="id"/>          <!-- Startknoten -->
  <xsl:param name="syntagmatic"/> <!-- Wurzelknoten der syntagmatischen Analyse -->
  <xsl:apply-templates select="$syntagmatic//segment[@id=$id]"/>
</xsl:template>

<!-- liegt ein Itinearar vor, mach mit den Kindelementen weiter -->
<xsl:template match="segment[@class='itinerary']">
  <itinerar>
    <xsl:apply-templates/>
  </itinerar>
</xsl:template>

<xsl:template match="ref">
  <xsl:variable name="r" select="@xref"/>
  <!-- mach mit dem Segment weiter, dessen id gleich dem Attributwert von xref ist -->
  <xsl:apply-templates select="//segment[@id=$r]"/>
</xsl:template>

<!-- Szenen und Sequenzen (und auch Planszenen) werden nicht auseinandergerissen,
daher erste und letzte Einstellung speichern -->
<xsl:template match="segment[@class='scene' or @class='sequence'
                             or @class='scene-shot']">
  <segment begin="{./ref[1]/@xref}" end="{./ref[position()=last()]/@xref}"/>
</xsl:template>
```

Code 4.36: Benanntes Template `itinerar` und rekursives Durchlaufen der syntagmatischen Analyse

Code 4.37 zeigt einen Ausschnitt aus `change_alternation_smil.xsl`. Das benannte Template `itinerar` wird aufgerufen und das Ergebnis in der gleichnamigen Variablen `itinerar` gespeichert. Aktuelle Parameter sind das Dokument der syntagmatischen Analyse sowie die `id` des Itinerars aus der Konfigurationsdatei (gespeichert im Attribut `ref-id` des Elementes `<itinerary>`). Die Ausgabe wird nun entsprechend der Planung im »Drehbuch« (s. Abb. 4.17) realisiert. Zur Erzeu-

<sup>1</sup> Alternativ könnte für Planszenen eine eigene Template-Regel definiert werden.

gung der Alternanz dient die `<xsl:for-each select="$itinerar/itinerar/segment">`-Schleife, in der für die Stationen des Itinerars im Wechsel mit dem alternanten Segment die entsprechenden `<video>`-Elemente in den `<seq>`-Container eingefügt werden.

Die HTML-Version (`changeAlternation_html.xsl`) unterscheidet sich von der SMIL-Version dahingehend, dass für die Videosegmente Instanzen des (selbstdefinierten) JavaScript-Objektyps `Video` im Array `source` gespeichert und die Funktion `videoSeq(source)` beim Laden der HTML-Seite aufgerufen wird (s. auch 4.5.4, 4.9).

```

<!-- Werte des alternanten Segmentes als globale Variable -->
<xsl:variable name="as_clipBegin" select="/presentation/alternanteSegment/@clipBegin"/>
<xsl:variable name="as_clipEnd"   select="/presentation/alternanteSegment/@clipEnd"/>
<xsl:variable name="as_sound"    select="/presentation/alternanteSegment/@sound"/>

<!-- Aufruf des benannten Template itinerar, Speicherung des Ergebnisses in der Variablen -->
<xsl:variable name="itinerar">
  <xsl:call-template name="itinerar">
    <xsl:with-param name="syntagmatic" select="$syntagmaticAnalysis"/>
    <xsl:with-param name="id" select="itinerary/@ref-id"/>
  </xsl:call-template>
</xsl:variable>
...
<seq>
  <!-- Situierung-->
  <video region="vid1" src="{vname1}"
    clipBegin="{introduction/@clipBegin}" clipEnd="{introduction/@clipEnd}">
    <xsl:if test="introduction/@sound='0'">
      <xsl:attribute name="soundLevel">0</xsl:attribute>
    </xsl:if>
  </video>

  <!-- für jedes Station im Itinerar -->
  <xsl:for-each select="$itinerar/itinerar/segment">
    <xsl:variable name="begin"   select="@begin"/>
    <xsl:variable name="end"     select="@end"/>
    <xsl:variable name="clipBegin" select="$units/unit[@id=$begin]/begin"/>
    <xsl:variable name="clipEnd"  select="$units/unit[@id=$end]/end"/>
    <par>
      <video region="vid1" src="video/snowboard.mp4"
        clipBegin="{clipBegin}" clipEnd="{clipEnd}" /> <!-- immer mit Ton -->
      <!-- Anzeige der Station -->
      <smilText region="text" textFontSize="14px">Station:
        <xsl:value-of select="$units/unit[@id=$begin]/location/loc_detail"/> /
        <xsl:value-of select="$units/unit[@id=$last]/location/loc_detail"/>
      </smilText>
    </par>

    <!-- alternante Segmente -->
    <xsl:if test="position() != last()">
      <video region="vid1" src="video/snowboard.mp4"
        clipBegin="{as_clipBegin}" clipEnd="{as_clipEnd}">
        <xsl:if test="$as_sound='0'">
          <xsl:attribute name="soundLevel">0</xsl:attribute>
        </xsl:if>
      </video>
    </xsl:if>
  </xsl:for-each>

  <!-- Resumee -->
  <video region="vid1" src="video/snowboard.mp4"
    clipBegin="{resumee/@clipBegin}" clipEnd="{resumee/@clipEnd}">

```

```

<xsl:if test="resume/@sound='0'">
  <xsl:attribute name="soundLevel">0</xsl:attribute>
</xsl:if>
</video>
</seq>

```

Code 4.37: Ausschnitt aus *change\_alteration\_smil.xml*

## 4.9 Wegfall der Kohärenzrelation

Im Genre-Experiment in Abschnitt 4.8 wurde die Kohärenzrelation im layoutierten Dokument modifiziert. Was heißt es nun, die Möglichkeit einer Kohärenzrelation beobachterseits möglichst auszuschließen?

Entfällt die Kohärenzrelation, bedeutet dies allgemein, dass das Geschehen in den Segmenten der verschiedenen diegetischen (Haupt-)Orte beobachterseits in keinem inhaltlichen Bezug zueinander steht. In einer solchen Situation greifen nun speziell die Randbedingungen für die Genre-Experimente: Nach Randbedingung (S5) darf dann keine Alternanz etabliert werden. Dies hat zur Konsequenz, dass ein Video bei Wegfall der Kohärenzrelation in zwei oder mehr Einzelvideos, in denen jeweils die Einstellungen eines diegetischen (Haupt-)Ortes in einer kontinuierlichen Folge gezeigt werden, »zerfällt«. Für das Beispielvideo »Snowboard« erhält man somit zwei Videos: In einem Video werden die Berg-Segmente, im anderen die Werkstatt-Segmente jeweils ohne Unterbrechung gezeigt (s. Abb. 4.18). Die Randbedingungen greifen nun *ex negativo* und verhindern einen ganzen Typ von Layouts.

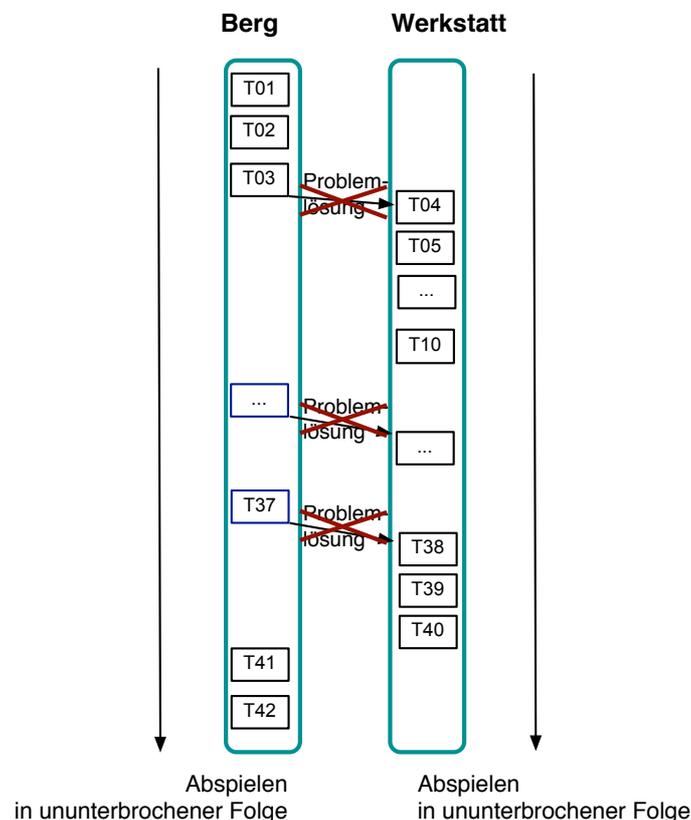


Abb. 4.18: Wegfall der Kohärenzrelation

Abb. 4.19 zeigt das geplante Screendesign für die Realisierung des experimentellen Genres »Einzelne Orte«. Die Navigationslinks im linken Bereich ermöglichen die Auswahl der abzuspielenden Einstellungsfolge, im unteren Bereich erhält der Beobachter die Information, im wievielten Segment des gezeigten Teilvideos er sich gerade befindet.

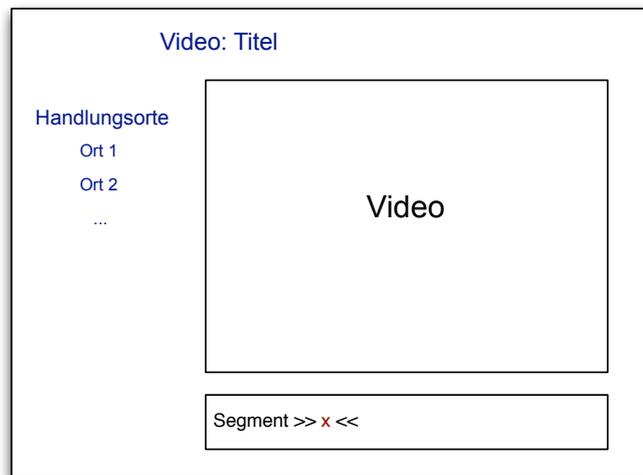


Abb. 4.19: Screendesign für das Genre »Einzelne Orte«

### Erfüllung der Randbedingungen

Da alle Einstellungen eines diegetischen (Haupt-)Ortes in ihrer chronologischen Reihenfolge und fortlaufend gezeigt werden, sind die Randbedingungen (S1), (S2), (S3) und für das Werkstatt-Segment auch (S4) erfüllt. (S5) greift wie erörtert. Die Randbedingungen für die parallele Anordnung von *Content Portions* sind nicht von Bedeutung, da der angezeigte Text lediglich der Navigation bzw. Information dient.

### Technische Umsetzung in HTML

Code 4.38 zeigt einen Ausschnitt aus dem XSLT-Stylesheet `separate_location_html.xsl`. Für die HTML-Lösung wird im JavaScript-Bereich das Array `source` deklariert. Es ist ein zweidimensionales Array, wobei eine Zeile jeweils die Informationen eines Handlungsortes enthält. Die einzelnen Elemente enthalten jeweils eine Instanz des (selbstdefinierten) JavaScript-Objektyps `Video` (s. 6.3.2) mit den Informationen über ein Video-Segment (Eigenschaften `id`, `begin`, `dauer`, `sound`, `txt`). Tab. 4.4 zeigt schematisch das Ergebnis für das Video »Snowboard«.

source	0	1	2	3	4
0 (=Berg)	Berg1	Berg2	Berg3	Berg4	Berg5
1 (=Werkstatt)	Werkstatt1	Werkstatt2	Werkstatt3	Werkstatt4	

Tab. 4.4: Zweidimensionales Array `source` im Beispielvideo »Snowboard«

Um das Array `source` mit Werten zu füllen, werden daher zwei verschachtelte `<xsl:for-each>`-Schleifen verwendet. Die äußere Schleife iteriert über die verschiedenen Orte (Element `<loc-`

\_name>), die innere über die Segmente des jeweiligen Ortes (vorab berechnet durch das benannte Template segment, s. 4.5.1).

Im <body>-Bereich werden der Navigationslinks durch eine <xsl:for-each>-Schleife über die verschiedenen Orte erzeugt. Jeder Link erhält den Eventhandler onClick mit dem Attributwert videoSeq(source[{\$dim}]), so dass bei Aktivierung des Links die Videosegmente des ausgewählten Ortes in einer kontinuierlichen Folge abgespielt werden.

```
<script>
  <!-- Array Deklaration -->
  var source = new Array();

  <!-- Iteration über die verschiedenen Orte -->
  <xsl:for-each select="distinct-values(/movie/units/unit/location/loc_name)">
    <xsl:variable name="loc" select="."/>
    <xsl:variable name="dim" select="position()-1"/>

    source[<xsl:value-of select="$dim"/>]=new Array(); <!-- für jeden Ort eine Zeile -->

    <!-- für jedes Segment des aktuellen Ortes -->
    <xsl:for-each select="$segTree/segments/segment[@location=$loc]">

      <!-- An jeder Position ein selbstdefiniertes Objekt Video
           mit Eigenschaften id, begin, dauer, sound, txt -->
      source[<xsl:value-of select="$dim"/>][source[<xsl:value-of select="$dim"/>].length]
        =new Video("video1",
          <xsl:value-of select="fu:inSec(@clipBegin)"/>,
          <xsl:value-of select="fu:dauer(@clipBegin,@clipEnd)*1000"/>,1,
          'Segment >> <xsl:value-of select="position()"/> &lt;&lt;');

    </xsl:for-each>
  </xsl:for-each>
</script>

<body>
  <!-- Navigation -->
  <div id="navi">
    <h3>Handlungsorte</h3>
    <xsl:for-each select="distinct-values(/movie/units/unit/location/loc_name)">
      <xsl:variable name="loc" select="."/>
      <xsl:variable name="dim" select="position()-1"/>
      <!-- Funktion videoSeq erhält als Parameter das Array des aktuellen Ortes -->
      <p class="navlink" id="a{$dim}"
        onclick="i=0; clearTimeout(timer); videoSeq(source[{$dim}]);
          toggleClass('a{$dim}','white','yellow')"
        <xsl:value-of select="$loc"/>
      </p>
    </xsl:for-each>
  </div>
  ...
</body>
```

Code 4.38: Ausschnitt aus separate\_location\_html.xsl

### Technische Umsetzung in SMIL

Code 4.39 zeigt einen Ausschnitt aus separate\_location\_smil.xsl. Eine <xsl:for-each>-Schleife, die über die verschiedenen Orte iteriert, erzeugt die Navigationslinks. Das Ziel eines Links ergibt sich jeweils aus dem Ausdruck #par{\$loc}, wobei die Variable loc den aktuellen Ort enthält. Zudem verfügen die Links über eine eindeutige xml:id mit dem Wert a{\$loc}.

Da eine Videosequenz erst abgespielt werden soll, wenn der Beobachter auf einen Navigationslink klickt, wird ein `<excl>`-Container notiert. In diesem wird für jeden Ort durch eine `<xsl:for-each>`-Schleife ein `<par>`-Container erzeugt. Jeder `<par>`-Container verfügt über das Attribut `begin="a{$loc}.activateEvent"`, daher wird er bei einem Mausklick auf das Element mit dieser `xml:id` aktiviert (ereignisbasierte Synchronisation).

Zudem enthält jeder `<par>`-Container einen (den eigenen!) Navigationslink in der Farbe gelb, der den beim Start angezeigten Link temporär überschreibt, um so den Ort der aktuell abgespielten Videosequenz farblich zu kennzeichnen. Für die Videosequenz wird der Zeitcontainer `<seq>` benötigt. Durch eine `<xsl:for-each>`-Schleife wird in diesem Zeitcontainer für jedes Segment des aktuellen Ortes ein `<video>`-Element und parallel dazu ein `<smilText>`-Element mit der Benutzerinformation notiert. Das Attribut `end="{concat($loc,position())}.end"` des `<smilText>`-Elementes bewirkt, dass der Text für die Abspieldauer des Videosegmentes angezeigt wird.

```

<!-- Navigationsbereich -->
<smilText region="zwi" textColor="#ff0000" textFontSize="14px">Handlungsorte </smilText>

<par>
  <xsl:for-each select="distinct-values(/movie/units/unit/location/loc_name)">
    <xsl:variable name="loc" select="."/>

    <!-- Navigationslinks -->
    <a href="#par{$loc}" xml:id="a{$loc}">
      <smilText region="r{$loc}" textColor="#ffffff"><xsl:value-of select="$loc"/></smilText>
    </a>
  </xsl:for-each>
</par>

<!-- Container für die Videosequenzen -->
<excl dur="indefinite">
  <!-- Iteration über die verschiedenen Orte -->
  <xsl:for-each select="distinct-values(/movie/units/unit/location/loc_name)">
    <xsl:variable name="loc" select="."/>

    <!-- Container für Videosequenz - wird abgespielt, wenn auf den Link geklickt wird-->
    <par xml:id="par{$loc}" dur="indefinite" begin="a{$loc}.activateEvent">

      <!-- Navigationslink in gelb darstellen -->
      <a href="#par{$loc}" xml:id="a{$loc}">
        <smilText region="r{$loc}" textColor="#ffff00"><xsl:value-of select="$loc"/></smilText>
      </a>
      <!-- Videosequenz -->
      <seq>
        <!-- für jedes Segment des aktuellen Ortes -->
        <xsl:for-each select="$segTree/segments/segment[@location=$loc]">
          <par>
            <video region="vid1" src="{vname1}"
              xml:id="{concat($loc,position())}"
              clipBegin="{@clipBegin}" clipEnd="{@clipEnd}" fill="freeze"/>
            <!-- für jedes Segment des aktuellen Ortes -->
            <smilText region="text" end="{concat($loc,position())}.end" fill="remove">
              Segment >> <xsl:value-of select="position()"/>
            </smilText>
          </par>
        </xsl:for-each>
      </seq>
    </par>
  </xsl:for-each>
</excl>

```

```
</xsl:for-each>
</excl>
```

Code 4.39: Ausschnitt aus *separate\_location\_smil.xml*

Abb. 4.20 zeigt einen Screenshot der erzeugten HMTL-Lösung.

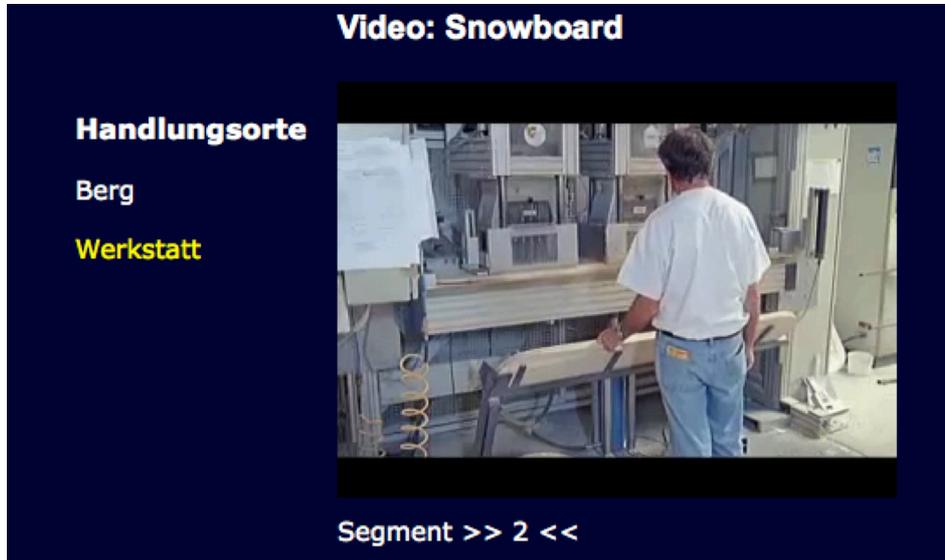


Abb. 4.20: Screenshot des Genre-Experimentes »Einzelne Orte«

### Wirkung auf die Beobachter

Beim Betrachten der einzelnen Videos wird deutlich, wie durch den Wegfall der Kohärenzrelation mögliche Beobachterinterpretationen beeinträchtigt werden.

Das Teilvideo »Werkstatt« zeigt den kompletten Produktionsprozess ohne Unterbrechung. Hier entsteht kein Störfühl. Jedoch wirkt dies insgesamt langatmig und langweilig, da die dramaturgische Anreicherung des *Accomplishments* durch den Wegfall der Kopplung der Fehlversuche Johannas mit dem jeweils erreichten Produktionsstand entfällt.

Beim Teilvideo »Berg« ist die Situation eine andere. Nach jedem Segment hat ein Beobachter das Gefühl eines »Sprungs« in der Handlung und das obwohl sowohl die handelnde Person als auch der Handlungsort identisch sind. Ursache hierfür sind die Veränderungen am Objekt »Snowboard«, wie das Petri-Netz in Abb. 4.21 zeigt. In einem Werkstatt-Segment sorgt jede Handlung dafür, dass das Snowboard von einem Ausgangszustand in den nächsten Zustand wechselt (s. auch Abb. 2.30) und dieser wiederum Ausgangszustand des nächsten Handlungsschrittes ist. In einem Berg-Segment wird dagegen der Zustand des Snowboards nicht verändert. Das nun folgende Berg-Segment zeigt den zeitlich folgenden Bearbeitungszustand des Snowboards, was von einem Beobachter, wie oben erwähnt, als »Sprung« empfunden wird, da die Erklärung, wie und warum sich das Snowboard verändert wurde, fehlt.

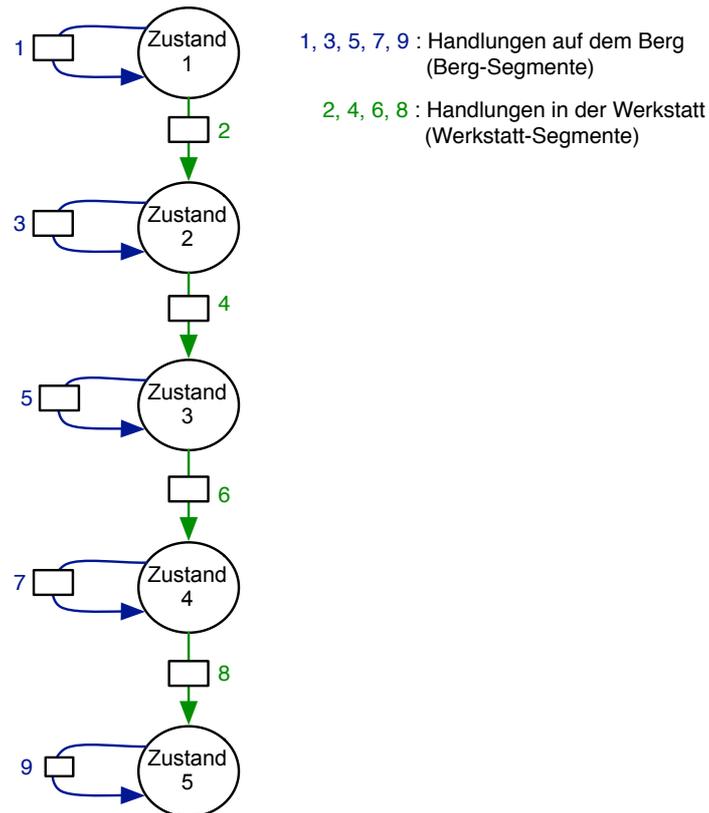


Abb. 4.21: Petri-Netz zum Beispielvideo »Snowboard«

Fokussiert ein Beobachter sich auf die einzelnen Tests, so wurde in 3.7 vermutet, dass jeder Test des Snowboards als *Activity* konzeptionalisiert werden könnte, wenn das Dokument nicht in der layoutierten Ordnung, sondern die Handlungsstränge getrennt betrachtet werden. Betrachtet man den Handlungsstrang der »Tests des Snowboards« mit eingeschaltetem Ton, gibt der Sprechertext jedoch eindeutige Hinweise auf einen natürlichen Endpunkt eines jeden Tests, wie Tab. 4.5 zeigt. Das heißt, in diesem Fall kann jeder Test weiterhin als *Accomplishment* konzeptionalisiert werden.

Test 1:	»[...] Aber funktionieren tut das nicht, auch wenn sie sich quer draufstellt.«
Test 2:	»[...] Aber so der wahre Jakob, ist das auch noch nicht.«
Test 3:	»[...] Na, auch mit Anstrengung, höchstens ein auf den Po fallen.«
Test 4:	»[...] Ja, geht schon, aber doch noch nicht so ganz. Irgendwas fehlt noch.«
Test 5:	»So Johanna, jetzt aber los! [...]

Tab. 4.5: Sprechertexte in »Tests des Snowboards«

Betrachtet man die Einstellungsfolge jedoch ohne Ton, sind im filmischen Bild Inszenierungen eines natürlichen Endpunktes in den Tests 1, 2, 3 und 4 nicht wirklich erkennbar. Insbesondere haben Test 1 (zweite Hälfte), Test 2, 3 und 4 einen quasi identischen Inhalt: Johanna hält das Brett, sie legt das Brett auf den Boden, stellt sich drauf, versucht zu rutschen, fällt hin. Eine Konzeptionalisierung als *Activity* »ein bisschen auf dem Snowboard rumrutschen und schauen, ob es funktioniert« ist möglich.

#### 4.10 Zusammenfassung

Mit den vorliegenden *Content Portions* des Videos »Snowboard« und unter Einhaltung der in 4.2.2 formulierten Randbedingungen wurden die folgenden Genre-Experimente umgesetzt:

Als Text-Bild-Kombination wurden zwei Formen einer Bildergeschichte realisiert: Eine »einfache« Bildergeschichte, die je Einstellung ein Bild und ggf. den zugehörigen (Sprecher-)Text zeigt, und eine detaillierte Bildergeschichte, die je Textsegment eine Seite mit Bild und ggf. Text und einer Visualisierung der vorliegenden Text-Bild-Relation durch ein Icon präsentiert.

Im Genre-Experiment »Arbeitsfortschritt« wird ein Werkstatt-Segment gemeinsam mit zwei Bildern präsentiert, wobei das eine Bild Johanna mit dem Snowboard vor dem gezeigten Produktionsschritt, das andere Johanna mit dem Snowboard nach dem gezeigten Produktionsschritt zeigt. Das gleichzeitige Anzeigen dieser *Content Portions* betont die vorliegenden Kohärenzrelationen.

Ein Betonung des kohärent Relationierten kann erreicht werden, wenn die relationierten *Content Portions* gemeinsam präsentiert werden. Dies wird im Genre-Experiment »Produktion und Test« umgesetzt. Hier wird je ein Werkstatt-Segment und gleichzeitig das im ursprünglichen Layout folgende Berg-Segment abgespielt. Ein Beobachter sieht also einen Produktionsschritt und gleichzeitig den Test mit dem durch diesen Produktionsschritt soweit fertiggestellten Snowboard.

Im Genre-Experiment »Johanna beobachtet die Snowboard-Produktion und wartet auf ihr Snowboard« wird die im Beispielvideo vorliegende Alternanz so modifiziert, dass sich die Rolle von Johanna und die beobachterseits konzeptionalisierte Kohärenzrelation ändert.

Bei Wegfall (bzw. nicht Vorhandensein) einer Kohärenzrelation darf nach den in 4.2.2 formulierten Randbedingungen keine Alternanz etabliert werden. Das Genre-Experiment »Einzelne Orte« setzt diese Situation um. Hier werden nun für das Beispielvideo »Snowboard« die Berg- bzw. Werkstatt-Segmente in kontinuierlicher Folge abgespielt.

## 5 Literatur

[Allen 1983]

Allen, J. F.: Maintaining Knowledge About Temporal Intervals. *Commun. ACM* 26, 11 (1983), S. 832–843.

[Allen, Hayes 1985]

Allen, J. F., Hayes, P. J.: A Common-Sense Theory of Time. *Proceedings of AAAI-85*, (1985), S. 528–531.

[Ast 2002]

Ast, M. S.: *Geschichte der narrativen Filmmontage: theoretische Grundlagen und ausgewählte Beispiele*. Tectum-Verlag, Marburg, 2002.

[Bateman 2007]

Bateman, J. A.: Towards a grande paradigmatic of film: Christian Metz reloaded. *Semiotica* 167, (2007), S. 13–64.

[Bateman 2013]

Bateman, J. A.: Multimodal analysis of film within the gem framework. *Ilha do Desterro A Journal of English Language, Literatures in English and Cultural Studies* N° 64 (2013), S. 49–84.

[Bateman 2014]

Bateman, J. A.: *Text and Image: A critical introduction to the visual/verbal divide*. Routledge Chapman & Hall, London; New York, 2014.

[Bateman, Schmidt 2011]

Bateman, J. A., Schmidt, K.-H.: *Multimodal film analysis: How Films mean*. Routledge Chapman & Hall, New York, 2011.

[Bennet, Partee 1972]

Bennett, M., Partee, B. H.: *Toward the logic of tense and aspect in English*. Indiana University, Bloomington, 1972.

[Bordwell 1985]

Bordwell, D.: *Narration in the fiction film*. Methuen, London, 1985.

[Bordwell, Thompson 2010]

Bordwell, D., Thompson, K.: *Film art: an introduction*. McGraw-Hill, New York, 2010.

[Branigan 1992]

Branigan, E.: *Narrative comprehension and film*. Routledge, London u.a., 1992.

[Branigan 1984]

Branigan, E.: *Point of view in the cinema: a theory of narration and subjectivity in classical film*. Mouton, Berlin u.a., 1984.

[Brinker 2010]

Brinker, K.: *Linguistische Textanalyse: eine Einführung in Grundbegriffe und Methoden*. Schmidt, Berlin, 2010.

[Bulterman, Rutledge 2008]

Bulterman, D. C., Rutledge, L.: *SMIL 3.0: Interactive Multimedia for Web, Mobile Devices and Daisy Talking Books*. Springer Berlin, Berlin, 2008.

[Burch 1973]

Burch, N.: *Theory of film practice*. Secker and Warburg, London, 1973.

[Bußman, 2002]

Bußmann, H. (Hrsg.): Lexikon der Sprachwissenschaft. Kröner, Stuttgart, 2002.

[Chatman 1980]

Chatman, S.: What novels can do that films can't (and vice versa). *Critical Inquiry* 7, 1 (1980), S. 121–140.

[Colin 1995]

Colin, M.: The grande syntagmatique revisited. In W. Buckland (Hrsg.): *The Film Spectator: From Sign to Mind*. Amsterdam University Press, Amsterdam, 1995, S. 45–85.

[Davis 1990]

Davis, E. *Representations of commonsense knowledge*. Morgan Kaufmann Publ., San Mateo, Calif., 1990.

[Doane 2002]

Doane, M. A.: *The emergence of cinematic time: modernity, contingency, the archive*. Harvard University Press, Cambridge, Mass. [u.a.], 2002.

[Eder 2002]

Eder, J.: *Dramaturgie des populären Films: Drehbuchpraxis und Filmtheorie*. LIT Verlag, Hamburg u.a., 2007.

[Fillmore 1967]

Fillmore, C. J.: The Case for Case. In: Bach, E., Harms, R.T. (Hrsg.): *Universals in Linguistic Theory*, (1967), New York: Holt, Rinehart and Winston, S. 1–88.

[Finke 2005]

Finke, M.: *Unterstützung des kooperativen Wissenserwerbs durch Hypervideo-Inhalte*. 2005. Dissertation Universität Darmstadt, <http://tuprints.ulb.tu-darmstadt.de/id/eprint/661> [letzter Aufruf: 04.01.2015]

[Garey 1957]

Garey, H. B.: Verbal aspect in French. *Language* 33, (1957), S. 91–110.

[Genette 2010]

Genette, G.: *Die Erzählung*. Fink, Paderborn, 2010.

[Goguen 1999]

Goguen, J.: An introduction to algebraic semiotics, with application to user interface design. In: Nehaniv, C. L. (Hrsg.): *Computation for metaphors, analogy, and agents*. Springer, Berlin u.a., 1999, S. 242–291.

[Goodman 1969]

Goodman, N.: *Languages of Art: An Approach to a Theory of Symbols*. Oxford University Press, London, 1969.

[Goodman 1984]

Goodman, N.: *Weisen der Welterzeugung*. Suhrkamp, Frankfurt am Main, 1984.

[Goodman 1987]

Goodman, N.: *Vom Denken und anderen Dingen*. Suhrkamp, Frankfurt am Main, 1987.

[Goodman, Elgin 1989]

Goodman, N., Elgin, C. Z.: *Revisionen - Philosophie und andere Künste und Wissenschaften*. Suhrkamp Verlag, Frankfurt am Main, 1989.

[Grimm 1996]

Grimm, P.: *Filmnarratologie. Eine Einführung in die Praxis der Interpretation am Beispiel des Werbespots*. Diskurs-Film-Verlag Schaudig und Ledig, München, 1996.

- [Halliday, Matthiessen 2000]  
Halliday, M. A. K., Matthiessen, C. M. I. M.: *Construing experience through meaning: a language-based approach to cognition*. Continuum, London; New York, 2000.
- [Halliday, Matthiessen 2014]  
Halliday, M. A. K., Matthiessen, C. M. I. M.: *Halliday's introduction to functional grammar*. Routledge, Abingdon, 2014.
- [Hickethier 2007]  
Hickethier, K.: *Film- und Fernsehanalyse*. Metzler, Stuttgart u.a., 2007.
- [ISO/IEC 8613-2:1994]  
ISO/IEC. Open Document Architecture Consortium. *The Open Document Architecture*. ISO/IEC 8613-2.
- [ISO/IEC 15938-5:2003]  
ISO/IEC. Information Technology – Multimedia Content Description Interface – Part 5: Multimedia Description Schemes. ISO/IEC 15938-5/2003
- [Kawin 1992]  
Kawin, B.F.: *How movies work*. University of California Press, Berkeley, 1992.
- [Kong 2006]  
Kong, K. C. C.: A taxonomy of the discourse relations between words and visuals. *Information Design Journal* 14, 3 (2006), S. 207–230.
- [Kress, van Leeuwen 1996]  
Kress, G., Van Leeuwen, T.: *Reading images: the grammar of visual design*. Routledge, London u.a., 1996.
- [Krifka 1989a]  
Krifka, M.: *Nominalreferenz und Zeitkonstitution: zur Semantik von Massentermen, Pluraltermen und Aspektklassen*. Fink, München, 1989.
- [Krifka 1989b]  
Krifka, M.: *Nominalreferenz, Zeitkonstitution, Aspekt, Aktionsart: Eine semantische Erklärung ihrer Interaktion*. In: Abraham, W., Janssen, T. (Hrsg.): *Tempus-Aspekt-Modus*. Niemeyer, Tübingen, 1989.
- [Krifka 1992]  
Krifka, M.: *Thematic relations as links between nominal reference and temporal constitution*. In: Sag, I. A., Szabolcsi, A.: *Lexical Matters*, Center for the Study of Language (CSLI), (1992), S. 29–53.
- [Kuchenbuch 2005]  
Kuchenbuch, T.: *Filmanalyse: Theorien - Methoden - Kritik*. Böhlau, Wien u.a., 2005.
- [Li, Ying 2003]  
Li, S., Ying, M.: *Region Connection Calculus: Its models and composition table*. *Artificial Intelligence* 145, 1–2 (2003), S. 121–146.
- [Mann, Taboada 2014]  
Mann, W. C., Taboada, M.: *Rhetorical Structure Theory. Intro to RST / Rhetorical Structure Theorie*. <http://www.sfu.ca/rst/01intro/intro.html>. [letzter Aufruf: 04.01.2015]
- [Mann, Thompson 1987]  
Mann, W., Thompson, S.: *Rhetorical Structure Theory: A theory of text organization*. In: *Technical Reports ISI/RS-87-190*

- [Martinec, Salway 2005]  
Martinec, R., Salway, A.: A system for image-text relations in new (and old) media. *Visual Communication* 4, 3 (2005), S. 337–371.
- [Martínez, Scheffel 2007]  
Martínez, M., Scheffel, M.: Einführung in die Erzähltheorie. Beck, München, 2007.
- [Martin 1992]  
Martin, J. R.: *English Text: System and Structure*. John Benjamins Publishing Co, Amsterdam, 1992.
- [Metz 1964]  
Metz, C.: Le cinéma: langue ou language? *Communications* 4, (1964), S. 52–90.
- [Metz 1965]  
Metz, C.: Une étape dans la réflexion sur le cinéma. *Critique* 214, (1965), S. 227–248.
- [Metz 1966]  
Metz, C.: La grande syntagmatique du film narratif. *Communications* 8, (1966), S. 120–124.
- [Metz 1972]  
Metz, C.: *Semiologie des Films*. Fink, München, 1972.
- [Moens, Steedman 1988]  
Moens, M., Steedman, M.: Temporal ontology and temporal reference. *Computational linguistics* 14, 2 (1988), S. 15–28.
- [Mollá-Aliod 1997]  
Mollá-Aliod, D.: *Aspectual Composition and Sentence Interpretation: A formal approach*. PhD Thesis 1997.  
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.36.41&rep=rep1&type=pdf>.  
[letzter Aufruf: 04.01.2015]
- [Möller-Naß 1988]  
Möller-Naß, K.-D.: *Filmsprache: eine kritische Theoriegeschichte*. MAkS-Publikationen, Münster, 1988.
- [Monaco 2009]  
Monaco, J.: *Film verstehen: Kunst, Technik, Sprache, Geschichte und Theorie des Films und der Neuen Medien ; mit einer Einführung in Multimedia*. Rowohlt, Reinbek, 2009.
- [Muckenhaupt 1986]  
Muckenhaupt, M.: *Text und Bild: Grundfragen der Beschreibung von Text-Bild-Kommunikation aus sprachwissenschaftlicher Sicht*. Narr, Tübingen, 1986.
- [Pörings, Schmitz 2003]  
Pörings, R., Schmitz, U.: *Sprache und Sprachwissenschaft: eine kognitiv orientierte Einführung*. Narr, Tübingen, 2003.
- [Prince 1973]  
Prince, G.: *A grammar of stories: an introduction*. Mouton, The Hague [u.a.], 1973.
- [Randell et al. 1992]  
Randell, D. A., Cui, Z., Cohn, A. G.: A Spatial Logic based on Regions and Connection. *Proceedings 3rd International Conference on Knowledge Representation and Reasoning*, Morgan-Kaufmann (1992), S. 165–176.
- [Reimer 1991]  
Reimer, U.: *Einführung in die Wissensrepräsentation. Netzartige und schema-basierte Repräsentationsformate*. Teubner, Stuttgart, 1991.

[Reisz, Millar 2010]

Reisz, K., Millar, G.: The technique of film editing. Focal Press, Burlington, MA, 2010.

[Roth 1983]

Roth, L.: Film semiotics, Metz, and Leone's trilogy. Garland Publishing, New York, 1983.

[Sawhney et al. 1996]

Sawhney, N., Balcom, D., Smith, I.: HyperCafe: Narrative and Aesthetic Properties of Hypervideo. Proceedings of the the seventh ACM conference on Hypertext, ACM (1996), S. 1–10.

[Scheffel 2009]

Scheffel, M.: Was heißt (Film-)Erzählen? Exemplarische Überlegungen mit einem Blick auf Schnitzlers Traumnovelle und Stanley Kubricks Eyes Wide Shut. In: Kaul, S., Palmier, J.-P., Skrandies, T. (Hrsg.): Erzählen im Film: Unzuverlässigkeit - Audiovisualität - Musik. Transcript, Bielefeld, 2009, S. 15–31.

[Schmidt 2004]

Schmidt, K.-H.: Zur chronologischen Syntagmatik von Bewegtbilddaten (II): Polyspatiale Alternanz. Kodikas / Kode - Ars Semeiotica Vol. 27/2004 (3-4), (2004), S. 255–283.

[Schmidt 2008]

Schmidt, K.-H.: Zur chronologischen Syntagmatik von Bewegtbilddaten (III): Deskriptive Syntagen. Kodikas / Kode - Ars Semeiotica Vol. 31/2008 (3-4), (2008), S. 217–270.

[Schmidt, Strauch 2002]

Schmidt, K.-H., Strauch, T.: Zur chronologischen Syntagmatik von Bewegtbilddaten. Eine semiologische Reklassifikation der Großen Syntagmatik von Metz (anhand einer Neuanalyse des Spielfilms "Adieu Philippine"). Kodikas / Kode - Ars Semeiotica Vol. 25/2002 (1-2), (2002), S. 65–96.

[Schulz, Hahn 2002]

Schulz, S., Hahn, U.: Necessary Parts and Wholes in Bio-Ontologies. Principles of knowledge representation and reasoning proceedings of the eighth international conference (KR2002), (2002), S. 387–394.

[Singh, Singh 1992]

Singh, M., Singh, M. P. Computing the Temporal Structure of Events in Natural Language. ECAI, (1992), S. 528–532.

[Singh, Singh 1995]

Singh, M., Singh, M. P.: The Temporal Structure of Narratives: A Semantic Approach. Proceedings of the 2nd Conference of the Pacific Association for Computational Linguistics (Pacling), (1995). <http://www.cs.ncsu.edu/faculty/mpsingh/papers/others/pacl-narrative-95.pdf> [letzter Aufruf: 04.01.2015]

[Singh, Singh 1997]

Singh, M., Singh, M. P.: On the Temporal Structure of Events. Proceedings of AAAI-97 Workshop on Spatial and Temporal Reasoning, (1997), S. 49–54.

[Smirnova, Mortelmans 2010]

Smirnova, E., Mortelmans, T.: Funktionale Grammatik: Konzepte und Theorien. de Gruyter, Berlin u.a., 2010.

[Smith 2005]

Smith, B.: The logic of biological classification and the foundations of biomedical ontology. Invited Papers from the 10th International Conference in Logic Methodology and Philosophy of Science, Oviedo, Spain, King's College Publications (2005), S. 505–520.

[Smith, Rosse 2004]

Smith, B., Rosse, C.: The role of foundational relations in the alignment of biomedical ontologies. *Medinfo* 11, (2004), S. 444–448.

[Stöckl 2004]

Stöckl, H.: *Die Sprache im Bild – das Bild in der Sprache: zur Verknüpfung von Sprache und Bild im massenmedialen Text : Konzepte, Theorien, Analysemethoden*. Walter de Gruyter, 2004.

[Stötzel 1990]

Stötzel, D. U.: *Das Magazin „Die Sendung mit der Maus“: Analyse einer Redaktions- und Sendungskonzeption*. Harrassowitz, Wiesbaden, 1990.

[Straßner 2002]

Straßner, E.: *Text-Bild-Kommunikation - Bild-Text-Kommunikation*. Niemeyer, Tübingen, 2002.

[Taboada, Habel 2013]

Taboada, M., Habel, C.: Rhetorical relations in multimodal documents. *Discourse Studies* 15, 1 (2013), S. 65–89.

[van Leeuwen 1991]

van Leeuwen, T.: Conjunctive structure in documentary film and television. *Continuum* 5, 1 (1991), S. 76–114.

[van Leeuwen 2005]

van Leeuwen, T.: *Introducing social semiotics*. Routledge, London u.a., 2005.

[Vendler 1957]

Vendler, Z.: Verbs and Times. *The Philosophical Review* 66, 2 (1957), S. 143–160.

[Vendler 1967]

Vendler, Z.: *Linguistics in philosophy*. Cornell University Press, Ithaca, New York, 1967.

[Winston 1987]

Winston, M. E., Chaffin, R., Herrmann, D.: A taxonomy of part-whole relations. *Cognitive Science* 11, 4 (1987), S. 417–444.

## 6 Anhang

### 6.1 Transkription des Videos »Snowboard«

[http://www.youtube.com/watch?v=xDu\\_-CO6eCE\\_](http://www.youtube.com/watch?v=xDu_-CO6eCE_) [letzter Aufruf: 04.01.2015]

T	Beginn	Bild	Kamera	Aktion /Besonderes	Sprechertext
1	0:00		Einst: T	Johanna geht auf den Berg und sieht sich um.  S: Musik  Ort: Berg	Da kommt Johanna. Ausgerüstet wie der Teufel mit Sturzhelm und Schneebrille, und die will den Berg 'runterfahren. Aber womit? Vielleicht will sie Ski fahren? Und dazu hat sie jetzt erst mal
2	0:18		Einst: N	Johanna in Großaufnahme mit Brett vor dem Körper.  S: Musik  Ort: Berg	ein Brett dabei. Aber zum Skifahren gehören doch zwei Bretter?
3	0:23		Einst: T	Johanna legt Brett auf den Boden und versucht damit zu rutschen.  Ort: Berg	Es gibt aber doch auch diese Schnee Bretter "Snowboards". Da braucht man ein Brett. Vielleicht ist es so 'was? Aber funktionieren tut das nicht, auch wenn sie sich quer draufstellt.
4	0:30		Einst: HN	Bewegung von Akteur1 vom Vorratsstapel zu Maschine 1  S: Maschinen-geräusche  Ort: Werkstatt / Station 1 (Vorratsstapel, Fräse)	Also müssen wir 'mal dahingehen, wo solche Dinge hergestellt werden.
5	0:36		Einst: N	Ort: Werkstatt / Station1 (Fräse)	Da werden ja mehrere solcher Bretter genommen.

6	0:43		Einst: G	Ort: Werkstatt / Station1 (Fräse)	Bestimmte Anzahl. Die kommen in so 'nen Vorratsstapel. Dann wird unten von diesem Vorratsstapel immer ein Brett abgezogen. Das wandert dann unter diesem Rad dahinten durch, nach hinten.
7	0:51		Einst: HT	Akteur1 bewegt sich zum hinteren Bereich der Maschine  Ort: Werkstatt / Station1 (Fräse)	Und danach in so 'ne grüne Kiste. Vorne rein und hinten wieder raus. Und immer drei Bretter gehören zusammen. Und der Strich bedeutet nicht, dass da abgeschnitten werden soll, sondern da ist später hinten.
8	1:04		Einst: G  KBew: heranzoomen  Einst: D	Ort: Werkstatt / Station1  Maschine 1 in Großaufnahme  Maschine fräst zwei Bretter  Ort: Inneres der Maschine 1	Was passiert in der grünen Kiste?  Da laufen die Bretter unter einer Fräse durch. Das ist das runde Ding, was sich da dreht. Und was macht diese Fräse? Hier kann man's sehen: Die macht das Brett dünner an manchen Stellen, vorne dünn, dann in der Mitte etwa dicker, das ist die Stelle, wo man später drauf stehen soll, und dann zum Ende des Bretts hin, fräst sie das Brett wieder dünner.
9	1:27.1		Einst: HN KBew: Schwenkverfolgt Bewegung des Akteurs  Einst: G	Akteur1 nimmt Bretter (drei), die an der Wand lehnen und legt sie auf einen Tisch. Er wiederholt diesen Vorgang.  Ort: Werkstatt / Station 2 (Tisch)	Und dann werden immer drei Bretter nebeneinander gelegt. Aber insgesamt kommen sechs drauf, also zwei mal drei: ein mal drei, zwei mal drei. Die Bretter werden nebeneinander gelegt und dann kommt aus dieser Maschine ein spezieller Kleber.
10	1:44		Schnitt Einst: HN	Nach Schnitt wird Akteur1 von vorne gezeigt.  Ort: Werkstatt / Station 2 (Tisch)	Und damit werden die Bretter zusammengeklebt. An dieser Stelle muss man darauf achten, dass der Kleber gut in die Zwischenräume zwischen die Bretter kommt, also in die

					<p>Ritzen. Dann kommt noch ein Gewicht drauf. Und bei den anderen drei passiert das gleiche noch mal.</p>
11	1:58		<p>Einst: HT Einst:</p> <p>KBew: Schwenk um Johanna zu verfolgen</p>	<p>Johanna geht auf den Berg und hat die drei zusammengeleimten Bretter in der Hand.</p> <p>Johanna testet das Snowboard</p> <p>S: Musik</p> <p>Ort: Berg</p>	<p>So, das sieht doch jetzt schon viel besser aus. Das Brett ist breiter, man kann in der Mitte auch besser drauf stehen. Mal sehen, ob das geht. Naaaa. Vielleicht doch nicht so. Vielleicht, wenn man sich quer draufstellt? Bisschen rutschen tut's. Vielleicht auch allen Vierern mit anstoßen? Aber so der wahre Jakob, ist das auch noch nicht.</p>
12	2:19.3		<p>Einst: N</p>	<p>Akteur2 legt Bretter in die Maschine</p> <p>Ort: Werkstatt / Station 3 (Fräse)</p>	<p>Also kommt das Ding jetzt unter diese Maschine. Da wird das Brett eingeklemmt. Und was passiert da?</p>
13	2:29		<p>Einst: D KBew: Zoom-out und kleiner Schwenk</p>	<p>Ort: Werkstatt / Station 3 (Fräse, "Inneres")</p>	<p>Da laufen dann wieder Fräsen- nicht nur eine - sondern mehr, außen am Rand entlang und geben dem Ding die Form: vorne und hinten rund und an der Seite etwas schmaler.</p>
14	2:42		<p>Einst: D zoom-in, kleiner Schwenk nach unten</p>	<p>Akteur3</p> <p>Ort: Werkstatt / Station 4 (Bohrmaschine)</p>	<p>Danach unter diese wunderbare Bohrmaschine, die auf einen Schlag gleich 8 Löcher bohrt.</p>
15	2:57		<p>Einst: HT</p>	<p>Akteur3 steht vor den Maschinen. Beide Bohrmaschinen sind erkennbar.</p> <p>Ort: Werkstatt / Station 4 (Bohrmaschine)</p>	<p>Und neben dieser einen Bohrmaschine, ist eine zweite, die macht das Gleiche noch mal, noch mal 8 Löcher. So, da wollen wir 'mal sehen. Jetzt ist das</p>

16	3:06		Einst: T KBew: Schwenk verfolgt Johanna	Johanna auf dem Berg, hält das Brett in der Hand, versucht wieder zu rutschen, fällt hin.  S: Musik  Ort: Berg	Brett stromlinienförmig schön rund und hat viele Löcher. Gucken wir mal, ob Johanna damit besser zurecht kommt. Jetzt ist auch klar – es wird kein Ski, sondern Snowboard. Na, auch mit Anstrengung, höchstens ein auf den Po fallen.
17	3:18		Einst: G Weg- zoomen	Akteur4 legt Metallteile in die Löcher  Ort: Werkstatt / Station 5	Als Nächstes kommen in alle Löcher solche Metallteile. Das sind Metallgewinde, aber von den Gewinden sieht man nichts, die werden nämlich mit der Unterseite nach oben eingedrückt. und zwar in jedes dieser Löcher. Das passiert auf dieser Maschine
18	3:32		Einst: HN	Akteur 5 entfernt Schutzpapier  Ort: Werkstatt / Station 6	Das sind Kunststoffbahnen mit Blümchen drauf. Das ist die Unterseite, auf der man später fahren soll.
19	3:39		Einst: HN	Akteur 6 und Akteur7 legen Glasfaserschicht auf. Ort: Werkstatt / Station 7	Da kommt dann noch eine Schicht aus Glasfasern drauf. Beides zusammen wird dann in dieser Maschine zusammengepresst. Und wenn das passiert ist, dann werden diese beiden Sachen
20	3:55		Einst: D	Ort: Werkstatt / Station 8 (Fräse)	– genau wie eben das Brett – mit 'ner Fräse in die richtige Form gebracht. Da sieht man's: vorn und hinten rund. Und wenn das passiert ist dann kommen alle Sachen zusammen, dann wird das Brett gebaut
21	4:08		Einst: N	Ort: Werkstatt / Station 9  hier wurden kleine Filmstücke herausgeschnitten, um so die Sequenz etwas zu kürzen. (»jump cut«)	in dieser Form. Und weil das die Lauffläche ist, kommt die natürlich als erstes in die Form – soll ja schließlich unten sein. Dann kommt oben eine kleine runde Stahlkappe dran, wird festgehalten an der anderen Seite auch – da sieht man's.

22	4:22		Einst: N	Akteur8 Ort: Werkstatt / Station 9	Dann kommen zwei Gummistreifen drauf, die werden mit Kleber eingeschmiert und die kommen erst mal rechts und links an die langen Seiten. Die sollen die Schwingungen abfedern, wenn man später damit fährt.
23	4:37		Einst: N	Ort: Werkstatt / Station 9	Danach kommen jetzt lange Stahlkappen für die Seiten rein, die werden richtig festgekloppt, damit die auch richtig sitzen.
24	4:47		Einst: G	Ort: Werkstatt / Station 9	Dann noch mal Kleber drauf.
25	4:53		Einst: N	Ort: Werkstatt / Station 9	Und jetzt über die Stahlkanten nochmal ein Gummistreifen, einmal nach der ganzen Länge und dann wird ein Streifen in Stücke geschnitten.
26	5:04		Einst: G	Ort: Werkstatt / Station 9	Und aus diesen Stückchen wird dann so 'ne halbrunde Geschichte gelegt.
27	5:07		Einst: G	Ort: Werkstatt / Station 9	Dann kommen Glasfaserstreifen oben und unten an die Rundungen,
28	5:11		Einst: N	Ort: Werkstatt / Station 9	dann das Brett, das wird auch wieder mit Kleber eingestrichen, und dann an den Stellen, wo die Metallteile mit den Gewinden sind, nochmal Gummi, damit die nicht ausreißen. Dann das Ganze umdrehen und weil das Brett ja vorne dünn gefräst ist, passt es sich da auch sich da prima an die Form an. Man kann das gut da eindrücken. Festhalten.

29	5:32		Einst: N	Ort: Werkstatt / Station 9	Jetzt kommen an die Seiten noch Fiberglasstreifen. Die dienen einmal dazu, dass das Brett nicht durchbrechen kann und damit kann man bestimmen, wie stark das Brett federt, wenn man darauf fährt. Festkloppen.
30	5:46		Einst: N	Ort: Werkstatt / Station 9	Und dann kommt die Oberseite drauf, das was man nachher oben sieht. Diesmal ist nur ein Blümchen drauf.
31	5:51		Einst: N	Ort: Werkstatt / Station 9	Dann kommt der zweite Teil der Form.
32	5:54		Einst: G heranzoomen	Ort: Werkstatt / Station 10 (Presse)	Alles zusammen unter diese Presse.
33	5:56		Einst: D	Ort: Werkstatt / Station 10 (Presse)	Da bleibt das alles zusammen 15 bis 20 Minuten.
34	6:00		Einst: D	Ort: Werkstatt / Station 10 (Presse)	Wird heiß gemacht. Der Kleber klebt alles zusammen und was übersteht, das quillt an der Seite raus.
35	6:06.5		Einst: HT	Akteur8 nimmt Snowboard aus Maschine und streicht über die Unterseite, um damit zu zeigen, dass die Unterseite glatt ist.  Ort: Werkstatt / Station 10 (Presse)	Und nach etwa 15 bis 20 Minuten ist dann alles zusammen gebacken. Unterseite schön glatt
36	6:15		Einst: HN	Johanna kniet neben dem Brett, dreht es um, Blümchenmuster wird sichtbar.	Mal sehen, wie das jetzt geht? Also Johanna, nun gib mal Gas, nun ist doch alles richtig! Ja, geht schon, aber doch

				<p>Sie versucht zu rutschen, gelingt etwas, aber nicht richtig.</p> <p>S: Musik</p> <p>Ort: Berg</p>	<p>noch nicht so ganz. Irgendwas fehlt noch.</p>
37	6:27		Einst: HT	<p>S: Maschinen-geräusche</p> <p>Ort: Werkstatt / Station 11 (Stichsäge)</p>	<p>Ach so, die Oberseite ist noch zu lang. Die muss hier noch abgeseigt werden.</p>
38	6:33		Einst: N	<p>Person schleift</p> <p>S: Maschinen-geräusche</p> <p>Ort: Werkstatt / Station 12 (Schleifmaschine)</p>	<p>Und dann werden die Stahlkanten an der Seite geschliffen, einmal von Hand an dieser Maschine</p>
39	6:39		Einst: D	<p>Das Innere der Maschine wird gezeigt.</p> <p>S: Maschinen-geräusche</p> <p>Ort: Werkstatt / Station 13 (automatische Schleifmaschine)</p>	<p>und dann noch mal voll-automatisch hierdrin. Sieht so'n bisschen aus wie Feuerwerk, macht aber, dass die Kanten schön scharf sind, dass man einen guten Griff hat im Schnee.</p>
40	6:48		Einst: D	<p>Ort: Werkstatt / Station 14 (spezielle) Bohrmaschine</p>	<p>Dann werden die Löcher von den Gewinden aufgebohrt.</p>
41	6:54		Einst: D	<p>Johanna auf dem Berg, steigt in die Bindung des fertigen Snowboards ein.</p> <p>S: Musik</p> <p>Ort: Berg</p>	<p>Und in diese Löcher wird dann die Bindung eingeschraubt, ohne die man nicht fahren kann. Und in diese Bindung kann man dann mit den Skischuhen einsteigen und das Brett am Fuß festmachen, sonst verliert man das ja unterwegs.</p>
42	7:04		Einst: HT	<p>S: Musik</p> <p>Ort: Berg</p>	<p>So Johanna, jetzt aber los. Jetzt geht's wirklich los, jetzt gibt's kein Vertun mehr. Und schon saust sie los. Nur schade, dass mittlerweile die ganze schöne Aussicht verschwunden ist. Und Johanna verschwindet im Nebel.</p>

## 6.2 Tabellarische Übersicht der Analyse der Text-Bild-Relationen

Ort	Einstellung	»Off«-Text	Relationstyp	
B	01	01_01	Da kommt Johanna.	<i>expansion – extension – collection</i>
		01_02	Ausgerüstet wie der Teufel mit Sturzhelm und Schnebrille,	<i>expansion – elaboration – specification</i>
		01_03	und die will den Berg 'runterfahren.	<i>expansion – enhancement – motivation</i>
		01_04	Aber womit? Vielleicht will sie Ski fahren?	<i>decoration – hypophora*</i>
		01_05	Und dazu hat sie jetzt erstmal	<i>expansion – elaboration – explanation</i>
B	02	02_01	ein Brett dabei.	<i>expansion – elaboration – explanation</i>
		02_02	Aber zum Skifahren gehören doch zwei Bretter?	<i>decoration – hypophora*</i>
B	03	03_01	Es gibt aber doch auch diese Schneebretter "Snowboards". Da braucht man ein Brett.	<i>expansion – extension – collection</i>
		03_02	Vielleicht ist es so 'was?	<i>decoration – hypophora*</i>
		03_03	Aber funktionieren tut das nicht, auch wenn sie sich quer draufstellt.	<i>expansion – elaboration – specification</i>
W	04	04_01	Also müssen wir 'mal dahingehen, wo solche Dinger hergestellt werden.	<i>expansion – enhancement – motivation</i>
W	05	05_01	Da werden ja mehrere solcher Bretter genommen.	<i>expansion – elaboration – explanation</i>
W	06	06_01	Bestimmte Anzahl.	<i>expansion – elaboration – specification</i>
		06_02	Die kommen in so 'nen Vorratsstapel.	<i>expansion – elaboration – explanation</i>
		06_03	Dann wird unten von diesem Vorratsstapel immer ein Brett abgezogen.	<i>expansion – elaboration – identification</i>
		06_04	Das wandert dann unter diesem Rad dahinten durch, nach hinten.	<i>expansion – elaboration – identification</i>
W	07	07_01	Und danach in so 'ne grüne Kiste.	<i>expansion – elaboration – explanation</i>
		07_02	Vorne rein und hinten wieder raus.	<i>expansion – elaboration – explanation</i>
		07_03	Und immer drei Bretter gehören zusammen.	<i>expansion – elaboration – specification</i>
		07_04	Und der Strich bedeutet nicht, dass da abgeschnitten werden soll, sondern da ist später hinten.	<i>expansion – enhancement – purpose</i>
W	08	08_01	Was passiert in der grünen Kiste?	<i>decoration – hypophora*</i>
		08_02	Da laufen die Bretter unter einer Fräse durch.	<i>expansion – elaboration – explanation</i>
		08_03	Das ist das runde Ding, was sich da dreht.	<i>expansion – elaboration – identification</i>
		08_04	Und was macht diese Fräse?	<i>decoration – hypophora*</i>
		08_05	Hier kann man's sehen:	<i>expansion – elaboration – identification</i>
		08_06	Die macht das Brett dünner an manchen Stellen, vorne dünn, dann in der Mitte etwas dicker,	<i>expansion – elaboration – specification</i>

		08_07	das ist die Stelle, wo man später drauf stehen soll,	<i>expansion – enhancement – purpose</i>
		08_08	und dann zum Ende des Bretts hin, fräst sie das Brett wieder dünner.	<i>expansion – elaboration – specification</i>
W	09	09_01	Und dann werden immer drei Bretter nebeneinander gelegt.	<i>expansion – elaboration – specification</i>
		09_02	Aber insgesamt kommen sechs drauf, also zwei mal drei: ein mal drei, zwei mal drei.	<i>expansion – elaboration – specification</i>
		09_03	Die Bretter werden nebeneinander gelegt,	<i>expansion – elaboration – explanation</i>
		09_04	und dann kommt aus dieser Maschine ein spezieller Kleber.	<i>expansion – elaboration – specification</i>
W	10	10_01	Und damit werden die Bretter zusammengeklebt.	<i>expansion – elaboration – explanation</i>
		10_02	An dieser Stelle muss man darauf achten, dass der Kleber gut in die Zwischenräume zwischen die Bretter kommt, also in die Ritzen.	<i>expansion – enhancement – manner</i>
		10_03	Dann kommt noch ein Gewicht drauf.	<i>expansion – elaboration – explanation</i>
		10_04	Und bei den anderen drei passiert das Gleiche noch mal.	<i>expansion – elaboration – explanation</i>
W	11	11_01	So, das sieht doch jetzt schon viel besser aus.	<i>decoration – comment*</i>
		11_02	Das Brett ist breiter, man kann in der Mitte auch besser drauf stehen.	<i>expansion – elaboration – specification</i>
		11_03	Mal sehen, ob das geht.	<i>decoration – preparation*</i>
		11_04	Naaaa. Vielleicht doch nicht so.	<i>decoration – comment*</i>
		11_05	Vielleicht, wenn man sich quer draufstellt?	<i>decoration – invitation*</i>
		11_06	Bisschen rutschen tut's.	<i>expansion – elaboration – explanation</i>
		11_07	Vielleicht auch allen Vieren mit anstoßen?	<i>decoration – invitation*</i>
		11_08	Aber so der wahre Jakob ist das auch noch nicht.	<i>decoration – comment*</i>
W	12	12_01	Also kommt das Ding jetzt unter <i>diese</i> Maschine.	<i>expansion – elaboration – identification</i>
		12_02	Da wird das Brett eingeklemmt.	<i>expansion – elaboration – explanation</i>
		12_03	Und was passiert da?	<i>decoration – hypophora*</i>
W	13	13_01	Da laufen dann wieder Fräsen – nicht nur eine – sondern mehr, außen am Rand entlang	<i>expansion – elaboration – explanation</i>
		13_02	und geben dem Ding die Form: vorne und hinten rund und an der Seite etwas schmaler.	<i>expansion – elaboration – specification</i>
W	14	14_01	Danach unter diese wunderbare Bohrmaschine, die auf einen Schlag gleich 8 Löcher bohrt.	<i>expansion – elaboration – specification</i>

W	15	15_01	Und neben dieser einen Bohrmaschine, ist eine zweite, die macht das Gleiche noch mal, noch mal 8 Löcher.		<i>expansion – elaboration – specification</i>
		15_02	So, da wollen wir 'mal sehen. Jetzt ist das	join	<i>decoration – preparation*</i>
B	16	16_01	Brett stromlinienförmig schön rund und hat viele Löcher.		
		16_02	Gucken wir mal, ob Johanna damit besser zurecht kommt.		<i>decoration – preparation*</i>
		16_03	Jetzt ist auch klar – es wird kein Ski, sondern Snowboard.		<i>expansion – elaboration – specification</i>
		16_04	Na, auch mit Anstrengung, höchstens ein auf den Po fallen.		<i>decoration – comment*</i>
W	17	17_01	Als Nächstes kommen in alle Löcher solche Metallteile.		<i>expansion – elaboration – collection</i>
		17_02	Das sind Metallgewinde,		<i>expansion – extension – collection</i>
		17_03	aber von den Gewinden sieht man nichts, die werden nämlich mit der Unterseite nach oben eingedrückt und zwar in jedes dieser Löcher.		<i>expansion – elaboration – specification</i>
		17_04	Das passiert auf dieser Maschine.		<i>expansion – elaboration – identification</i>
W	18	18_01	Das sind Kunststoffbahnen mit Blümchen drauf.		<i>expansion – extension – collection</i>
		18_02	Das ist die Unterseite, auf der man später fahren soll.		<i>expansion – extension – collection</i>
W	19	19_01	Da kommt dann noch eine <i>Schicht aus Glasfasern</i> drauf.		<i>expansion – extension – collection</i>
		19_02	Beides zusammen wird dann in dieser Maschine zusammengepresst.		<i>expansion – elaboration – identification</i>
		19_03	Und wenn das passiert ist, dann werden diese beiden Sachen	join	<i>expansion – elaboration – specification</i>
W	20	20_01	– genau wie eben das Brett – mit 'ner Fräse in die richtige Form gebracht.		
		20_02	Da sieht man's: vorn und hinten rund.		<i>expansion – elaboration – identification</i>
		20_03	Und wenn das passiert ist, dann kommen alle Sachen zusammen, dann wird das Brett gebaut	join	<i>expansion – elaboration – explanation</i>
W	21	21_01	in dieser Form.		
		21_02	Und weil das die Lauffläche ist, kommt die natürlich als erstes in die Form – soll ja schließlich unten sein.		<i>expansion – enhancement – reason*</i>
		21_03	Dann kommt oben eine kleine runde Stahlkappe dran, wird festgehalten an der anderen Seite auch		<i>expansion – extension – collection</i>
		21_04	– da sieht man's.		<i>expansion – elaboration – identification</i>

W	22	22_01	Dann kommen zwei Gummistreifen drauf, die werden mit Kleber eingeschmiert und die kommen erstmal rechts und links an die langen Seiten.	<i>expansion – extension – collection</i>
		22_02	Die sollen die Schwingungen abfedern, wenn man später damit fährt.	<i>expansion – enhancement – purpose</i>
W	23	23_01	Danach kommen jetzt lange Stahlkappen für die Seiten rein,	<i>expansion – elaboration – specification</i>
		23_02	die werden richtig festgekloppt, damit die auch richtig sitzen.	<i>expansion – enhancement – purpose</i>
W	24	24_01	Dann noch mal Kleber drauf.	<i>expansion – elaboration – explanation</i>
W	25	25_01	Und jetzt über die Stahlkanten nochmal ein Gummistreifen, einmal nach der ganzen Länge	<i>expansion – elaboration – explanation</i>
		25_02	und dann wird ein Streifen in Stücke geschnitten.	<i>expansion – elaboration – explanation</i>
W	26	26_01	Und aus diesen Stückchen wird dann so 'ne halbrunde Geschichte gelegt.	<i>expansion – elaboration – explanation</i>
W	27	27_01	Dann kommen <i>Glasfaserstreifen</i> oben und unten an die Rundungen,	<i>expansion – extension – collection</i>
W	28	28_01	dann das Brett, das wird auch wieder mit Kleber eingestrichen,	<i>expansion – elaboration – explanation</i>
		28_02	und dann an den Stellen, wo die Metallteile mit den Gewinden sind, nochmal Gummi, damit die nicht ausreißen.	<i>expansion – enhancement – purpose</i>
		28_03	Dann das Ganze umdrehen	<i>expansion – elaboration – explanation</i>
		28_04	und weil das Brett ja vorne dünn gefräst ist, passt es sich da auch sich da prima an die Form an.	<i>expansion – enhancement – reason*</i>
		28_05	Man kann das gut da eindrücken. Festhalten	<i>expansion – elaboration – explanation</i>
W	29	29_01	Jetzt kommen an die Seiten noch <i>Fiberglasstreifen</i> .	<i>expansion – extension – collection</i>
		29_02	Die dienen einmal dazu, dass das Brett nicht durchbrechen kann und damit kann man bestimmen, wie stark das Brett federt, wenn man darauf fährt.	<i>expansion – enhancement – purpose</i>
		29_03	Festkloppen.	<i>expansion – elaboration – explanation</i>
W	30	30_01	Und dann kommt die Oberseite drauf, das was man nachher oben sieht.	<i>expansion – extension – collection</i>
		30_02	Diesmal ist nur ein Blümchen drauf.	<i>expansion – elaboration – specification</i>
W	31	31_01	Dann kommt der zweite Teil der Form.	<i>expansion – elaboration – specification</i>
W	32	32_01	Alles zusammen unter diese Presse.	<i>expansion – elaboration – identification</i>
W	33	33_01	Da bleibt das alles zusammen 15 bis 20 Minuten,	<i>expansion – enhancement – spatio-temporal</i>
W	34	34_01	wird heiß gemacht.	<i>expansion – extension – collection</i>

		34_02	Der Kleber klebt alles zusammen und was übersteht, das quillt an der Seite raus.	<i>expansion – elaboration – explanation</i>
W	35	35_01	Und nach etwa 15 bis 20 Minuten ist dann alles zusammen gebacken.	<i>expansion – enhancement – spatio-temporal</i>
		35_02	Unterseite schön glatt.	<i>expansion – elaboration – specification</i>
B	36	36_01	Mal sehen, wie das jetzt geht.	<i>decoration – preparation*</i>
		36_02	Also Johanna, nun gib mal Gas, nun ist doch alles richtig!	<i>decoration – invitation*</i>
		36_03	Ja, geht schon, aber doch noch nicht so ganz. Irgendwas fehlt noch.	<i>decoration – comment*</i>
W	37	37_01	Ach so, die Oberseite ist noch zu lang. Die muss hier noch abgesägt werden.	<i>expansion – enhancement – reason*</i>
W	38	38_01	Und dann werden die Stahlkanten an der Seite geschliffen,	<i>expansion – elaboration – specification</i>
		38_02	einmal von Hand an dieser Maschine	<i>expansion – elaboration – identification</i>
W	39	39_01	und dann noch mal vollautomatisch hierdrin.	<i>expansion – elaboration – identification</i>
		39_02	Sieht so'n bisschen aus wie Feuerwerk,	<i>expansion – elaboration – specification</i>
		39_03	macht aber, dass die Kanten schön scharf sind, dass man einen guten Griff hat im Schnee.	<i>expansion – enhancement – purpose</i>
W	40	40_01	Dann werden die Löcher von den Gewinden aufgebohrt.	<i>expansion – elaboration – explanation</i>
B	41	41_01	Und in diese Löcher wird dann die Bindung eingeschraubt, ohne die man nicht fahren kann.	<i>expansion – enhancement – purpose</i>
		41_02	Und in diese Bindung kann man dann mit den Skischuhen einsteigen und das Brett am Fuß festmachen, sonst verliert man das ja unterwegs.	<i>expansion – enhancement – purpose</i>
B	42	42_01	So Johanna, jetzt aber los. Jetzt geht's wirklich los, jetzt gibt's kein Vertun mehr.	<i>decoration – invitation*</i>
		42_02	Und schon saust sie los.	<i>expansion – elaboration – explanation</i>
		42_03	Nur schade, dass mittlerweile die ganze schöne Aussicht verschwunden ist.	<i>decoration – comment*</i>
		42_04	Und Johanna verschwindet im Nebel.	<i>expansion – elaboration – explanation</i>

## 6.3 Dokumentation der Funktionen

### 6.3.1 XSLT-Funktionen

Die folgenden XSLT-Funktionen befinden sich in der Datei `functions.xml`.

`fu:bname(id)`

Aus der Einstellungsnummer wird der Name der Bilddatei berechnet

Parameter:

`id` Einstellungsnummer

`fu:inSec(zk)`

Eine Zeitangabe der Form `min:sec` wird in Sekunden umgerechnet

Parameter:

`zk` Zeichenkette der Form `min:sec`

`fu:dauer(zk1,zk2)`

Aus zwei Zeitangaben wird die Differenz (= Dauer in Sekunden) wird berechnet

Parameter:

`zk1, zk2` Zeichenketten der Form `min:sec`

### 6.3.2 JavaScript-Funktionen

Die folgende Funktion befindet sich in der Datei `cssfunction.js`.

`toggleClass(id,class1,class2)`

Austausch der Stylesheetklasse `class1` / `class2`

Parameter:

`id` id des Elementes, dessen Stylesheetklasse ausgetauscht werden soll  
`class1` Stylesheetklasse 1  
`class2` Stylesheetklasse 2

Die folgenden JavaScript-Funktionen befinden sich in der Datei `videofunction.xml`.

`videoPlayLoop(id, beginn, dauer, loop)`

spielt ein Video ab einem bestimmten Zeitpunkt für eine bestimmte Dauer ab.

Parameter:

`id` id des Videos  
`begin` Abspielbeginn (in Sekunden)  
`dauer` Abspieldauer (in Millisekunden)  
`loop` Endlosschleife (`true/false`)

`videoPlayStop(id1, id2, beginn, dauer)`

Das Video mit der Id `id1` wird ab einem bestimmten Zeitpunkt für eine bestimmte Dauer abgespielt, das Video 2 wird nach Ende von Video 1 gestoppt.

Parameter:

<code>id1</code>	id-Wert des Videos 1
<code>id2</code>	id-Wert des Videos 2
<code>begin</code>	Abspielbeginn (in Sekunden)
<code>dauer</code>	Abspieldauer (in Millisekunden)

`video(id,begin,dur,sound)`

JavaScript-Objektyp `Video` zur Speicherung von Informationen zu einem Video-Segment.

Attribute:

<code>id</code>	Id des Zielelementes
<code>begin</code>	Beginn (in Sekunden)
<code>dur</code>	Dauer in Millisekunden
<code>sound</code>	mit oder ohne Ton (1 oder 0)
<code>txt</code>	beschreibender Text

Methoden: Zu jedem Attribut gibt es eine `get`-Methode.

`videoSeq(x)`

Eine Folge von Videosegmenten wird abgespielt.

Parameter:

<code>x</code>	Array, das Instanzen des (selbstdefinierten) JavaScript-Objektyps <code>Video</code> enthält.
----------------	---

`videoDelay(id,delay,begin,dur)`

Ein Videosegment wird nach einer Verzögerungszeit angespielt.

Parameter:

<code>id</code>	id-Wert des Videos
<code>delay</code>	Verzögerungszeit
<code>begin</code>	Beginn (in Sekunden)
<code>dur</code>	Dauer in Millisekunden

In den Funktionen `videoPlayLoop`, `videoPlayStop`, `videoSeq` und `videoDelay` wird die Methode `setTimeout(anweisung,zeit)` des JavaScript `window`-Objektes<sup>1</sup> verwendet. Sie führt eine Anweisung nach einer bestimmten Verzögerungszeit aus. Übergeben werden zwei Parameter: der erste Parameter ist die JavaScript-Anweisung, die verzögert ausgeführt werden soll, der zweite Parameter die Verzögerungszeit in Millisekunden. In den drei programmierten Funktionen ist die Anweisung, die verzögert ausgeführt werden soll, das Pausieren der Videos (DOM-Methode `pause()`) und die Verzögerungszeit die Dauer, die ein Videosegment angespielt werden soll.

---

<sup>1</sup> Das `window`-Objekt existiert immer, da ja immer ein Browserfenster geöffnet ist. Es kann (und wird meistens) beim Methodenaufruf weggelassen.

In der Datei `picstory.js` befindet sich der JavaScript-Code für alle Interaktionen der Bilder geschichten im HTML-Format. Verwendet wird hierzu das JavaScript-Framework jQuery<sup>1</sup> in der Version 1.9, was folgende Vorteile hat:

- Vermeidung der Vermischung von HTML-Code mit JavaScript-Elementen (z. B. Event-Handler-Attributen),
- Methoden für visuelle Effekte, DOM-Manipulation (und vieles mehr) sind im Framework vorhanden.

Ein typischer jQuery-Aufruf sieht folgendermassen aus: `$(selector).action()`. Hierbei ist `$()` die Kurzform für die Funktion `jQuery`. Der angegebene Selektor wählt ein oder mehrere HTML-Elemente aus, auf die dann die Methode `action()` angewandt wird.

Verwendet werden folgende Methoden des jQuery-Frameworks:

- `css("color", "farbe")` CSS-Eigenschaft Farbe auf neuen Wert setzen
- `fadeIn()` / `fadeOut()` Langsames Ein- / Ausblenden
- `show()` / `hide()` Anzeigen / Verbergen eines Elementes
- `toggle()` Wechsel zwischen Sichtbar und unsichtbar

Beim dem Laden des Dokumentes (`$(document).ready()`) wird die erste »Seite« der Bilder geschichte angezeigt und der erste Eintrag des Inhaltsverzeichnisses farbig markiert. Die Eventhandler (Click-Handler) werden den Einträgen des Inhaltsverzeichnisses und den Links zum Vor- und Zurückblättern zugeordnet.

---

<sup>1</sup> <http://jquery.com> [letzter Aufruf: 04.03.2015]

## 6.4 Übersicht aller Dateien auf der CD

Die folgenden Dateien befinden sich auf der beigelegten CD.

Ordner	Dateiname	Beschreibung
<b>css</b>		<b>CSS-Dateien</b>
	change_alternation.css	Modifikation der Alternanz und der Kohärenzrelation (→ change_alternation.htm)
	para_relation.css	Darstellung der paradigmatischen Relationen (→ para_Txx_yy.htm)
	picstory.css	Bildergeschichten (→ picStory_detail.htm, picStory_simple.htm)
	separate_location.css	Wegfall der Kohärenzrelation (→ separate_location.htm)
	show_units.css	Anzeige der einzelnen Einstellungen (→ show_units.htm)
	video_synchron.css	Betonung des kohärent Relationierten (→ video_synchron1.htm, ..., video_synchron4.htm)
	work_progress.css	Betonung der Kohärenzrelation (→ work_progress1.htm, ..., work_progress4.htm)
<b>dtd</b>		<b>DTDs</b>
	config_alternation.dtd	Konfigurationsdatei zu Modifikation der Alternanz und der Kohärenzrelation
	paradigmatic_analysis.dtd	Paradigmatische Analyse
	paradigmatic_relation.dtd	Paradigmatische Relationen
	syntagmatic_analysis.dtd	Syntagmatische Analyse
	textImage_analysis.dtd	Analyse der Text-Bild-Relationen
	textImage_relation.dtd	Taxonomie der Text-Bild-Relationen
	units.dtd	Informationen über die Einstellungen (Transkriptionsprotokoll)
<b>html</b>		<b>Generierte HTML-Dateien</b>
	change_alternation.htm	Modifikation der Alternanz und der Kohärenzrelation
	para_T01_T02.htm ... para_T41_T42.htm	Darstellung der paradigmatischen Relation
	picStory_detail.htm	Detaillierte Bildergeschichte
	picStory_simple.htm	Einfache Bildergeschichte
	separate_location.htm	Wegfall der Kohärenzrelation
	show_units.htm	Abspielen der einzelnen Einstellungen
	video_synchron_v_1.htm ... video_synchron_v_4.htm	Betonung des kohärent Relationierten  v=1,...,4 für die Varianten
	work_progress1.htm ... work_progress4.htm	Betonung der Kohärenzrelation
<b>js</b>		<b>JavaScript-Dateien</b>
	cssfunction.js	Funktion zum Wechsel der CSS-Stylesheetklasse - toggleClass() (→ show_units.htm, separate_location.htm)
	jquery-1.9.0.min.js	jQuery Framework
	picstory.js	Interaktionen in den Bildergeschichten: vorwärts- und rückwärtsblättern, fadeIn/fadeOut-Effekt, Sprung vom Inhaltsverzeichnis zum Kapitel, Text anzeigen/verbergen (verwendet jQuery)

	videofunction.js	Funktionen zur Videosteuerung - videoPlayLoop(id,begin,dur,loop) - videoPlayStop(id1,id2,begin,dur) - JavaScript-Objekt Video - videoSeq(x) - videoDelay(id,delay,bgin,dur) (→ change_alternation.htm, para_Txx_Tyy.htm, separate_location.htm show_units.htm, video_synchron_x_y.htm, work_progressx.htm)
<b>pdf</b>		<b>Generierte PDF-Dateien</b>
	para_relation.pdf	Übersicht: paradigmatische Relationen
	text_image_report.pdf	Auswertung Text-Bild-Relationen
	transition_table.pdf	Wechseltabelle
<b>smil</b>		
- assets*	button_vor.gif,...	Bilder für Navigationsbuttons
- img*		
- icon		
	collection.png comment*.png ....	Icons zur Visualisierung der Text-Bild-Relation in den detaillierten Bildergeschichten (→ picStory_detail.htm, picStory_detail.smil)
-snowboard*		
	m00001.png ... m00042.png	Screenshot je Einstellung, Format: 102 x 77
- snowboard_d*		
	m00001_01.png ... m00042_04.png	Screenshot je Textsegment, Format: 320 x 240
-snowboard_g*		
	m00001.png ... m00042.png	Screenshot je Einstellung, Format: 320 x 240
<b>- video*</b>		
	snowboard.mp4 snowboard.ogv snowboard.webm	Video »Snowboard« in verschiedenen Formaten
<b>----</b>		
	change_alternation.smil	Modifikation der Alternanz und der Kohärenzrelation
	para_T01_T02.smil ... para_T41_T42.smil	Darstellung der paradigmatischen Relationen
	picStory_detail.smil	Detaillierte Bildergeschichte
	picStory_simple.smil	Einfache Bildergeschichte
	separate_location.smil	Wegfall der Kohärenzrelation
	video_synchron1.smil ... video_synchron4.smil	Betonung des kohärent Relationierten für die 4 Varianten
	work_progress.smil	Betonung der Kohärenzrelation
<b>xml</b>		
	config_snowboard.xml	Konfigurationsdatei zu Modifikation der Alternanz und der Kohärenzrelation
	paradigmatic_analysis.xml	Paradigmatische Analyse von »Snowboard«
	paradigmatic_relation.xml	Sammlung der vorkommenden paradigmatischen Re- lationen
	paraRelation_template.xml	Template zur Erfassung der paradigmatischen Analyse

		(erzeugt durch: paraRelation_gen.xml)
	segment_test_loc_detail.xml	Ausgabe von segment_test_loc_detail.xml (zum Test) XML-Ausgabe der Segmentierung nach Orten (loc_detail)
	segment_test.xml	Ausgabe von segment_test.xml (zum Test) XML-Ausgabe der Segmentierung nach Orten
	snowboard.xml	XML-Repräsentation der Einstellungen (Transkriptionsprotokoll) des Videos »Snowboard«
	syntagmatic_analysis.xml	Syntagmatische Analyse des Videos »Snowboard«
	textImage_analysis.xml	Analyse der Text-Bild-Relationen des Videos »Snowboard«
	textImage_relation.xml	Taxonomie der Text-Bild-Relationen nach Kong und eigene Erweiterungen
	textImage_template.xml	Template zur Erfassung der Text-Bild-Relationen (erzeugt durch: textImage_relation_gen.xml)
<b>xsl</b>		<b>XSLT-Stylesheets</b>
	blaettern.xml	Erzeugt Links zum vor- und zurückblättern (in HTML) inkludiert in: video_synchron.xml, work_progress.xml
	change_alternation_html.xml	Modifikation der Alternanz und der Kohärenzrelation - -> HTML
	change_alternation_html.xml	Modifikation der Alternanz und der Kohärenzrelation - -> SMIL
	functions.xml	Hilfsfunktionen (bname, inSec, dauer) inkludiert in: - change_alternation_html.xml - change_alternation_smil.xml - para_relation_show.xml - para_relation_show_fo.xml - picStory_detail_html.xml - picStory_detail_smil.xml - picStory_simple_html.xml - picStory_simple_smil.xml - video_synchron.xml - work_progress.xml
	itinerar_gen.xml	Berechnung des Itinerars (erste, letzte Einstellung jedes Ortes) aus der syntagmatischen Analyse inkludiert in: - change_alternation_html.xml - change_alternation_smil.xml
	paraRelation_gen.xml	Template zur Erfassung der Analyse der paradigmatischen Relationen
	paraRelation_show_fo.xml	Darstellung der paradigmatischen Relationen --> PDF
	paraRelation_show.xml	Darstellung der paradigmatischen Relationen --> HTML, SMIL
	picStory_detail_html.xml	Detaillierte Bildergeschichte --> HTML
	picStory_detail_smil.xml	Detaillierte Bildergeschichte --> SMIL
	picStory_simple_html.xml	Einfache Bildergeschichte --> HTML
	picStory_simple_smil.xml	Einfache Bildergeschichte --> SMIL
	segment_loc_detail.xml	Benanntes Template segment_loc_detail Berechnung einstellungsübergreifender Ortsgruppen über location/loc_detail inkludiert in: - paraRelation_gen.xml
	segment_test_loc_detail.xml	Testausgabe zu segment_loc_detail.xml
	segment_test.xml	Testausgabe zu segment.xml
	segment.xml	Benanntes Template segment Berechnung einstellungsübergreifender Ortsgruppen

		über location/loc_name, inkludiert in: - paraRelation_gen.xml - picstory_detail_smil.xml - picstory_detail_html.xml - picstory_simple_smil.xml - picstory_simple_html.xml - separate_location_html.xml - separate_location_smil.xml - video_synchron.xml - work_progress.xml
	separate_location_html.xml	Wegfall der Kohärenzrelation --> HTML
	separate_location_smil.xml	Wegfall der Kohärenzrelation --> SMIL
	show_units.xml	Ansteuerung der einzelnen Einstellungen --> HTML
	textImage_relation_gen.xml	Template zur Erfassung der Text-Bild-Relationen
	textImage_report.xml	Auswertung Text-Bild-Relationen --> PDF
	transition_table.xml	Erzeugung der Wechseltabelle --> PDF
	video_synchron.xml	Betonung des kohärent Relationierten --> HTML, SMIL
	work_progress.xml	Betonung der Kohärenzrelation --> HTML, SMIL

Zu \*:

Smil hat Probleme bei der Auflösung von relativen Pfaden, wenn diese zuerst ins Elternverzeichnis gehen. Daher sind die mit \* gekennzeichneten Ordner Unterordner des Ordners smil.