



Jens Philipp Lanwer

Erzählen im virtuellen Interaktionsraum

» NET.WORX **84**

NETWORX

IMPRESSUM

| | |
|------------------------|--|
| Herausgeber | Dr. Jens Runkehl, Prof. Dr. Peter Schlobinski, Dr. Torsten Siever |
| Editorial-Board | Prof. Dr. Jannis Androutsopoulos (Universität Hamburg) für den Bereich Medienanalyse; Prof. Dr. Christa Dürscheid (Universität Zürich) für den Bereich Mobile Kommunikation; Prof. Dr. Nina Janich (Technische Universität Darm- stadt) für den Bereich Werbesprache; Prof. Dr. Ulrich Schmitz (Universität Duisburg- Essen) für den Bereich Digitale Kommunikation |
| ISSN | 1619-1021 |
| Anschrift | <i>Nieders.:</i> Leibniz Universität Hannover, Deutsches Seminar, Königsworther Platz 1, 30167 Hannover <i>Nordrhein-Westfalen:</i> RWTH Aachen, Institut für Sprach- und Kommunikationswissenschaft, Eilf- schornsteinstraße 15, 52062 Aachen Internet: www.mediensprache.net/networx/ E-Mail: networx@mediensprache.net |

ZU DIESER ARBEIT

| | |
|--------------------------|---|
| Autor & Titel | Jens Philipp Lanwer (2019). Erzählen im virtuellen Interaktionsraum. |
| Version | 1.0 (2019-02-05) |
| DOI | 10.15488/4435 |
| Zitierweise | Lanwer, Jens Philipp (2019). Erzählen im virtuellen Interaktionsraum. < http://www.mediensprache.net/ networx/networx-84.pdf >. In: <i>Networx</i> . Nr. 84. Ver. 1.0. ISSN: 1619-1021. DOI: 10.15488/4435. |
| Zitiert nach | Runkehl, Jens und Torsten Siever (?2001). Das Zitat im Internet. Ein Electronic Style Guide zum Publizie- ren, Bibliografieren und Zitieren. Hannover |

MANUSKRIPTE

| | |
|-----------------------|---|
| Einsendung | Die Einsendung von Beiträgen und Mitteilungen sind an folgende E-Mail-Adresse zu richten: networx@ mediensprache.net oder an die Postadresse: Dr. Jens Runkehl, Institut für Sprach- und Kommunikati- onswissenschaft der RWTH Aachen, Eilfschornstein- straße 15, 52062 Aachen. |
| Autorenhinweis | Mit der Annahme des Manuskripts zur Veröffentlichung in der Schriftenreihe Networx räumt der Autor dem Projekt mediensprache.net das zeitlich, räumlich und inhaltlich unbeschränkte Nutzungsrecht ein. Dieses beinhaltet das Recht der Nutzung und Wieder- gabe. Ein Recht auf Veröffentlichung besteht nicht. |
| Begutachtung | Die Begutachtung eingesandter Beiträge wird von den Herausgebern sowie den Vertretern des Editorial Board vorgenommen. |

Networx

ist die Online-Schriftenreihe des
Projekts [mediensprache.net](http://www.mediensprache.net). Die
Reihe ist eine eingetragene Publika-
tion beim Nationalen ISSN-Zentrum
der Deutschen Bibliothek in Frank-
furt am Main und ist ausgestattet
mit einem DOI.

Einsenden?

Möchten Sie eine eigene Arbeit in
der Networx-Reihe veröffentlichen?
Dann senden Sie uns Ihren Text
an folgende E-Mail-Adresse: networx@mediensprache.net oder per
Snail-Mail an: Dr. Jens Runkehl,
Institut für Sprach- und Kommuni-
kationswissenschaft, Eilfschornstein-
straße 15, 52062 Aachen.

Homepage:

Alle Arbeiten der Networx-Reihe
sind kostenlos im Internet down-
loadbar unter:

[https://www.mediensprache.net/
networx/](https://www.mediensprache.net/networx/)

Copyright

© Projekt [mediensprache.net](http://www.mediensprache.net)
Die Publikationsreihe Networx sowie
alle in ihr veröffentlichten Beiträge
und Abbildungen sind urheberrecht-
lich geschützt. Jede Verwertung
außerhalb der engen Grenzen des
Urheberrechtsgesetzes ist ohne
ausdrückliche Zustimmung des Pro-
jekts [mediensprache.net](http://www.mediensprache.net) unzulässig
und strafbar. Das gilt insbesondere
für Vervielfältigungen, Übersetzun-
gen, Mikroverfilmungen und die
Einspeicherung und Verarbeitung in
elektronischen Systemen.

Informationsstand

Stand der hier angegebenen
Informationen – soweit nicht anders
vermerkt ist: **Januar 2019**

Inhaltsverzeichnis

| | | |
|----------|--|-----------|
| 1 | Einleitung | 4 |
| 2 | Das virtuelle Hier der Face-to-Screen-Interaktion | 5 |
| 3 | Blickverhalten in konversationellen Erzählungen | 10 |
| 4 | Multimodales Erzählen in Skype-Gesprächen | 14 |
| 4.1 | Daten und Transkription | 15 |
| 4.2 | Exemplarische Analyse einer Erzählsequenz | 16 |
| 4.2.1 | Sequenzvorlauf und thematischer Auslöser | 17 |
| 4.2.2 | Story preface | 21 |
| 4.2.3 | Erzähleinstieg | 23 |
| 4.2.4 | Ausbau der szenischen Rekonstruktion | 25 |
| 4.2.5 | Seitensequenz und Vorbereitung der Pointe | 27 |
| 4.2.6 | Darbietung der Pointe | 31 |
| 4.2.7 | Evaluation und Sequenzausstieg | 34 |
| 5 | Systematisierung der Beobachtungen und Ausblick | 38 |
| 6 | Literatur | 41 |

1. Einleitung

Eine multimodale Perspektive auf soziale Interaktion verändert „zwangsläufig klassische Gegenstände und Fragestellungen der Konversationsanalyse“ (Schmitt 2005:24). Dies betrifft ganz sicher auch die konversationsanalytisch ausgerichtete Erzählforschung: Analysen zum multimodalen Erzählen werfen insbesondere auf das Rezipientenverhalten im Kontext narrativer Sequenzen ein neues Licht, da sich aufgrund des notwendigen Zurückfahrens der verbal-mündlichen Aktivitäten gerade im Bereich des Nonverbalen spezifische, für das Zuhören und Zusehen (vgl. Schmitt 2005:26) konstitutive Verhaltensweisen wie Kopfnicken (Stivers 2008), Blickverhalten (C. Goodwin 1984; C. Goodwin 1986; Rossano 2012; Rossano 2013) oder Körperausrichtung (C. Goodwin 1984; C. Goodwin 1986) manifestieren. Besonders Befunde zum Blickverhalten in narrativen Sequenzen liefern Hinweise darauf, dass soziale Akteure in der Face-to-Face-Interaktion die Blickorientierung als eine Art *participation display* zwecks Herstellung eines spezifischen Beteiligungsrahmens nutzen (C. Goodwin 1981; C. Goodwin 2000; M.H. Goodwin 1997; C. Goodwin/M.H. Goodwin 2004; Levinson 1988). In Situationen des kollaborativen Erzählens spielt das Blickverhalten außerdem eine zentrale Rolle in Bezug auf die Vergabe der primären Erzählerrolle (Zima 2017). Die analytische Ausblendung entsprechender Displays führt zwangsläufig zu einer unvollständigen Beschreibung narrativer Praktiken unter Bedingungen der Kopräsenz. Eine multimodale Analyseperspektive eröffnet jedoch nicht nur einen weiter gefassten Blick auf Praktiken des konversationellen Erzählens in der Face-to-Face-Interaktion (im Folgenden FtF), sondern ermöglicht es auch, Erzählpraktiken im Rahmen mediatisierter Interaktionsformen wie der Videotelefonie, d.h. in der Face-to-Screen-Interaktion (im Folgenden FtS) unter Berücksichtigung der potenziell verfügbaren Ausdrucksmodi systematisch zu analysieren und auf diese Weise Einblicke zu liefern in multimodale Strategien der kooperativen Hervorbringung narrativer Sequenzen unter Bedingungen virtueller Kopräsenz (Fornel 1994:110).

Die Nutzung von web-basierten Diensten zur Durchführung von Videotelefonaten wie Facetime, Google Hangouts oder eben Skype dient häufig der Beziehungspflege über die Distanz – vor allem zwischen Personen, die eine mehr oder weniger enge soziale Beziehung pflegen. Entsprechende Techniken bedienen wachsende kommunikative Bedürfnisse „of geographically dispersed networks of families, friends and lovers“ (Harper/Watson/Licoppe 2017:301). Im Rahmen von Videotelefonaten werden zwecks Aufrechterhaltung der sozialen Bindung, häufig Erlebnisse in den räumlich getrennten Lebenswelten narrativ rekonstruiert und durch die gemeinsame Bearbeitung im Videotelefonat als Bestandteil des *personal common ground* (Clark 1996:112–116) etabliert. Videotelefonate weisen entsprechend nicht selten erhebliche narrative Anteile auf, die instruktive Einblicke in soziale Konstruktionsprozesse im Zeitalter des Web 2.0 eröffnen. Zum anderen bietet die Analyse von Videotelefonaten die Möglichkeit, auf Interaktionsereignisse bei systematischer Berücksichtigung des Sichtbaren analytisch zuzugreifen, ohne die ‚Natürlichkeit‘ dieser Ereignisse durch das Platzieren irgendeiner Art von Aufnahmetechnik zu beeinträchtigen. Videokamera und Mikrofonierung sind

Teil der verwendeten Kommunikationstechnik selbst. Die Aufzeichnung greift lediglich auf ohnehin schon technisch erfasstes Datenmaterial zu. Hinzu kommt, dass der analysierende Forscher hier denselben, sich aus der Kameraperspektive ergebenden Einschränkungen unterliegt, wie sie für das äußere Wahrnehmungsfeld der Interagierenden zu veranschlagen sind. Dies scheint besonders vor dem Hintergrund der konversationsanalytischen Forderung nach einer möglichst guten Annäherung an die Teilnehmerperspektive der Interagierenden bedeutsam.

Im vorliegenden Beitrag soll in einer Detailanalyse einer Erzählsequenz aus einem Videotelefonat mit Skype aufgezeigt werden, dass sich Muster im Blickverhalten der beteiligten Akteure entdecken lassen, die darauf hindeuten, dass die Blickausrichtung – vor allem in der wechselseitigen Bezogenheit der Akteure – als Ressource der Handlungskoordination genutzt wird, um eine konversationelle Erzählung als *embodied joint project* kooperativ hervorzubringen. Dies lässt vermuten, dass die Affordanzen der verwendeten Kommunikationstechnik es den Akteuren ermöglichen, in einer Art virtuellem Interaktionsraum miteinander zu interagieren. Um diesen Punkt interaktionstheoretisch adäquat fassen zu können, gilt es zunächst das in der interaktional ausgerichteten Multimodalitätsforschung etablierte Konzept des Interaktionsraumes (Deppermann/Schmitt 2007; Hausendorf 2010; Hausendorf 2013; Hausendorf 2015; Mondada 2007; Mondada 2013) auf die Modellierung des ‚Raumes‘ der FtS zu übertragen (Abschnitt 2). Im Anschluss wird ein kurzer Überblick relevanter Forschungsergebnisse bezüglich des Blickverhaltens sozialer Akteure im Rahmen der kooperativen Herstellung konversationeller Erzählungen gegeben (Abschnitt 3), vor deren Hintergrund schließlich die Befunde der eigenen exemplarisch-explorativen Analyse (Abschnitt 4) systematisiert und interpretiert werden können (Abschnitt 5).

2. Das virtuelle Hier der Face-to-Screen-Interaktion

Der Einstieg in eine soziale Interaktion erfordert die Etablierung eines Grundmaßes an Koorientierung von mindestens zwei Interaktionspartnern: „Without it, there is no interaction“ (Hausendorf 2013:290). In der durch Kopräsenz gekennzeichneten FtF-Situation sind die Ressourcen zur Herstellung dieser Koorientierung hinsichtlich der verfügbaren Sinnesmodalitäten maximal komplex: „Das Bewußtseinserleben des anderen [ist mir] durch ein Maximum an Symptomfülle zugänglich“ (Schütz/Luckmann 1979:95). Außerdem manifestiert sich die Koorientierung der Interaktionspartner hier aufgrund einer „echten Gleichzeitigkeit der Erlebnisabläufe“ (Schütz/Luckmann 1979:92) in einer maximalen Synchronisation der Sinneswahrnehmungen der Akteure. Eine wie auch immer geartete Form der Synchronisation von Wahrnehmungen sozialer Akteure kann allgemein als Grundvoraussetzung für die Etablierung reflexiver Wahrnehmungswahrnehmung und damit als konstitutiv für soziale Interaktion gelten: „Interaktion beginnt damit, [...] dass wahrgenommen werden kann, dass wahrgenommen wird“ (Hausendorf 2015:52). Konstitutiv hierfür scheint primär ein geteiltes Jetzt der Sinneserfahrungen. So ist bspw. für verbal-mündliche Interaktion in erster Linie von

Bedeutung, dass Interagierende ihre Aufmerksamkeit gemeinsam auf prosodisch konturierte Zeitabschnitte des akustischen Sprachsignals richten (vgl. Chafe 1994:53–70; Langacker 2001:154–163). Für Gespräche als im Wesentlichen vokalsprachlich gestützte Form der Interaktion ist somit insbesondere die Gleichzeitigkeit der Erlebnisabläufe in Bezug auf den akustischen Kanal entscheidend, während die Kopräsenz in einem gemeinsamen Hier nicht in gleicher Weise erforderlich bzw. im Fall des Telefonierens sogar gänzlich entbehrlich ist (vgl. Auer 2017:3).

Anders, als es Hausendorf (2015:47) postuliert, wird hier jedoch nicht davon ausgegangen, dass das Telefonieren daher zeigt, „dass und wie durch technische Hilfsmittel Wahrnehmungswahrnehmung auf die Hörbarkeit des Gehörtwerdens reduziert werden kann.“ Die Tätigkeit des Hörens hat nicht in dem Sinne ein ‚unmittelbares‘ äußerlich wahrnehmbares Korrelat, wie es für das Sehen zu veranschlagen ist. Und selbst, wenn wir dieses Korrelat in Analogie zum Sehen in der Ausrichtung der betreffenden Sinnesorgane, d.h. der Ohren sehen wollten, müssten wir zugestehen, dass diese zum einen im Rahmen eines Ferngesprächs nicht sichtbar sind und dass zum anderen auch in der FtF-Interaktion bekanntermaßen nicht die Ohr- sondern die Blickzuwendung als Index des (aufmerksamen) Zuhörens fungiert. In der FtF-Situation impliziert das Zusehen das Zuhören (C. Goodwin 1981; Heath 1982). Für die Herstellung eines Interaktionsraumes als eine durch reflexive Wahrnehmungswahrnehmung konstituierte Sphäre der zwischenmenschlichen Begegnung ist daher immer ein geteiltes Hier als Basis visueller Koorientierung konstitutiv.

Ein Interaktionsraum kann mit Hausendorf als „mutually shared ‚here‘ for perception, movement and action“ (Hausendorf 2013:277), begriffen werden, das von den Interagierenden „im Medium der Wahrnehmungswahrnehmung [...] hergestellt“ (Hausendorf 2015:43) wird. Ein ‚virtueller‘ Interaktionsraum ist in dieser Sichtweise ebenso wie ein ‚realer‘¹ als Resultat interaktiver Aushandlungsprozesse zu begreifen und zu beschreiben. Die jeweiligen Voraussetzungen der interaktiven Herstellung sind jedoch durchaus andere: Der Unterschied liegt vor allem darin, dass der geteilte Handlungsraum bei faktischer, körperlich-materieller Kopräsenz auf Basis eines geteilten Wahrnehmungsraumes in seiner gesamten multimodalen Komplexität (Schmitt 2005:52) hergestellt werden kann. In der FtF ist der kinästhetische Zugriff der Akteure auf ihre Umwelt eingebettet in eine physisch geteilte Ökologie, was sich in einer Überlappung der jeweiligen Welten in unmittelbarer Reichweite (graue Schattierungen in Abbildung 1) manifestiert (vgl. auch Stukenbrock 2013:403). Die Akteure begegnen sich in einem „gemeinsamen Sektor des lebensweltlichen Raums und der Weltzeit“

¹ Die begriffliche Unterscheidung zwischen einem ‚realen‘ und einem ‚virtuellen‘ Interaktionsraum ist hier nicht in dem Sinne zu verstehen, dass ersterer für die Interagierenden als *interactive achievement* per se realer ist. Der Punkt ist vielmehr der, dass im zweiten Fall die Kopräsenz der Interagierenden physisch als solche nicht greifbar ist. Die Akteure befinden sich nicht faktisch in ein und derselben Ökologie, in der sie einen Interaktionsraum herstellen. De Fornel (1994:110) spricht hier auch von einer „co-présence virtuelle“. Ein virtueller Interaktionsraum wird auf der Basis einer solchen virtuellen Kopräsenz erzeugt, kann aber in seiner Beschaffenheit als sozio-kognitives Konstrukt für die Interagierenden durchaus ‚real‘ sein.

(Schütz/Luckmann 1979:90–91), sodass für sie auf der Basis reflexiver Wahrnehmungswahrnehmung auf allen sensorischen Ebenen ein geteilter Wahrnehmungsraum (GWR), ein geteilter Bewegungsraum (GBR) und ein geteilter Handlungsraum (GHR) gewissermaßen vollumfänglich herstellbar wird (Abbildung 1a). Im Rahmen der FtS sind die Akteure bei geteilter Weltzeit in ihrem kinästhetischen Zugriff hingegen auf die sie jeweils umgebende Ökologie beschränkt. Nur durch die medientechnische Kopplung, die auch den Bereich des Sichtbaren betrifft, eröffnet sich ein reflexiver sensorischer Zugriff auf den Körper des jeweils anderen. Der Körper des anderen wird dadurch potenziell auch hier zum auslegbaren „Ausdrucksfeld, das mir sein bewußtes Leben erschließt“ (Schütz/Luckmann 1979:90–91).

Bereits Goffman (1963:16) merkt an: „When two-way television is added to telephones, the unique contingencies of direct interaction will finally be available for those who are widely separated.“ Die Videotelefonie ermöglicht es daher, einen wichtigen Teil der Bedingungen der Reziprozität der FtF medientechnisch nachzubilden und einen geteilten Wahrnehmungs- und Bewegungsraum zu konstituieren (Abbildung 1b); allerdings mit sensorischen Einschränkungen. Zwar sind über den geteilten Wahrnehmungsraum die Bewegungen des jeweils anderen der eigenen Wahrnehmung zugänglich und damit koordinier- und semiotisierbar. Es gibt aber beispielsweise – und dies liegt auf der Hand – nicht die Möglichkeit der körperlichen Berührung, die bei physischer Kopräsenz bspw. zum Ausdruck sozialer Nähe oder emotionaler Anteilnahmen genutzt werden kann (M.H. Goodwin 2017). Der geteilte Bewegungsraum ist in FtS grundsätzlich auf die visuellen Konstitutionsbedingungen angewiesen, während sich in der FtF-Situation ein solcher auch auf der Basis ‚taktiler Reziprozität‘ herstellen lässt. Die „re-creation of face-to-face communication in virtual space“ (Keating/Mirus 2004:694) ist lediglich verbunden mit einer Etablierung visueller Reziprozität.

Gänzlich anders ist dies schließlich im Falle konventioneller Telefonie. Die Interagierenden haben hier keinerlei Zugriff auf das körperliche Verhalten des jeweils anderen. Das übermittelte akustische Signal verweist zwar auf körperliche Aktivitäten des Gesprächspartners. Der in Bezug auf die Wahrnehmungsaktivitäten der Akteure ‚anzeichenhafte‘ Körper ist als solcher der Wahrnehmung jedoch nicht zugänglich. Die Grundlage der Interaktionskonstitution kann daher keine Form reflexiver Wahrnehmungswahrnehmung sein, die auf der reflexiven Wahrnehmbarkeit verkörperter Wahrnehmungsaktivitäten beruht. Die Etablierung und Aufrechterhaltung des interaktionskonstitutiven geteilten Aufmerksamkeitsfokus basiert allein auf dem Umstand, dass das verbale Verhalten der Akteure einerseits auf ein jeweils dahinterliegendes Bewusstsein zurückverweist und andererseits in seiner pragmatischen und zeitlich-sequenziellen Passung auf sich aktual vollziehendes, wahrnehmungsgelitetes Handeln dieses Bewusstseins hindeutet (vgl. Lanwer i.Dr.). Im Rahmen eines gewöhnlichen Telefongesprächs sind Rückmeldesignale deshalb von besonderer Relevanz, da diese die Aufrechterhaltung des aufgrund der mangelnden visuellen Zugänglichkeit fragilen geteilten Wahrnehmungsraumes (Abbildung 1c) indizieren. Es ist wichtig, sich dies noch einmal klar zu machen, um den Unterschied zwischen konventioneller Telefonie und Videotelefonie adäquat fassen zu können: Videotelefonie ist nicht einfach eine Form

der Telefonie mit Bild, sondern eine Form medial-vermittelter Interaktion, die – anders als die mediatisierte Interaktion am Telefon – die FtF in ihrer konstitutiven Kopräsenz ‚imitiert‘ und damit eine Form virtueller Kopräsenz ermöglicht.

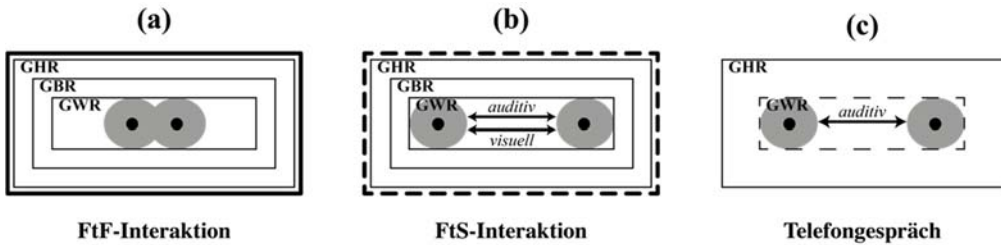


Abbildung 1: Interaktionskonstitution unter verschiedenen Bedingungen der wechselseitigen Zugänglichkeit

Es lassen sich jedoch verschiedene Besonderheiten einer solchen virtuellen Kopräsenz ausmachen, die für den spezifischen Zuschnitt eines Interaktionsraums in der FtS kennzeichnend sind: Die Akteure können gewissermaßen den Blick des Gegenübers lenken bzw. beeinflussen. In einer bspw. dyadischen FtS kann die Kamera, die das Sichtfeld von Akteur A in den (physischen) Raum von Akteur B erweitert, in ihrer Ausrichtung (nur!) durch Akteur B manipuliert werden und vice versa. Die Akteure können somit wechselseitig manuell (zumindest indirekt) auf die Blickausrichtung des jeweils anderen einwirken und auf diese Weise geteilte visuelle Aufmerksamkeitsfoki in einer Weise etablieren und steuern, wie es in der FtF nicht möglich ist. Dies kann bspw. eingesetzt werden, um Wahrnehmungsobjekte der visuellen Koordination zugänglich zu machen (Licoppe/Morel 2012:413–417), und so die perspektivisch Begrenztheit des geteilten Wahrnehmungsraums zu kompensieren.

Die Mediatisierung der visuellen Informationen führt allgemein zu einer nicht unwesentlichen Reduktion des Wahrnehmbaren, die sich vor allem auch darin niederschlägt, dass die Interagierenden zumeist selbst (wenn überhaupt) nur ausschnitthaft sichtbar sind. Häufig sind allein Kopf und oberer Torso der Wahrnehmung des jeweils anderen visuell zugänglich. Licoppe/Morel (2012) sprechen daher auch von einem *talking heads arrangement* (vgl. Abbildung 2). Diese ‚fragmentarische Präsenz‘ muss außerdem fortlaufend sichergestellt werden. Die Interagierenden müssen stets die eigene Positionierung in Relation zur Ausrichtung der Kamera regulieren, um sich visuell verfügbar zu halten. Im Falle der Videotelefonie ist aufgrund der physischen Trennung der Wahrnehmungsräume der beteiligten Akteure ‚[d]ie Herstellung wechselseitiger Sichtbarkeit [...] nicht selbstverständlich [...], sondern [sie hängt ab] vom korrekten Einsatz der Technik‘ (Mondada 2010:299), die die getrennten Ökologien verbindet. Diese Relevanz der Regulierung der eigenen Sichtbarkeit manifestiert sich darin, dass Interagierende, wie es Licoppe/Morel (2012:426) beobachten können, gewissermaßen darauf verpflichtet *on screen* zu sein, soweit dies nicht lokale Zeigepraktiken mittels Manipulation der Kamera konterkariert. Das Monitoring der eigenen Sichtbarkeit wird in

der FtS letztlich in stärkerem Maße als in der FtF zu einer interaktionsrelevanten Aufgabe, die von den Gesprächsteilnehmern fortlaufend bearbeitet werden muss, um einen virtuellen Interaktionsraum für multimodale Interaktion in seiner Konstitution überhaupt sicherstellen und aufrechterhalten zu können (vgl. hierzu auch Lanwer i.Dr.).



Abbildung 2: Talking heads arrangement

Die Konstitution eines virtuellen Interaktionsraumes ist in der FtS also generell fragiler als in der FtF (in Abbildung 1b durch die gestrichelte Linie angedeutet). Hinweise darauf, dass sich bspw. eine mangelnde Kontrolle über die eigene Sichtbarkeit in einer ausbleibenden Nutzung visueller Displays niederschlagen kann, finden sich bspw. bei Heath/Luff (1993). In diesem Umstand manifestiert sich sehr deutlich, dass ein virtueller Interaktionsraum als ein geteiltes Hier für multimodale Interaktion nicht einfach medientechnisch hergestellt wird, sondern dass dieser Raum Resultat einer sozialen Konstruktionsleistung ist, die zudem nur unter bestimmten Bedingungen der reflexiven Zugänglichkeit einigermaßen stabil erfolgen kann. Web-basierte Dienste zur Durchführung von Videotelefonaten bieten daher in der Regel eine Monitoringfunktion an, die es ermöglicht, das Kamerabild der eigenen Aufnahmeeinheit in einem Zusatzfenster auf dem eigenen Bildschirm anzuzeigen. Diese Option eröffnet eine Wahrnehmung des selbst aus der Perspektive des anderen, was letztlich zu einer speziellen Form der Reziprozität der Perspektiven führt. Es werden jedoch freilich nur die aus der Perspektive des jeweiligen Gegenübers verfügbaren Sinnesinformationen in die Sphäre der eigenen Wahrnehmung hineinprojiziert, aber nicht das, was das jeweilige Gegenüber ‚tatsächlich‘ sieht. Die Reziprozität der Perspektiven bleibt also auch hier lediglich eine ‚nützliche Hypothese‘ (vgl. Lanwer i.Dr.). In die Hypothesenbildung können jedoch Informationen über die dem jeweils anderen mutmaßlich zugänglichen Sinnesinformationen einbezogen werden, die in der FtF-Situation in der Form per se nicht zur Verfügung stehen.²

Wird ein Interaktionsraum im Medium der reflexiven Wahrnehmungswahrnehmung erst einmal hergestellt, können schließlich körperliche Ressourcen zur intersubjektiven Koordination (Depperman/Schmitt 2007) eingesetzt werden. Was das Blickverhalten

² Die Monitoringfunktion birgt zugleich ein gewisses Ablenkungspotenzial, da diese auch als Spiegel genutzt werden kann, was bspw. dazu führen kann, dass sich die Akteure auffällig oft die Frisur richten o.ä. Zugleich können aufgrund der Basis der reflexiv zugänglichen visuellen Informationen Gesprächsthemen initiiert werden, die die Äußerlichkeiten der beteiligten Akteure betreffen (vgl. hierzu Abschnitt 4.2.7).

anbetrifft, ist dabei zu beachten, dass es in der FtS aufgrund der technischen Rahmenbedingungen nicht in gleicher Weise wie in der FtF zur Etablierung von *mutual gaze* (im Sinne eines wirklichen Blickkontakts) kommen kann. Dies würde voraussetzen, dass beide Akteure in die Kameralinse ihrer Aufnahmeeinheit schauen, ohne dabei den Blickkontakt zum Bildschirm zu verlieren (Arminen/Licoppe/Spagnolli 2016:298), was bei derzeitigem Stand der technischen Entwicklung im Rahmen der privaten Nutzung von Computern oder Mobilgeräten als Kommunikationstechnik nicht möglich ist. Dennoch scheint das Blickverhalten der Akteure – wie es die Analysen in Abschnitt 4 dokumentieren – in der FtS nicht irrelevant für die interaktive Hervorbringung bspw. von Erzählsequenzen zu sein.

3. Blickverhalten in konversationellen Erzählungen

Für das Gelingen multimodaler Interaktion ist es von Bedeutung, dass die beteiligten Akteure sich in Bezug auf ihr (körperliches) Verhalten räumlich wie zeitlich koordinieren (Deppermann/Schmitt 2007). Der Aspekt der räumlichen Koordination betrifft in der FtS – wie wir gesehen haben (vgl. Abschnitt 2) – vor allem die Organisation der reflexiven Sichtbarkeit, d.h. die Herstellung bzw. Aufrechterhaltung eines geteilten Wahrnehmungsraumes. Die zeitliche Koordination betrifft hingegen primär die Herstellung eines geordneten und in seiner sequenziellen Ordnung sinnhaften Handlungs-geschehens. Eine solche Koordination ist entsprechend auch für die kooperative Realisierung einer konversationellen Erzählung von zentraler Bedeutung: Die Einbettung eines ausgedehnten Erzählstrangs in den alternierenden Zug-um-Zug-Verlauf eines Gesprächs ist an erster Stelle nicht primär ein Problem der Organisation des Rederechts, sondern zunächst ein Problem der Koordination von Handlungsplänen im Rahmen der Hervorbringung eines größeren Handlungspakets im Sinne eines *joint project* (Clark 1996:191–220). Die Realisierung eines *story preface* signalisiert nicht nur die Absicht einer zeitlich ausgedehnten Beanspruchung des Rederechts, sondern die Absicht zur Realisierung einer Erzählhandlung sowie daran gekoppelte Erwartungen hinsichtlich des für die Handlung im Vollzug konstitutiven Partizipationsverhaltens der Koaktanten (vgl. C. Goodwin 1984; Sacks 1971). Von den Rezipienten einer Erzählung ist während des *telling* nicht pauschal ein Verzicht auf das Rederecht zu erwarten, sondern vor allem ein erzählspezifisches Rückmeldeverhalten (Schegloff 1982), welches überhaupt erst die Einnahme der Rolle eines Rezipienten anzeigt und damit den aktiven Sprecher in seiner dazu komplementären Rolle als Erzähler ratifiziert. Was für den Erzähleinstieg gilt, kann in gleicher Weise auch für den Ausstieg geltend gemacht werden. Hier kann vom Erzähler bspw. ein *display of story completion* (Jefferson 1978:231) eingesetzt werden, um im sequenziellen Verlauf das Ende des *telling* erkennbar zu machen und damit eine evaluative Positionierung des/der Koaktanten zum Erzählgeschehen zu evozieren (Stivers 2013:201), die ein Verstehen der Narration in Rückbezug auf das im *story preface* realisierte Erzählversprechen anzeigt. Für die kooperative Hervorbringung dieser einzelnen Handlungspakete sind entsprechend lokal ausgehandelte reflexive Passungen im Partizipationsverhalten konstitutiv.

Die an einer Interaktion beteiligten Akteure sind allgemein (d.h., nicht nur in Erzählsequenzen) fortlaufend damit befasst, sich wechselseitig spezifische Partizipationsrollen (*participant roles*) anzuzeigen bzw. diese interaktiv auszuhandeln und auf diese Weise den für die jeweilige Handlung im Vollzug konstitutiven Beteiligungsrahmen im Sinne eines *participation framework* (C. Goodwin 1981; C. Goodwin 2000; M.H. Goodwin 1997; C. Goodwin/M.H. Goodwin 2004; Levinson 1988) herzustellen. Zahlreiche Untersuchungen vor allem von C. Goodwin liefern empirische Evidenz dafür, dass entsprechende Beteiligungsrahmen bei Kopräsenz der Akteure immer auch unter Rückgriff auf körperliche Ressourcen, wie Proxemik, Körperpostur aber bspw. auch das Blickverhalten, hergestellt werden (C. Goodwin 1981; C. Goodwin 1984; C. Goodwin 1986; C. Goodwin 2000; C. Goodwin 2007 u.a.). C. Goodwin (2000:1499) spricht daher auch von einem *embodied participation framework*. Indirekte Evidenz für die Relevanz des Blickverhaltens hinsichtlich der multimodalen Herstellung eines solchen Beteiligungsrahmens liefern Studien, die dokumentieren, dass bestimmte Muster im Blickverhalten systematisch mit der Einnahme der Sprecher- bzw. Hörerrolle einhergehen (für einen Überblick vgl. Rossano 2012:54–56, 2013). Der Unterschied zwischen dem Blickverhalten von Sprecher und Hörer liegt darin, dass Hörer den aktiven Sprecher häufiger und ausdauernder fokussieren als dies umgekehrt der Fall ist. Dies scheint insbesondere darauf zurückführbar, dass prospektive Sprecher häufig spätestens zum Zeitpunkt der Turnübernahme den Blick von ihren Koaktanten abwenden und sich als aktive Sprecher dann erst gegen Turnende wieder auf diese zurückorientieren. Hörer fokussieren hingegen nicht selten mehr oder weniger über den gesamten Turnverlauf den aktiven Sprecher. Das sprecherseitige Abwenden des Blickes zum Turnbeginn dient – so Kendon (1967:35–36) – der kognitiven Entlastung im Rahmen der Äußerungsplanung, die Refokussierung auf das Gegenüber zum Turnende dagegen der Organisation der Übergabe des Rederechts. Eine etwas anders gelagerte Deutung findet sich hingegen bei C. Goodwin (1981:55–94), der die hörerseitige Blickorientierung auf den Sprecher als *participation* bzw. *engagement display* beschreibt, welches die Einnahme einer Hörerrolle anzeigt und damit den aktiven Sprecher als solchen legitimiert. Das sprecherseitige Abwenden des Blickes zu Turnbeginn erfüllt – so C. Goodwin – den Zweck, dem/den Koaktanten Zeit einzuräumen, die für die prospektive Hörerrolle konstitutive Ausrichtung des Blickes vorzunehmen, bevor die Blickzuwendung des prospektiven Sprechers erfolgt. Geschieht dies nicht, wird, wie es C. Goodwin (1981:77–86) empirisch aufzeigen kann, die Blickzuwendung einzelner Koaktanten regelmäßig durch Reparaturen im verbalen Kanal forciert und auf diese Weise – so die Interpretation des Befundes – der interaktive Beteiligungsrahmen in seinem multimodalen Zuschnitt (wieder-)hergestellt.

In Bezug auf konversationelles Erzählen ergibt sich dabei die Besonderheit, dass Hörer in bestimmten Sequenzumgebungen (in dem bei Goodwin behandelten Beispiel geht es um eine Seitensequenz) ihren Blick zeitweise auch ‚unsanktioniert‘ vom Erzähler abwenden. Dies erklärt C. Goodwin (1984:130–131) damit, dass einerseits im Normalfall bereits während der Realisierung des *story preface* sowie zum Zeitpunkt des

Erzähleinstiegs verbal und/oder nonverbal die Einnahmen einer Rezipientenrolle signalisiert worden ist und dass andererseits Abweichungen vom ansonsten für die Rezipientenrolle konstitutiven Blickverhalten durch alternative nonverbale Displays, wie bspw. Kopfnicken, kompensiert werden können (vgl. C. Goodwin 1986). Grundlegend kritisiert worden ist C. Goodwins Analyse von Rossano (2012), der davon ausgeht, „that gaze is organized also with respect to sequential environment rather than just participation role“ (Rossano 2012:66–67). Das empirische Fundament dieser (vermeintlichen) Gegenhypothese bildet eine Studie zum Blickverhalten von Sprecher und Hörer in zwei verschiedenen sequenziellen Kontexten. Rossano vergleicht in seiner Arbeit das Blickverhalten in Frage-Antwort-Sequenzen mit dem in *extended tellings*³ und kann feststellen, dass prospektive Rezipienten zum Zeitpunkt des Einstiegs in eine entsprechende Sequenz, den Blick früh und ausdauernd auf den prospektiven Erzähler richten, um „recognition and acceptance“ (Rossano 2012:69) einer *action sequence* zu signalisieren, während sich in Frage-Antwort-Sequenzen eine hörerseitige Blickabwendung bei sprecherseitiger Zuwendung beobachten lässt.

Dass in verschiedenen Handlungszusammenhängen unterschiedliche Muster im Blickverhalten von Sprecher und Hörer zu beobachten sind, ist allerdings einerseits keine neue Erkenntnis: Bereits Kendon (1967:45–46) stellt fest, dass Interagierende bspw. den (bisher) aktiven Sprecher zwar anschauen, wenn sie durch die Realisierung von Rückmeldesignalen lediglich Aufmerksamkeit signalisieren, ihren Blick hingegen im Rahmen evaluierender Einschübe von diesem abwenden. Andererseits finden sich in der Forschungsliteratur Beobachtungen zu handlungsspezifischen Formen des Blickverhaltens *innerhalb* von Erzählsequenzen bspw. beim Beschreiben von Ereignissen (C. Goodwin 1986) oder im Rahmen von Redeinszenierungen (Sidnell 2006). Eine konversationelle Erzählung als komplexe Handlungssequenz im Sinne eines *joint project* ist – wie es auch in den Analysen Goodwins zu Tage tritt – durch eine zeitlich emergente, dynamische Beteiligungsstruktur gekennzeichnet, die mit der sequenziellen Abfolge verschiedener für die gesamte Aktivität konstitutiver Handlungsschritte zusammenhängt (vgl. C. Goodwin 1984; C. Goodwin 1986). Die Studie von Rossano blendet diesen Aspekt systematisch aus, indem lediglich der Moment des Einstiegs in eine Erzählsequenz in den Blick genommen wird. Rossano kann daher zwar durchaus beobachten, dass der Erzähleinstieg rezipientenseitig durch *sustained gaze* ratifiziert wird. Es bleibt jedoch unbeleuchtet, wie lange und vor allem an welchen Stellen im Verlauf einer Erzählung die rezipientenseitige Blickzuwendung aufgelöst und an welchen wiederaufgenommen wird. Gleiches gilt für die Blickorganisation des Erzählers.

Viel entscheidender scheint jedoch, dass Partizipationsrahmen ebenso wie Handlungstypen in ihrer multimodalen Komplexität wohl generell nicht allein in Bezug auf das Blickverhalten beschrieben werden können, auch wenn dieses mit konstitutiv sein mag. Es drängt sich daher die Frage auf, welchen spezifischen Beitrag das Blickverhalten zur Handlungskoordination in Erzählsequenzen leistet. Instruktiv sind hier die

³ Bei *extended tellings* handelt es sich – so Stivers (2013:200) – in der Regel um *storytellings*.

Überlegungen von Streeck (2014), der in Bezug auf verschiedene Studien zum Blickverhalten in der Interaktion zu der Hypothese gelangt, dass die Herstellung von Blickkontakt eine Art Mechanismus zur reflexiven Quittierung intersubjektiv koordinierter kognitiver Aufmerksamkeit darstellt und *mutual gaze* somit zentral für die Erzeugung sozialer Realität ist: „gaze is understood [...] as a minimal (and often sufficient) ‚contractual‘ interaction format by which an individual act is explicitly ratified as a social fact“ (Streeck 2014:51). Für die Analyse narrativer Sequenzen ist diese Überlegung von besonderer Bedeutung, da – wie oben angesprochen – nicht allein entscheidend ist, dass der prospektive Erzähler den *floor* behaupten kann, sondern vor allem auch, dass die Koaktanten eine das Erzählgeschehen (kognitiv) aktiv mitkonstruierende Haltung einnehmen und anzeigen. Hörerseitig regelt „[d]ie Wahl zwischen Anschauen und Wegblicken [...] [allgemein] die Kopartizipation“ (Weiß/Auer 2016).

In Bezug auf das Blickverhalten in Erzählungen wird diese Hypothese durch eine Studie von Bavelas/Coates/Johnson (2002) flankiert, auf die sich auch Streeck in seiner Argumentation bezieht. Bavelas/Coates/Johnson analysieren in einer eher quantitativ ausgerichteten Untersuchung zum Blickverhalten von Erzähler und Rezipient im Rahmen konversationeller Erzählungen in dyadischen Interaktionen Momente der Herstellung von *mutual gaze*. In der Diskussion ihrer Ergebnisse charakterisieren Bavelas/Coates/Johnson die Herstellung von *mutual gaze* als Etablierung eines ‚Blickfensters‘ (*gaze window*), das für die Regulierung des verbalen Rückmeldeverhaltens von zentraler Bedeutung ist. Bavelas/Coates/Johnson konzentrieren sich in ihrer Analyse jedoch nur auf Momente der bestehenden Koorientierung aufeinander. Die Autoren ignorieren gänzlich den zeitlich emergenten Charakter der Herstellung und Auflösung von *mutual gaze* ebenso wie die potenzielle Unabhängigkeit der Blickorganisation der einzelnen Akteure. Es bleibt bspw. unklar, ob Rezipienten in den untersuchten Daten den jeweiligen Erzähler durchgängig ‚anstarren‘, oder ob diese ihren Blick zuweilen auch abwenden, wie es Goodwin für eine Seitensequenz beobachten kann und wie es bspw. auch Kendons Beobachtungen zu evaluativen Einschüben vermuten lassen würden. Es bleibt damit generell offen, wie die Interagierenden ihr Blickverhalten im Rahmen der sequenziellen Hervorbringung einer konversationellen Erzählung koordinieren, um auf diese Weise wiederum die Sequenzierung der Narration selbst mit zu konstituieren. Soll eine Erzählung als eine Art *embodied joint project* beschrieben werden, dessen sequenzielle Hervorbringung durch körperliche Praktiken mitbestimmt ist, wie es die Studien von Goodwin und Rossano vermuten lassen, muss jedoch gerade diese Reflexivität des betreffenden Kontextualisierungsverfahrens in den Fokus geraten. Anders gewendet bedeutet dies, dass der Nachweis entsprechender Kontextualisierungspraktiken in Videotelefonaten als Evidenz für die Herstellung und Nutzung eines virtuellen Interaktionsraumes gewertet werden kann.

4. Multimodales Erzählen in Skype-Gesprächen

Gegenstand der folgenden Analyse ist eine Erzählsequenz im Rahmen eines Videotelefonats, das mit der Software Skype geführt worden ist.⁴ In einer explorativ-exemplarischen Analyse soll das Blickverhalten der Interagierenden im Verlauf der Erzählsequenz in seiner zeitlich-sequenziellen Entwicklung nachgezeichnet werden. Auf diesem Wege soll analysiert werden, ob die Interagierenden die Affordanzen der spezifischen Interaktionsökologie dazu nutzen, eine konversationelle Erzählung als *embodied joint project* multimodal herzustellen. Da im Grunde analytisch geprüft werden soll, ob die Interagierenden im Rahmen der Erzählsequenz überhaupt in einem virtuellen Interaktionsraum agieren, wird bei der Analyse zunächst eine strikte Trennung zwischen der Analyse des verbalen Verhaltens einerseits und der Analyse des Blickverhaltens andererseits vorgenommen. Ziel dabei ist es, zu prüfen, ob das Blickverhalten in seiner zeitlichen Organisation ebenso wie in der wechselseitigen Abstimmung in einer Weise auf den anhand des verbalen Verhaltens beschreibbaren sequenzstrukturellen Aufbau der Narration bezogen ist, dass davon ausgegangen werden kann, dass das Blickverhalten als Ressource zur Herstellung eben dieses Aufbaus von den Teilnehmern genutzt wurde. Im Fokus steht daher die Analyse der sequenziellen Platzierung von Phasen der Koorientierung der Akteure aufeinander in Relation zur verbalen Entfaltung der Erzählung. Unter den technischen Bedingungen der Bildtelefonie kann es jedoch, wie in Abschnitt 2 bereits ausgeführt wurde, nicht zu *mutual gaze* im eigentlichen Sinne kommen. Für die Analysen sind daher Phasen relevant, in denen sich die Akteure wechselseitig in einer Art auf die Bildschirmrepräsentation des jeweils anderen orientieren, dass davon auszugehen ist, dass beide Akteure die Blickzuwendung des anderen jeweils zur Kenntnis nehmen. Entsprechende Phasen lassen sich auch als eine Art *mutual screen-gaze* charakterisieren und sind in den untersuchten Daten in der Regel anhand von Blickorientierung und Kopfausrichtung gut identifizierbar.⁵ Im Folgenden wird in Bezug auf entsprechende Phasen in Anlehnung an Bavelas/Coates/Johnson (2002) ebenfalls von der Etablierung eines Blickfensters gesprochen. Eine wichtige Voraussetzung für die Herstellung eines Blickfensters ist – wie in Abschnitt 2 ausgeführt – die Kontrolle der eigenen Sichtbarkeit. Im Falle von Videotelefonie mit Skype, wie sie hier untersucht wird, kann die eigene Sichtbarkeit optional mithilfe eines Monitoringfensters sichergestellt werden (vgl. hierzu auch Lanwer i.Dr.). In den erhobenen Daten machen alle Gesprächsteilnehmer von dieser Option zumindest (technisch gesehen) Gebrauch.

⁴ Die Daten sind im Rahmen des durch das „Programm zur Förderung des exzellenten wissenschaftlichen Nachwuchses“ der Universität Duisburg-Essen geförderten Projektes „Aufbau einer Forschungs- und Lehrdatenbank mit Skypekommunikation“ im Zeitraum von März 2015 bis Februar 2016 erhoben und korpuslinguistisch aufbereitet worden.

⁵ Die sich dabei ergebenden Probleme bzgl. der Frage, ob die Akteure tatsächlich das Abbild des Gegenübers auf dem Monitor fokussieren oder an diesem vorbei oder gar durch dieses hindurchschauen, sind analytisch gesehen, die gleichen, wie sie sich ohne den Gebrauch von Eye-Tracking-Apparaturen im Rahmen der Untersuchung von FtF-Settings ergeben und stellen damit kein spezifisches Problem der Analyse von FtS-Settings dar, das hier gesondert behandelt werden müsste resp. könnte.

4.1. Daten und Transkription

Die analysierte Beispielsequenz entstammt einem Korpus von insgesamt elf Aufzeichnungen von mit der Software Skype (www.skype.com) durchgeführten Videotelefonaten, die bis auf eine Ausnahme alle dyadisch angelegt sind. An dem untersuchten Gespräch sind zwei männliche Akteure beteiligt (im Folgenden A und B). A befindet sich zum Zeitpunkt des Skype-Telefonats in seiner Wohnung in Maribor (Slowenien), wo er sein Auslandssemester verbringt. Akteur B, der ältere Bruder von B, sitzt in seinem Büro in einer westfälischen Stadt, wo er von A angerufen wurde. Das Gespräch ist mittels der Mitschnittsoftware Evaer (www.evaer.com) auf dem Computer von A aufgezeichnet worden. Die Software ermöglicht es, auf einem der beteiligten Geräte stationär das Ton- und Bildsignal eines Skype-Gesprächs mitzuschneiden. Die Ausgabe des visuellen Kanals erfolgt im *Split-Screen*-Format, wie es das Datenmaterial in der Beispielanalyse zeigt. Das Audiosignal wird hingegen in einer Tonspur zusammengefasst. Da bei der Datenübertragung mit gewissen, wenn auch zumeist geringen, Latenzzeiten zu rechnen ist, hat der Mitschnitt mit Blick auf die Alignierung von Bild und Ton eine Beschaffenheit, die in dieser Form nur dem Akteur, auf dessen Gerät die Aufzeichnung vorgenommen worden ist, exakt entspricht. In der analysierten Gesprächssequenz lassen sich jedoch keine latenzbedingten Irritationen im Gesprächsverhalten erkennen, sodass davon auszugehen ist, dass die Latenzen in der Übertragung äußerst gering ausgefallen und beidseitig unbemerkt geblieben sind.

Das Korpus umfasst insgesamt ca. sechs Stunden vollständig transkribierter AV-Daten. Die Rohdaten sind mit dem EXMARaLDA Partitur Editor (exmaralda.org) zeitaligniert und nach den GAT2-Konventionen (Selting et al. 2009) in Form eines Minimaltranskripts verschriftlicht worden. Der im Folgenden analysierte Gesprächsausschnitt wurde für die vorliegende Untersuchung in Form eines ausgebauten Basistranskripts nachtranskribiert und um eine Notation des Blickverhaltens angereichert. Das verwendete Transkriptionssystem bezieht sich in seiner Systematik auf C. Goodwin (1981) sowie Mondada (2014 u.a.). Hier soll daher neben ein paar grundlegenden Erläuterungen zur Transkriptionsweise nur kurz auf einige Besonderheiten in der Darstellung eingegangen werden: Die nonverbale Notation erfolgt in Bezug auf das Verbaltranskript interlinear. Veränderungen in der Blickausrichtung werden mit gepunkteten Linien, Phasen der Beibehaltung einer Blickrichtung mit gestrichelten Linien notiert. Zwecks Notation der genauen Ausrichtung des Blicks wurde grobschematisch einerseits zwischen Zuwendung und Abwendung des Blicks in Bezug auf den jeweiligen Computerbildschirm unterschieden. Andererseits wird bei Blickabwendung zwischen nach links (L), nach rechts (R), nach unten (U)

oder nach oben (O) ausgerichtetem Blick differenziert. Die Fokussierung des Bildschirms wird mit *B* kodiert. In Fällen, in denen die Akteure auf dem Bildschirm speziell das Monitoringfenster anvisieren (was in der Regel aufgrund einer peripheren Blickausrichtung (wie in Abbildung 9.2) gut ersichtlich ist), wird ein *M* verzeichnet. Phasen der Bildschirmzuwendung werden im nonverbalen Transkript durch Fettung typographisch hervorgehoben. Die zeitliche Alignierung zwischen nonverbalem und verbalem Transkript wird durch Gliederungsmarkierungen angezeigt (# im verbalen und > im nonverbalen Transkript). Abschnitte im Verbaltranskript, die mit Phasen beidseitiger Fokussierung des Bildschirms, d.h. mit einem Blickfenster zusammenfallen, werden farbig hinterlegt, wie es das angeführte Beispiel zeigt:

```
070  A:      °hh#h #un_#dann k#Am der RAU:S-#
      ---->..>B-->.....>U----->
      B:      ---->..>B----->
```

Die hier abgedruckte Zeile ist entsprechend so zu lesen, dass (in der Aufzeichnung) beide Aktanten, während A einatmet, ihre bisherige Blickrichtung zeitgleich auflösen, und sich synchronisiert mit der Silbe *un* [ʊn] der verbalen Äußerung jeweils ihrem Computerbildschirm zuwenden. B behält im Weiteren die Blickrichtung zum Bildschirm bei, während sich A wieder abwendet und zeitgleich mit der Silbe *kam* [ka:m] nach unten schaut. Die auf diese Weise angefertigte Transkription liefert für den untersuchten Gesprächsausschnitt eine vollständige Notation der Blickausrichtung beider Interaktionsteilnehmer in ihrem Verhältnis zu den jeweils realisierten verbalen Äußerungen.

4.2. Exemplarische Analyse einer Erzählsequenz

Die analysierte Erzählsequenz hat eine Dauer von insgesamt 1:21 Min. Der verbale Anteil des Erzählers beträgt 1:16 Min., der des Rezipienten lediglich 0:06 Min. Die Redeanteile des Rezipienten bestehen ausschließlich aus kurzen evaluativen Einschüben, wie *ja hAt_a dich verARSCHT*; (Z. 101). oder *°hh jA ↑SCHÖN*; (Z. 108). Außerdem finden sich rezipientenseitig silbische Lacher (0:02 Min.). Die Einbeziehung visueller Informationen in die Datenanalyse liefert daher – wie sich zeigen wird – einen Zugang zum Material, der vor allem in Bezug auf das beobachtbare Rezipientenverhalten einen deutlich detaillierteren Blick eröffnet. Ziel der folgenden Analyse ist es jedoch – wie oben bereits angedeutet –, die Systematik der Blickorganisation der beteiligten Akteure in ihrer wechselseitigen Orientierung sowie in ihrer zeitlichen Entfaltung über den gesamten Verlauf der narrativen Sequenz und in Bezug auf ihre sequentielle

Organisation zu beschreiben. Die Analyse orientiert sich dabei an der folgenden grobschematischen Gliederung in einzelne Sequenzabschnitt:

1. Sequenzvorlauf und thematischer Auslöser
2. *Story preface*
3. Erzähleinstieg
4. Ausbau der szenischen Rekonstruktion
5. Seitensequenz und Vorbereitung der Pointe
6. Darbietung der Pointe
7. Evaluation und Sequenzausstieg

4.2.1. Sequenzvorlauf und thematischer Auslöser

Der narrativen Sequenz geht eine Unterredung über mögliche Urlaubsziele für den kommenden Winter voraus, in der Akteur B bekundet, ggf. für einen Skiurlaub noch einmal nach Maribor reisen zu wollen, wo er sich zum Zeitpunkt des Gesprächs aufgrund eines Auslandssemesters aufhält. Akteur A erwägt, sich mit B für die Reise ggf. zusammenzuschließen, und erkundigt sich im weiteren Gesprächsverlauf zwecks Evaluation des Reiseziels nach dem dortigen Preisniveau.

```

001 B: SCHAUN_wa #ma;#
      <B.....>U-->
      A: <B----->
002 B: SONST\#=
      ----->
      A: ----->
003 B: =jA: ich kUck ma wAt man so MACHen kann;#
      ----->
      A: ----->
→ 004 B: slowEnien is wahrscheInlich GÜNstiger auch dAnn
      ne,#
      -----
      --->
      A: -----
      ---->
005   (0.3) #
      A: ----->
      B: ----->

```

```

→ 006 A: °h#h (.) JA.#
      ----->
      B: -->.....>
007   (0.3)#
      A: ----->
      B: B----->
→ 008 A: Also: wEnn_e hie#:#:r (.) ä#:#:h RAUSgehs-#
      ----->.....>R----->
      B: ----->
009 A: da #kOst_n #BIE::R;#
      ----->
      B: ---->.....>R----->
010   (0.2)#(0.8)#(0.3)#
      A: ----->.....>B----->
      B: ----->
011 A: im CLUB,#=
      ----->
      B: ----->
012 A: =zwEI euro #SIEbzi#g nullfÜnf-#=
      ----->
      B: ----->.....>B----->
      Abb                                #3-1

013 A: =NE,#
      ----->
      B: ----->
014   (0.1)#(0.6)#
      A: ----->
      B: ----->.....>
→ 015 B: BOAH.#
      .....>
      A: ----->
      Abb                                #3-2
016 A: und wEnn_e in_e KNEIpe sitzt-#=
      ----->
      B: U----->
017 A: =ZA:HLs_e #Oft schon für nullfÜnf nur zwei
      ----->
      zwei DREI:ßig;#
      ----->
      B: .....>B----->
      ----->
      Abb                                #3-3

```

```

→ 018 A: AL:so:.#
      ----->
      B: .....>
      Abb          #3-4
→ 019 B: dAt_is GEI#L;#
      U----->
      A: ----->...>
020   (0.3) #
      A: .....>
      B: ----->
021 B: JA.#
      ---->
      A: R-->
022 A: JA.#
      ---->
      B: ---->

```

Nachdem A in Zeile 001 mit *SCHAUN_wa ma*; zunächst potenziell den Abschluss seiner Überlegungen bzgl. der Urlaubsplanung anzeigt, evoziert er mit *slowEnien is wahrschEinlich GÜNstiger auch dAnn ne*, (Z. 004) eine Vertiefung der Evaluation des anvisierten Reiseziels. Die Äußerung stellt hier zunächst den ersten Strukturteil einer Frage-Antwort-Sequenz dar, deren zweiten Strukturteil Akteur B nach einer kurzen Pause von 0.3 Sek. mit der Antwortpartikel *JA*. mit tief-fallender Kontur liefert. Es handelt sich hier jedoch – wie der weitere Sequenzverlauf zeigt – um einen besonderen Typus von Paarsequenz, den Schegloff (2007:170) als *topic-proffering sequence* bezeichnet. Die Spezifik des Sequenztyps liegt darin, dass der erste Strukturteil der Paarsequenz als thematisches Angebot fungiert. Eingeleitet werden entsprechende Sequenzen häufig – wie auch in diesem Fall – mit Ja-Nein-Fragen. Die Realisierung einer entsprechenden Frage zielt nicht primär auf eine polare Antwort ab, sondern auf eine inhaltlich ausgebaute Reaktion des adressierten Aktanten, dem durch die sequenzeröffnende Ja-Nein-Frage ein Wissensvorsprung in Bezug auf das durch den aktiven Sprecher eingebrachte Gesprächsthema attestiert wird, woraus ein gewisser Zugzwang resultiert (siehe hierzu auch König 2017). Nach einer erneuten Pause von 0.3 Sek., in (Z. 007), löst B diesen Zugzwang mit *Also: wEnn_e hie:r (.) ä:h RAUSgeh's*– (Z. 008) ein, indem er u.a. durch das mit Dehnung auf dem zweiten Vokal realisierte explikative *Also:* (Deppermann/Helmer 2013) die Bereitschaft zur Ausdifferenzierung seiner Antwort signalisiert. Im weiteren Verlauf entfaltet sich ein Wechselspiel zwischen von A gelieferten Informationen über Bierpreise in verschiedenen Etablissements und darauf bezogenen Evaluationen durch B. Der Sequenzabschluss erfolgt durch die prosodisch replikative Realisierung der Affirmationspartikel *JA*. (Z. 021 bzw. 022)

als *stance alignment* (Du Bois 2007:142–145) im Hinblick auf eine affektiv-positive Evaluation des Preisniveaus in Slowenien.



Abbildung 3: Bildstrecke Z. 012–019

Was das Blickverhalten anbetrifft, lässt sich hier zunächst das von Rossano für Paarsequenzen beschriebene Muster in allerdings invertierter Form beobachten: B, der den ersten Strukturteil der Paarsequenz realisiert, hat seinen Blick während der initialen IP in Zeile 004 noch nicht auf A gerichtet, während dieser bereits deutlich vor Turnübernahme Akteur B fokussiert. B erwidert die Blickzuwendung erst im Moment der Turnübernahme durch A in Zeile 006, wodurch es zu einer kurzzeitigen Etablierung eines Blickfensters kommt. Dieses wird jedoch seitens A noch in der laufenden IP wieder aufgelöst. Im Verlauf der Folge-IP *da kOst_n BIE::R*; wendet auch B seinen Blick seitlich ab. Während A die Orientierung auf das Gegenüber in Zeile 10 bereits in *Mid-Turn*-Position wieder herstellt, bleibt der Blick von B bis kurz vor Einlösung der syntaktischen und semanto-pragmatischen Projektionen gegen Turnende vom Bildschirm abgewandt. Erst kurz vor Turnende wendet dann auch Akteur B seinen Blick wieder dem Bildschirm zu. Die von B im Anschluss an die von A hervorgebrachte *tag question NE*, (Z. 013) sowie nach einer kurzen Pause von 0.7 Sek. realisierte Evaluation *BOAH*. (Z. 015) fällt dennoch nicht in ein Blickfenster, da B kurz vor Beginn der verbalen Äußerung die Blickorientierung auf den Monitor bereits wieder auflöst. B wendet praktisch mit Beginn der Realisierung seines evaluativen Turns den Blick nach rechts unten ab (vgl. Abbildung 3.1 bis 3.2) und kehrt mit seinem Blick erst im Verlauf des zweiten, eine Preisinformation enthaltenden Turns und *wEnn_e in_e KNEIpe sItzt= ZA:HLs_e Oft schon für nullfÜnf nur zwei ZWANzig zwei DREI:ßig*; (Z. 016–017) wieder zurück (Abbildung 3.3). Die Blickzuwendung seitens Akteur B wird allerdings erneut unmittelbar vor Evaluation der gelieferten Information wieder aufgehoben. Auch die evaluative Äußerungen *dAt is GEIL*; (Z. 019) wird von B mit einem nach rechts unten abgewandten Blick realisiert (Abbildung 3.4). Bei A lässt sich hingegen seit der Wiederaufnahme der Blickzuwendung in Zeile 010 eine durchgängige Fokussierung des Bildschirms beobachten, die erst nach Bs Evaluation der zweiten Preisinformation aufgehoben wird. Das sequenzschließende *stance alignment* findet daher in einem Fenster beidseitiger Abgewandtheit statt. Der Sequenzabschnitt fungiert insgesamt als thematischer *trigger* (Schegloff 1978) für die darauffolgend von A initiierte Narration.

4.2.2. Story preface

Auf das *stance alignment* in Zeile 021–022 folgt nach einer Pause von 0.1 Sek. mit *un_dann hAm wa hIEr so_ne POMmesbUde-* (Z. 024) ein Redebeitrag von A, der zunächst als Fortsetzung der Auflistung von Bierpreisen gedeutet werden könnte, da mit *so_ne POMmesbUde* nach *im CLUB* (Z. 011) und *in_e KNEIpe* (Z. 016) ein weiteres Etablissement eingeführt wird, dessen Bierpreise es potenziell zu evaluieren gilt. Die Fortentwicklung des Gesprächs lässt jedoch erkennen, dass hier auf die Einbettung einer narrativen Rekonstruktion hingearbeitet wird, die durch das bisherige Gesprächsthema ausgelöst worden ist:

```

023      (0.1) #
        A: <<U-->
        B: <<R-->
024 A: un_dann hAm wa hIEr so_ne POMmesbUde-#=
        ----->
        B: ----->
025 A: =hAb_i entDE#CKT;#
        ----->....>
        B: ----->
026 A: da_äh #pOmmesbu#de CURrywur#st,#
        .....>B----->.....>L-->
        B: ----->
        Abb      #4-1
027 A: un_dIE da:\#
        ----->
        B: .....>
        Abb      #4-2
→ 028 A: da hAb_ich gEstern wieder mich schön HINGesetzt-#=
        .....>
        B: B----->
029 A: =mit_der EIn DEUTschen;#
        U----->
        B: ----->
030 A: °hhhhh#
        ----->
        B: ----->
031      (0.2) #
        A: ----->
        B: ----->
032 A: hh°#
        ---->
        B: ----->
→ 033 A: ?hh° wAr SO: gEIl;#
        ----->
        B: ----->

```

Mit *hAb_i entDECKT*; in Zeile 025 wird die vorangegangene Äußerung *un_dann hAm_wa hIEr so_ne POMmesbUde*= zu einer Apokoinu-Konstruktion, genauer gesagt zu einer *modified mirror-image construction* (Betz 2008:33) ausgebaut, deren *pre*- und *post-pivot* in einer Reparaturum-Reparans-Relation zueinander stehen: Der reformulierende Teil *hAb_i entDECKT*; (*post-pivot*) modifiziert den epistemischen Status des präsentierten Wissens, indem die generische Formulierung *hAm_wa hIEr* (*pre-pivot*) durch eine Verbalphrase ersetzt wird, die den Aspekt der Exploration der neuen Lebenswelt und damit die Neuheit der Entdeckung betont. Nach zwei weiteren Reformulierungsschleifen realisiert A schließlich ein mehr oder weniger prototypisches *story preface*.



Abbildung 4: Bildstrecke Z. 026–028

Das Blickverhalten der Interagierenden verläuft dabei in hohem Maße aligniert mit der verbalen Struktur der Teilsequenz. Beobachtbar ist zum einen, dass eine kurze Phase der Fokussierung des Bildschirms durch A (Abbildung 4.1) mit der reformulierenden Äußerung *da_äh pOmmesbude CURrywurst*, (Z. 026) zeitlich koordiniert verläuft. B erwidert den Blick auf den Bildschirm an dieser Stelle jedoch nicht. In Bezug auf das Blickverhalten von B ist hingegen auffällig, dass dieser seinen Blick dem Bildschirm ziemlich exakt mit Beginn der Intonationsphrase (IP) *da hAb_ich gEstern wieder mich schön HINge setzt*– (Z.028) zuwendet (Abbildung 4.2), die mit dem Lokaldeiktikon *da* und dem Temporaladverb *gEstern* die raumzeitliche Verortung des Erzählgeschehens liefert und zusätzlich mit der Personenreferenz *ich* Akteur A als einen der Protagonisten der rekonstruierten Szene einführt. In der darauffolgende IP, die expandierend an die IP in Zeile 028 anschließt, wird mit der inkrementellen Präpositionalphrase *mit_der EIn DEUTschen*; (Z. 029), eine weitere Protagonistin etabliert. Nach einer kurzen Pause von 0.3 Sek. liefert A mit *?hh° wAr SO: gEIl*; (Z. 033) schließlich das Erzählversprechen, welches die projizierte narrative Rekonstruktion mit Blick auf den zu evaluierenden Gehalt vorab als *gEIl* charakterisiert und damit sowohl in Bezug auf das Erkennen des Erzählabschlusses als auch im Hinblick auf die Evaluation der Rekonstruktion verstehensrelevante Hinweise liefert. Während der Blick von B dem Erzähler (A) zugewandt ist, schaut dieser nach links unten aus dem Bild. Hinsichtlich des Blick-

verhaltens lässt sich insgesamt beobachten, dass sich kein einziger Moment der Koordination auf die Computerbildschirme ausmachen lässt. Das Blickverhalten lässt aber dennoch vor allem in Bezug auf die Sequenzierung der verbalen Beiträge der Interagierenden eine gewisse Systematik erkennen.

4.2.3. Erzähleinstieg

Auf das verhältnismäßig prototypische *story preface*, das Hinweise auf die Situierung des Erzählgeschehens ebenso wie auf die zu evaluierende Qualität der narrativen Rekonstruktion liefert, folgt die Überleitung zum eigentlichen *telling*:

```

→ 034 A: un_dAnn ICH so-#
      <U----->
      B: <B----->
035 A: dann: hAb_ich so #geZEIGT,#=
      ----->.....>
      B: ----->
036 A: =ja hIEr i#ch #BRAUCH #sO_ne::-#
      .....>O-->.....>U----->
      B: ----->
037 A: hier pOmmes CURrywurst,#
      ----->
      B: ----->
038 A: AUF\#
      ---->
      B: ---->
039 (0.2) #
      A: ---->
      B: ---->
040 A: auf ENglish mEINT_a,#
      ----->
      B: ----->
041 (0.7) #
      A: ---->
      B: ---->
042 A: JA-#
      --->
      B: ---->
043 A: könn_wa woh DEUTSCH sprEchen;#
      ----->
      B: ----->
      Abb #5-1
044 A: un_ICH so-#
      .....>
      B: ----->

```

```

→ 045 A: ja #dAt_is GUT,#
      ...>B----->
      B: ----->
      Abb #5-2
046 A: weil_d#a hIEr sp #sprEchen mE:ga viele DEU#TSCH
      #nE:,#
      ----->.....
      >U--->
      B: ----->.....>U-----
      >..>B>
      Abb #5-3 #5-4
047 (0.5) #
      A: ----->
      B: ----->
048 B: ?H# [M; #]
      -->[...>]U----->
049 A: [MEI#]Nt_a;#
      ---[----]----->

```

Der Erzähleinstieg wird in Zeile 034 mit der redessituierenden Phrase *un_dAnn ICH so-* (Golato 2000) als unmittelbare Eröffnung der szenischen Rekonstruktion umgesetzt. Anstelle inszenierter Rede schließt sich in Zeile 035 mit dann: *hAb_ich so geZEIGT*, jedoch zunächst eine Beschreibung gestischen Verhaltens an. Die Wiedergabe der mit der Zeigegeste verbundenen Verbaläußerung wird dann aber in den zwei darauffolgenden IPs unmittelbar nachgeliefert. Die metapragmatische Gegenrede *JA- könn_wa woh DEUTSCH sprEchen*; (Z. 042–043) wird in Zeile 040 mit dem metasprachlichen Kommentar *auf ENGLisch mEINT_a*, eingeleitet. Das Angebot des Servicemitarbeiters, die Konversation auch auf Deutsch führen zu können, wird von A in der erzählten Welt mit *ja dAt_is GUT*, (Z. 045) goutiert und in der Erzählwelt mit *weil_da hIEr_sp sprEchen mE:ga viele DEUTSCH nE:*, (Z. 046) als zunächst nicht sonderlich überraschend evaluiert. Die Information über die Verbreitung von Deutschkenntnissen im slowenischen Maribor von Akteur B in Zeile 048 mit dem glottal einsetzenden *?HM*; quitiert.



Abbildung 5: Bildstrecke Z. 043–046

Was das Blickverhalten der Akteure anbetrifft, lässt sich zunächst festhalten, dass B eine durchgängige Fokussierung auf den Bildschirm (d.h. *sustained gaze*) zeigt, während A seinen Blick fortdauernd nach unten abgewandt hat, wie es Abbildung 5.1 exemplarisch zeigt. Erst im Moment der Inszenierung der goutierenden Äußerung in Zeile 045 orientiert sich auch A wieder auf den Bildschirm seines Computers (Abbildung 5.2). Ergänzend zur raumzeitlichen Verortung und der Einführung der Protagonisten im *story preface* wird hier der Umstand für das Zustandekommen eines deutschen Gespräches in Slowenien skizziert. Im Moment der Rekonstruktion der Evaluation dieser Grundsituation in der erzählten Welt durch Akteur A wendet dieser seinen Blick wieder dem Computerbildschirm zu. Aufgrund der fortdauernden Zugewandtheit seitens Akteur B kommt es unmittelbar zur Etablierung eines Blickfensters. Relativ zu Beginn der die Möglichkeit des Deutschsprechens als verhältnismäßig gewöhnlich evaluierenden Äußerung in Zeile 046 wendet Akteur B dann allerdings bereits kurz nach Herstellung des Blickfensters seinen Blick wieder vom Bildschirm ab. B blickt zwar gegen Ende der betreffenden IP in Zeile 046 noch einmal zurück zum Bildschirm, fokussiert diesen aber nur 0.2 Sek. und wendet dann seinen Blick unmittelbar mit Quittierung der gelieferten Information wieder ab. Ursache hierfür ist vermutlich das impulsartige Geräusch, das A durch das Bündig-Stoßen eines Papierstapels verursacht, wie es auch in Abbildung 5.4 zu sehen ist. A hat hier seinen Blick jedoch wieder abgewandt, sodass es nicht erneut zu einer Koordination auf die jeweiligen Bildschirme kommt. Auch das quittierende ?HM; fällt daher nicht in ein Blickfenster.

4.2.4. Ausbau der szenischen Rekonstruktion

Nachdem die Grundkonstellation der rekonstruierten Szene dargelegt worden ist, baut Akteur A schließlich mittels einer Fortsetzung der Dialoginszenierung die szenische Rekonstruktion aus:

| | | | |
|-------|-----|--|--|
| 048 | B: | ?H#[M; #] | |
| | | [...>]U----> | |
| 049 | A: | [MEI#]Nt_a;# | |
| | | <U-[----]-----> | |
| 050 | A: | °h ja KUCK mA1,# | |
| | | -----> | |
| | B: | -----> | |
| → 051 | A: | wEnn_de jetz hier no_#ne COla #bEI nim#mst-# | |
| | | ----->.....> B--- > | |
| | B: | ----->.....> B--- > | |
| | Abb | #6-1 #6-2 | |
| 052 | A: | dann kOst_dat Alles zwar EIn euro MEHR,# | |
| | | -----> | |
| | B: | -----> | |

053 A: aber hAst_e ne cOla mit DRIN;#
 ----->
 B: ----->

→ 054 A: sa_ich<<behaucht>jA:> ne cO:l#a eigentlich NICH,#=
>R----->
 B: ----->
 Abb #6-3

055 A: =SA_ich;#
 ----->
 B: ----->

056 A: aber i_#nEhm noch wohl so_n PILS-#
 ----->.....>
 B: ----->
 Abb #6-4

057 A: MEINT_a,#
 B----->
 B: ----->

058 A: °h Oh da mAch_ich dir_n frIsch (.) gezApftes
 hElles BIER?#
 ----->
 ----->
 B: ----->
 ----->

059 (0.2)#
 A: ----->
 B: ----->

060 A: so_n schÖnes PILSken,#
 ----->
 B: ----->

061 (0.1)#
 A: ----->
 B: ----->

062 A: dAnn KANNs_#e da hIEr-#
 ----->.....>
 B: ----->

063 A: i_so jA: dann MACH dat mal;#
>
 B: ----->

Im weiteren Gesprächsverlauf inszeniert A eine Wechselrede zwischen ihm selbst und dem deutschsprechenden Imbissmitarbeiter. Gegenstand des Redewechsels ist eine Preisaktion, auf die der Mitarbeiter in der erzählten Welt mit °h ja KUCK mal, wEnn_de jetz hier no_ne COla bEI nimmst– dann kOst_dat Alles zwar EIn euro MEHR, (Zeile 050–052) aufmerksam macht. Akteur A rekonstruiert folgend eine Verhandlung

bezüglich der Möglichkeit die *COla* in dem Angebot durch *so_n PILS* (Z. 056) zu ersetzen, was – gemäß Redeinszenierung – erfolgreich verläuft, so dass A die Bestellung in der erzählten Welt mit *JA: dann MACH dat mal* (Z. 063) abschließt.

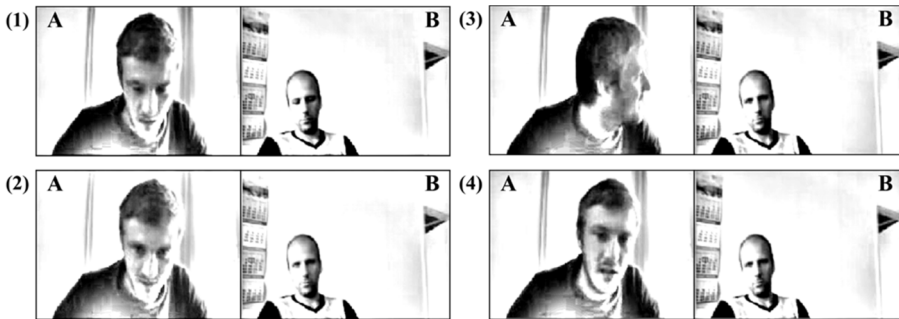


Abbildung 6: Bildstrecke Z. 049–057

In Bezug auf das Blickverhalten lässt sich zunächst beobachten, dass beide Akteure sich zum Abschluss der Turnkomponente *wENN_de jetzt hier no_ne COla bEI nimmst-* (Z. 052) zeitgleich wieder auf ihren Bildschirm zurückorientieren (vgl. Abbildung 6.1 bis 6.2), sodass ein Moment der Koorientierung aufeinander entsteht. Während B diese Blickausrichtung vorerst beibehält, alterniert die Ausrichtung des Blicks von A im Weiteren systematisch zwischen nach rechts abgewandtem Blick und Fokussierung auf den Bildschirm. Dabei wird den zwei Figuren des animierten Dialogs jeweils eine Blickrichtung zugewiesen: Die Inszenierung der eigenen Rede wie bspw. bei *sa_ich<<behaucht>JA:> ne cO:la eigentlich NICH,>* (054) erfolgt mit abgewandtem Blick (Abbildung 6.3), die Rekonstruktion der Beiträge des fiktiven Gesprächspartners ebenso wie die verbale Einleitung *MEINT_a*, (Z. 057) bei Orientierung auf den Computerbildschirm (Abbildung 6.4). Aufgrund der fortdauernden Blickzuwendung seitens Akteur B ergibt sich hieraus ein Wechselspiel zwischen Phasen, in denen ein Blickfenster etabliert wird, und Phasen einseitiger Displayfokussierung ausschließlich durch den Rezipienten.

4.2.5. Seitensequenz und Vorbereitung der Pointe

An die rekonstruierte Teilszene schließt eine Seitensequenz an, in der A bemerkt, dass er die Gelegenheit genutzt habe, um – wie er es formuliert – seine Begleiterin zum „Saufen“ zu zwingen:

```

064  A: °hhh#
      R--->
      B: <B-->
065  A: (0.5)#
      A: .....>
      B: ----->

```

- 066 A: un_dann hAb_ich für die #Andre dEUtsche noch_n
RA:Dler mItbestellt,#
.....>U-----
----->
B: -----
----->
- Abb #7.1
- 067 A: obwohl_die #gAr_nich WOLLte,#
----->.....>
B:>U----->
Abb #7-2 #7-3
- 068 A: eh HE hE#
B----->
B: ----->
- 069 A: °hh hab_i die #schö #schÖ:n zum saufen
geZWUNgen;#
----->.....>U-----
----->
B: -----
----->
- Abb #7-4
- 070 A: °hh#h #un_#dann k#Am der RAU:S-#
--->..>B--->.....>U----->
B: --->..>B----->
Abb #7-5
- 071 A: i_#so ja sA_ma wArum sprIchs_denn SO: gut
dEUtsch;#
..>B-----
----->
B: -----
----->
- 072 A: i_sa i_hAb hier sch#on VIEle dEUtsch sprechen
hÖren;#
----->.....
.....>
B: -----
----->
- 073 A: aber dann #mit so_m ÖSterrEIchischen akZENT;#
.....>R----->
B: ----->
- 074 A: oder mit so_n bIsschen SLA:wisch-#=
----->
B: ----->
- 075 A: =MEIN#t_a-#
----->.....>
B: ----->

076 A: °h#h er hätte halt vor drEI jahren Angefangen
 dEUtsches FERNsehn zu kUcken-#=
 ..>B-----
 ----->
 B: -----
 ----->

077 A: un_da hAt_a sich dat so_n bIsschen BEIgebracht;#
 ----->
 B: ----->

078 (0.1)#(0.2)#
 A: ----->.....>
 B: ----->

Die Fortsetzung erfolgt nach einem kurzen Atmer von A sowie einer Pause von 0.5 Sek. durch die Realisierung des erzähltypischen Verknüpfungssignals *und dann* in *un dann hAb_ich für die Andre dEUtsche noch_n RA:Dler mltbestellt*, (Z. 065). Zeitlich synchron mit der Realisierung der Präpositionalphrase *für die Andre dEUtsche* legt Akteur B ein Grinsen auf (Abbildung 7.1), wodurch sich B bereits vor Schilderung der genauen Rahmenbedingungen der rekonstruierten Bestellaktion von A zu dieser (potenziell) affektiv-positiv positioniert. Der Blick von A ist zu diesem Zeitpunkt jedoch nach unten abgewandt, sodass unklar ist, ob das viuselle Display von B überhaupt als solches wahrgenommen werden kann. Erst im Verlauf der Realisierung des postponierten Konzessivsatzes *obwohl die gAr_nich WOLLte*, (Z. 067), der den Umstand aufdeckt, dass seine Begleitung eigentlich kein Alkohol haben trinken wollen, wendet sich A wieder dem Bildschirm zu. Nach kurzem Lachen (Z. 068) resümiert A schließlich, dass er seine Begleitung also *schÖ:n zum saufen geZWUNgen* (Z. 069) habe. Diese Evaluation des rekonstruierten Geschehens quittiert B mit einem erneuten Grinsen (Abbildung 7.4). Es kommt damit zu einem konvergenten *stance alignment*. Durch die Evaluation wird die Teilsequenz als ‚anekdotischer Einschub‘ bzw. als *side sequence* kontextualisiert. Der verhandelte Sachverhalt wird als für sich inhaltlich abgeschlossen (Jefferson 1972:317) und in seiner konvergenten Evaluation als dem *personal common ground* der Interagierenden hinzugefügt behandelt. Dies gilt aber nicht für die narrative Sequenz als Ganzes, was sich in gewisser Weise auch im Blickverhalten der Interagierenden widerspiegelt.



Abbildung 7: Bildstrecke Z. 065–069

Der Blick von B ist zu Beginn der Nebensequenz dauerhaft auf A gerichtet, der selbst seinen Blick erst mit Einstieg in die Evaluation bzw. zum Zeitpunkt des Einsetzens des vokalen Lachens in Zeile 068 auf den Bildschirm richtet (Abbildung 7.3). Zu diesem Zeitpunkt hat B seinen Blick allerdings wieder abgewendet. Bereits mit Realisierung des Konzessivsatzes orientiert B seinen Blick nach rechts unten (Abbildung 7.2), sodass es im Rahmen der Evaluation generell nicht zur Etablierung eines Blickfensters kommt. Auch das Grinsen von B erfolgt bei nach rechts unten abgewandtem Blick. B wendet seinen Blick dem Computerbildschirm erst beim folgenden *und dann* in der IP *°hhh un_dann kAm der RAU:S-* (Z. 070) wieder zu, die eine Rückkehr zum eigentlichen Erzählgeschehen markiert (Abbildung 7.5). In diesem Moment wendet auch A seinen Blick kurz dem Bildschirm zu, sodass sich A und B in einem Blickfenster wechselseitig bzgl. der Fortsetzung der Erzählaktivität verständigen können. Diese Fortsetzung manifestiert sich u.a. auch in der Rückkehr zum Modus der Redeinszenierung. Im weiteren Gesprächsverlauf wird jedoch lediglich die Wiedergabe der eigenen Gesprächsbeiträge sowie die der Beiträge der Begleitung als direkte Rede gestaltet, während die Redewiedergabe des Imbissmitarbeiters in Form indirekter Rede wie in *°hh er hätte halt vor drEI jahren Angefangen dEUtsches FERNsehn zu kUcken-* (Z. 076) bzw. paraphrastisch wie in *un_da hAt_a sich dat so_n blsschen BEI gebracht;* (Z. 077) erfolgt. A blickt dabei erneut nur im Falle der Wiedergabe der Redebeiträge des Imbissangestellten auf den Bildschirm seines Computers, woraus sich erneut ein Wechselspiel in der Koordination ergibt. Thematisch spitzt sich der rekonstruierte Dialog auf die Frage nach dem Ursprung der Deutschkompetenz des Imbissmitarbeiters zu, der auf die Frage *wArum sprIchs_denn SO: gut dEUtsch* (Z. 071) – so A – zunächst angegeben habe, sein Deutsch durch den Konsum von deutschem Fernsehen erworben zu haben.

4.2.6. Darbietung der Pointe

Die Hinführung zur Pointe der Erzählung erfolgt schließlich durch die Rekonstruktion verschiedener affektiver Positionierungen in Bezug auf die Behauptung des Imbissmitarbeiters, sein Deutsch durch Fernsehgucken erworben zu haben:

```

079 A: ch_SO-#
      .....>
      B: <B----->
080   (0.5) #
      A: R---->
      B: ----->
081 A: du hAst dir das nUr aus_m FERNnsehn, #
      ----->
      B: ----->
082 A: dann mEInte: (.) DIE au_noch so, #
      ----->
      B: ----->
083 A: °h ja du sprIchst au gAr_nich so AS#si (.)
      lLanguage\#
      ----->.....
      .....>
      B: -----
      ----->
→ 084 A: ä:::h äh AS#si ASSi dEUtsch;#=
      B----->.....>
      B: ----->
085 A: =und dann AUCh so:\#
      U----->
      B: ----->
086 A: von Er #te EL so, #=
      ----->.....>
      B: ----->
087 A: =MEINT_a, #
      .....>
      B: ----->
→ 088 A: °h (0.3) <<veränderte Sprechstimme>>#ja ick_KANN au
      wohl Anders sprEch#en wA, #>
      ----->B-----
      ----->.....>
      B: -----
      ----->
Abb                                     #8-1
                                     #8-2
089 A: un_ICH_so, #
      R----->
      B: ----->

```

090 (0.5) #
 A: ----->
 B: ----->

091 A: °h WIE; #
 ----->
 B: ----->

092 A: er_so jA #ich kOmm aus berLIN-#
>B----->
 B: ----->
 Abb #8-3

093 A: ich bin in be in berLIN AUfgewachsen und
 <<lachend>geBORN>; #
 ----->
 B: ----->

094 A: HE HN hN #
 ----->
 B: ----->

095 B: eh hE [HE hE]
 ----->.....>U----->
 ----->.....>

096 A: [dann hAt] #der bis_er #zwAn#zig war in #
 berL#IN ge#wOhnt-#
 ----->.....>U-->....>B----->
>L>...>B----->
 Abb #8-4 #8-5 #8-6

097 A: dann hAt_er in MÜNchen noch (.) zEhn jahre
 gewOhnt-#
 ----->
 B: B----->
 ----->

098 A: un_jEtz lEBt_a in Maribor; #
 ----->
 B: ----->

099 A: h° #
 -->
 B: -->

Die Rekonstruktion der evaluativen Beiträge leitet A initial mit der lautlich reduzierten redessituierenden Phrase *ch_SO-* (Z. 079) ein. Nach einer kurzen Pause von 0.5 Sek. Schließt sich die mit klitischem *ch_* (,ich') projizierte Rekonstruktion der eigenen Rede in Zeile 081 an. Durch den Gebrauch der Fokuspartikel *nUr* in Kombination mit dem Fokusakzent auf *FERNsehen* wird eine affektive Positionierung der Überraschtheit in Bezug auf den geschilderten Erwerbsweg im Spiegel der beobachteten Sprachfähigkeiten kontextualisiert. Die folgende evaluative Sprachbewertung bestimmter Varietäten des Deutschen als *Assi*: (.) *language* bzw. *Assi dEUtsch*, die einerseits offensichtlich

mit ‚Fernsehdeutsch‘ assoziiert werden und sich andererseits nicht mit dem Sprachgebrauch des Mitarbeiters zu decken scheinen, wird durch die redesituierende Phrase *dann mEInte: (.) DIE au noch sO*, (Z. 082) der Begleitung von A zugeschrieben, sodass die Evaluation der Gesamtsituation auf zwei Schultern verteilt und damit die Erzählwürdigkeit (*tellability*) des rekonstruierten Ereignisses noch einmal betont wird. Die Verwendung von *Assi dEUtsch* wird in der Fortsetzung der Inszenierung vor allem dem Sender *Er te EL* (Z. 086) attestiert. Die Reaktion seitens des Mitarbeiters auf entsprechende Bemerkungen wird von A schließlich erneut mit *MEINT_a*, (Z. 087) eingeleitet. Nach kurzem Einatmen und einer Pause von 0.3 Sek. Realisiert A dann die Äußerungseinheit *ja ick KANN au wohl Anders sprEchen wA*, (Z. 088), die verschiedene sprachliche Merkmale, wie Verschluss- anstelle von Reibelaut in *ick* oder die Verwendung der *tag question wA*, aufweist, die regelhaft zur Stilisierung des Berlinischen eingesetzt werden (Lanwer 2011). Flankiert werden entsprechende Gestaltungsmerkmale durch eine Modulation der Sprechstimme. Die Varietätenstilisierung stellt im Verlauf des erzählerischen Spannungsbogens den dramaturgischen Höhepunkt dar. Durch die performative Zuschreibung einer arealsprachlichen Varietätenkompetenz spitzt A die Komplikation noch einmal zu. Seine Reaktion auf die Performance des Mitarbeiters in der erzählten Welt inszeniert A mit *un ICH_sO, (0.5) °h WIE*; (Z. 089–091) als erstaunt, bevor er mit *dann hAt der bis er zwAnzig war in berLIN gewOhnt-* (Z. 096) die Erklärung für die demonstrierte Fähigkeit des Berlinisch-Sprechens liefert und damit die Komplikation auflöst. Die IP *ich bin in be in berLIN AUfgewachsen und <<lachend>geBORN>*; (Z. 093) liefert letztlich die Pointe, was von A durch anschließendes Lachen angezeigt und von B mit Lachen quittiert wird (Z. 094–095). Mit dem Lachen von B überlappend liefert A noch einmal eine Paraphrase der Pointe. Im Weiteren wird diese dann ausgebaut, indem A biographische Stationen der deutschsprachigen Sprachsozialisation des Imbissmitarbeiters chronologisch aufzählt (Z. 096–098).

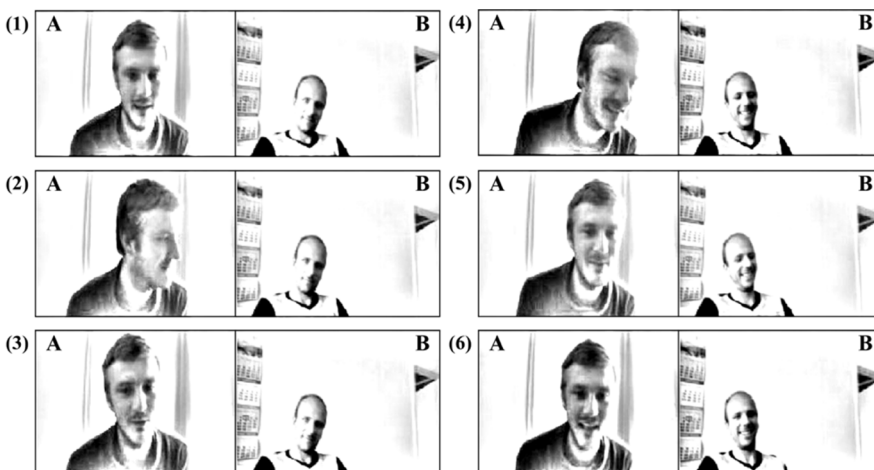


Abbildung 8: Bildstrecke Z. 088–097

In Bezug auf das Blickverhalten lässt sich zunächst festhalten, dass Akteur B nahezu über die gesamte Zeitstrecke den Bildschirm fokussiert, während A auch im Rahmen der Einleitung und Darbietung der Pointe das bereits beschriebene Muster der mit der Inszenierung verschiedener Figuren koordinierten Blickausrichtung fortsetzt (Abbildung 8.1 bis 8.2). Dies gilt – wie auch schon in der Teilsequenz zuvor – sowohl für direkte als auch für indirekte Rede. Auffällig ist hier, dass beide Akteure, während die auf die Pointe folgenden die Lachsequenz noch in ein Blickfenster fällt, mehr oder weniger zeitgleich mit der Realisierung der anschließenden, die Komplikation auflösenden IP in Zeile 096 den Blick vom jeweiligen Bildschirm abwenden (Abbildung 8.3 bis 8.4). Akteur A kehrt jedoch noch in der laufenden IP kurzzeitig mit seinem Blick zurück zum Computerbildschirm. B schaut in diesem Moment grinsend nach rechts unten (Abbildung 8.5), sodass das Grinsen als affiliierendes *stance display* mit Blick auf eine para- sowie nonverbale Erstevaluation des Erzählgeschehens fungieren kann, dass das bereits im gemeinschaftlichen Lachen realisierte *stance alignment* noch einmal bekräftigt. Mit Beginn der biographischen Ausstaffierung der Pointe kehren dann beide Akteure mit ihrem Blick zum Bildschirm zurück (Abbildung 8.6), sodass letztlich ein Großteil der Pointe und deren Wiederholung in ein Blickfenster fällt.

4.2.7. Evaluation und Sequenzausstieg

An die Darbietung der Pointe und die durch Lachen und Grinsen realisierte para- bzw. nonverbale Erstevaluation des Erzählgeschehens schließt sich eine *response sequence* an, in der die narrative Rekonstruktion auch noch mal verbal evaluiert wird:

```

100      (0.1) #
        A: <B---->
        B: <B---->
        Abb      #9-1
→ 101 B: ja hAt_a dich verA#R#SCHT;#
        ----->.....>
        A: ----->.>.....>
102 A: hA#t_a uns schön verARSCHT,#
        ..>R----->
        B: .....>
        Abb      #9-2
103 A: sO_n TYpischer #berLIIn#er #PENn#e#r.#
        ----->.....>B-->.....>L>...>
        B: B----->
104      (0.1) # (0.2) #
        A: B----->
        B: ----->.....>
105 B: hh° #h° #h° h° h° h° #
        ....>B-->.....>
        A: ----->
        Abb      #9-3

```

```

106      (0.1)#
      A: ----->
      B: U----->
      Abb      #9-4
→ 107 B: GEIL;#
      ----->
      A: ----->
108 B: °hh jA SCHÖ#N,
      ----->
      A: ----->.>
109      (0.2)#
      A: .....>
      B: ----->
110 A: JA.#
      ...>
      B: ...>
      Abb      #9-5
111      #(0.6)#(0.5)#
      A: R----->
      B: .....>B----->
112 A: dAt wollt_[ich nur] SAgen;#
      -----[-----]----->
113 B: [JA;; ]
      -----[-----]----->
114      (1.0)#
      A: ----->
      B: ----->
115 A: [ja_ÄH\ ]
      [-----]----->
116 B: [ja WEnigst]ens\#
      [-----]----->
117      (0.1)#
      A: .....>
      B: ----->
→ 118 B: n ACH:-#
      ----->
      A: .....>
      Abb      #9-6
119 B: Ä:H-#
      ----->
      A: B----->

```

120 B: n_ge#BURTstagsge#schEnk für A#nika muss ich jetzt
 kAUfen;#
>R----->.....>B-----
 ----->
 A: -----
 ----->

121 (0.3) #
 A:>
 B: ----->

122 A: JA, #
 ----->
 B: ----->

123 A: MACH dAt; #
 ----->
 B: ----->

Die Evaluationssequenz wird ‚aktivitätskonform‘ von Akteur B, d.h. vom Rezipienten der Narration eingeleitet, der mit *ja hAt_a dich verARSCHT*; (Z. 101) sein Verständnis der Erzählung anzeigt, welches A in der IP *hAt_a uns schÖn verARSCHT*, (Z. 102) durch strukturelle Wiederaufnahme zahlreicher Elemente der initialen evaluativen Äußerung von B in Form eines *confirmatory repeat* bestätigt. Im Weiteren liefert A mit *sO_n TYpischer berLIner PENner*. (Z. 103) eine affektive Positionierung in Bezug auf das Verhalten des Imbissmitarbeiters, worauf B zunächst mit stimmlosem Lachen (Z. 105) goutierend reagiert, bevor er die Erzählung durch mit fallender Intonation realisiertes *GEIL*; (Z. 107) bzw. *°hh JA SCHÖN*, (Z. 108) noch einmal final evaluiert. Inwiefern die Evaluation als *GEIL* dabei auf die im *preface* vorgenommene Charakterisierung lexikalisch zurückgreift, bleibt unklar, da eine evaluative Verwendung von *geil* bspw. auch im Sequenzvorlauf beobachtet werden konnte. Mit der Realisierung von *JA*. (Z. 110) bringt A die *response sequence* schließlich zum Abschluss, was sich zum einen in der auf der Partikel realisierten tieffallenden Intonationskontur, aber zum anderen vor allem in der im Folgenden entstehenden Gesprächspause von gut 11.0 Sek. manifestiert, was A dazu veranlasst den anvisierten Sequenzabschluss mit *dAt wOllt_ich nur SAgen*; (Z. 112) metapragmatisch zu kommentieren. Es folgen teilweise überlappende sowie abgebrochene Äußerungsfragmente beider Interaktionsteilnehmer, bevor B mit dem *ACH*: in Zeile 118 ein neues Gesprächsthema initiiert.



Abbildung 9: Bildstrecke Z. 101–119

Was das Blickverhalten im Rahmen der *response sequence* betrifft, fällt zuallererst auf, dass B mit der Übernahme des Rederechts in Zeile 101 seine Körperpositur verändert, indem er sich deutlich nach vorne lehnt, sodass die Blickzuwendung in gewisser Weise kurzzeitig intensiviert wird (Abbildung 9.1). Das bestehende Blickfenster wird jedoch unmittelbar zu Beginn der verbalen Evaluation aufgelöst, da beide Interaktionsteilnehmer parallel mit der Realisierung des infiniten Teils der Verbklammer der evaluierenden Äußerung *ja hAt a dich verARSCHT*; (Z. 101) ihren Blick jeweils vom Computerbildschirm abwenden. Auffällig ist dabei, dass B im Grunde nicht den Blick vom Bildschirm abwendet, sondern auf das Monitoringfenster fokussiert, in dem er sich selbst sehen kann.⁶ Die anschließenden verbalen Evaluationen durch A erfolgen somit in einer Phase der Blickabwendung beider Akteure. A wendet sich erst im Anschluss an die zweite evaluierende Äußerung *sO_n TYPischer berLIner PENner*. (Z. 103) dem Bildschirm wieder zu, sodass das goutierende Lachen von B nicht nur auditiv, sondern auch visuell von A erfasst werden kann. Im Verlauf der Lachsalve wendet auch B seinen Blick kurz zum Bildschirm, sodass es zu einem wenn auch kurzen Moment der Etablierung eines Blickfensters kommt (Abbildung 9.3). Noch während des Lachens wendet B seinen Blick jedoch wieder vom Bildschirm ab, wohingegen A dem Computerbildschirm bis Ende der zweiten von B realisierten IP, *°hh JA SCHÖN*, (Z. 108), zugewandt bleibt (Abbildung 9.4). Erst mit der folgenden Realisierung des terminalen *JA* wendet A seinen Blick deutlich nach rechts ab. Den metapragmatischen Hinweis auf den Sequenzabschluss in Zeile 112 gibt A in eben dieser abgewandten Position (Abbildung 9.5). B orientiert sich hingegen bereits in der Gesprächspause in Zeile 111 wieder zurück auf seinen Bildschirm und behält diese Blickrichtung auch in der drauffolgenden, durch Abbrüche und Überlappungen gekennzeichneten Umbruchphase bei. A erwidert die Blickzuwendung erst nach Realisierung des tininitiierten *ACH*: durch B, sodass

⁶ Zuvor war das sichtbar lichte Haar von B, was hier vermutlich die Aufmerksamkeit auf sich zieht, bereits Thema des Gesprächs.

die eine Absicht zur Turnübernahme seitens B indizierende Häsitiation $\dot{A}::H-$ in Zeile 119 in ein Blickfenster fällt (Abbildung 9.6). Mit Übernahme des Rederechts wendet B seinen Blick noch einmal kurz vom Bildschirm ab, bevor er diesen in *Mid-Turn*-Position refokussiert, mit dem Effekt, dass auch die durch A mit *JA*, (Z. 112) vollzogene Affirmation des Themenwechsels, die den Ausstieg aus der Erzählsequenz endgültig besiegelt, in einer Phase beidseitiger Bildschirmfokussierung realisiert wird.

5. Systematisierung der Beobachtungen und Ausblick

Auch wenn – oder gerade weil – die durchgeführte Untersuchung eher explorativen Charakter hat, scheint es sinnvoll den Versuch zu unternehmen, die Vielzahl der Beobachtungen zu systematisieren. Ausgegangen wird dabei im Folgenden von der eingangs formulierten Frage, inwiefern sich Mechanismen der multimodalen Herstellung einer konversationellen Erzählung als *embodied joint project* aufdecken lassen, die Evidenzen für das Herstellen eines virtuellen Interaktionsraumes unter spezifischen Bedingungen der Kopräsenz liefern. Insgesamt deutet das Blickverhalten der Akteure zunächst darauf hin, dass die Akteure in einem virtuellen Interaktionsraum agieren, indem die Blickorganisation als Koordinationsressource kommunikativ eingesetzt wird. Die Blickausrichtung der Interagierenden scheint hochgradig aufeinander abgestimmt, was vermuten lässt, dass diese nicht ad hoc, sondern unter Rückbezug auf bestimmte Erwartungserwartungen bezüglich eines ggf. auch erzählspezifischen Blickverhaltens erfolgt, durch dessen Einlösen die Erzählung wiederum als geordnete Handlungssequenz hervorgebracht werden kann. Es stellt sich vor dem Hintergrund der skizzierten Befunde allerdings die Frage, inwiefern es Sinn macht, davon auszugehen, dass die Aktivität des Erzählens als Ganzes durch ein spezifisches Blickverhalten der involvierten Akteure charakterisiert ist. In den untersuchten Daten ist durchaus beobachtbar, dass – wie es auch Rossano beschreibt – der Einstieg in das *telling* sowie dessen Ausgestaltung durch eine ausdauernde Blickzuwendung seitens des Rezipienten quittiert wird. Es lässt sich zudem beobachten, dass beide Akteure im Verlauf der narrativen Einheit in ihrem Blickverhalten hochgradig interdependent sind und dass sie dieses in der wechselseitigen Ausrichtung vor allem an sequenzstrukturellen Schlüsselstellen dazu nutzen, um sich mit Blick auf den kooperativen Handlungsvollzug zu koordinieren, was letztlich auch das Anzeigen bestimmter Beteiligungsrollen betreffen mag. Die Etablierung von Blickfenstern dient dabei nur in einzelnen Fällen zur Regulierung des verbalen Rückmeldeverhaltens, das aber ohnehin eher reduziert ausfällt. Dass das Blickverhalten am Ende also tatsächlich konstitutiv ist für einen erzählspezifischen Beteiligungsrahmen, lässt sich vor dem Hintergrund der durchgeführten Einzelanalyse schwer sagen.

Die Beobachtungen deuten aber zumindest darauf hin, dass das Blickverhalten hier eher an spezifische (Teil-)Handlungen, wie die Inszenierung fremder und eigener Rede oder der Realisierung verbaler, para- und nonverbaler Evaluationen, gebunden ist als an ein übergeordnetes Großprojekt „Erzählen“. Mit Blick auf Formen der Redewieder-

gabe lassen sich Muster beobachten, die in zumindest ähnlicher Form in der Forschungsliteratur bereits beschrieben sind: Sidnell (2006:392–393) kann aufzeigen, dass die Ausrichtung des Sprecherblicks systematisch mit Wechseln zwischen Redeinszenierung und szenischen Beschreibungen alterniert. Während szenische Beschreibungen bei Fokussierung auf den bzw. einen der Rezipienten erfolgen, realisiert der Erzähler Redeinszenierungen mit abgewandtem Bick. In der hier untersuchten Narration alterniert der Sprecherblick hingegen mit der Wiedergabe der Rede verschiedener Figuren. In Bezug auf das Blickverhalten im Rahmen von Evaluationen lässt sich festhalten, dass diese außerhalb der narrativen Sequenz das gleiche Muster erkennen lassen, wie es auch im Rahmen der Erzählung selbst beobachtbar ist. Zugleich decken sich die Beobachtungen mit der Analyse Kendons (1967), der für evaluative Einschübe ein Abwenden des Blickes als charakteristisch beschreibt. Es scheint in diesem Zusammenhang allerdings erwähnenswert, dass einerseits das Abwenden des Blickes im Rahmen der Evaluation in der *response sequence* deutlich später erfolgt, als dies im Rahmen der Seitensequenz oder im Zusammenhang des Sequenzvorlaufs der Fall ist. Die verbale Evaluation des Erzählgeschehens durch Akteur B fällt daher zu einem beträchtlichen Teil in ein Blickfenster, was ansonsten nicht der Fall ist. Andererseits ist zu bemerken, dass die Evaluation der ‚Mini-Pointe‘ innerhalb der Seitensequenz lediglich nonverbal durch Grinsen erfolgt, während wir in der *response sequence* eine para- und nonverbale Erstevaluation finden, der eine verbale Zweitevaluation folgt. Zusammengenommen deuten die Beobachtungen darauf hin, dass ggf. in unterschiedlichen Sequenzpositionen im Rahmen von Erzählungen verschiedene Formen der Evaluation erwartbar sind, die zudem mit unterschiedlichen Formen der Partizipation zusammenhängen können, die (auch) durch das Blickverhalten kontextualisiert werden.

Die Analyse hat zudem gezeigt, dass die Rolle des Blickverhaltens hinsichtlich der Konstitution eines *embodied participation frameworks* auch im Zusammenspiel mit anderen körperlichen Ressourcen wie Kopfausrichtung oder Screen-Proxemik zu sehen ist. Aus dem Vergleich mit dem Rezipientenverhalten vor und nach der narrativen Sequenz wird bspw. ersichtlich, dass zwar nicht die zurückgelehnte Körperhaltung zu Beginn der Erzählsequenz als Kontextualisierungshinweis für die Einnahme einer Rezipientenrolle dient, sehr wohl aber das Auflösen der zurückgelehnten Körperhaltung als Anzeige der Aufgabe der Rezipientenrolle interpretiert werden. Hier zeigt sich, dass im Rahmen von FtS die Proxemik im Sinne einer auf die Aufnahmeeinheit bezogenen Distanzregulierung als kommunikative Ressource nutzbar gemacht wird (vgl. Lanwer i.Dr.), was den Analysegegenstand potenziell ausweitet. Zum anderen wird deutlich, dass die Unterscheidung zwischen Sprecher- und Hörerrolle nicht hinreichend ist, um den dynamischen Beteiligungsrahmen einer Erzählsequenz zu erfassen. Für die Evaluation einer Erzählung ist gewissermaßen ein Sprechen als Rezipient konstitutiv, während die frühzeitige Aufgabe des Rezipientenstatus bspw. durch einen abrupten Themenwechsel den sauberen Abschluss der Sequenz gefährdet (Quasthoff 1981:310–311). Der *participation status* eines Akteurs ist also vermutlich vor allem auch dann nicht hinreichend, um das Blickverhalten der Interagierenden zu erklären – wie es Rossano (2012:67) moniert –, wenn dieser auf eine simple Unterscheidung zwischen

Sprecher- und Hörerrolle reduziert wird. Die Differenzierung übergeht die beobachtbare Vielfalt von Formen der Sprecher- und Hörer-Partizipation, wie sie vor allem von Goffman (1979) bzw. C. Goodwin (1984; 1986) beschrieben worden sind (vgl. für einen Überblick auch Levinson 1988). Außerdem lässt die dichotome Kategorisierung Verschiebungen der Beteiligungsrollen innerhalb eines Gesprächs, die im Rahmen konversationeller Erzählungen – wie beschrieben – zwar den Wechsel von Sprecher- und Hörerrolle aber bspw. nicht den von Erzähler- und Rezipientenrolle betreffen können, prinzipiell außer Acht. Noch eklatanter wird diese Problematik mit Blick auf die Analyse von *assisted storytelling* (Lerner 1992:247) bzw. allgemein in Bezug auf kollaboratives Sprechen (Schwitalla 1993).

Für zukünftige Untersuchungen scheint es daher sinnvoll, entsprechende Konstellationen mit in die Analyse einzubeziehen. Es ist zu erwarten, dass sich vor diesem Hintergrund auch der Widerspruch zwischen C. Goodwins und Rossanos Interpretation des Blickverhaltens auflöst: Es ist davon auszugehen, dass mit spezifischen Handlungssequenzen bzw. Handlungen in der Regel bestimmte Beteiligungsrollen verbunden sind. Die Tätigkeiten des Sprechens und Zuhörens erfolgen in unterschiedlichen Handlungszusammenhängen und Interaktionsökologien unter verschiedenen Voraussetzungen der Partizipation. Der sich hier anschließenden Frage, ob diese Voraussetzungen in der FtS unter Umständen andere sind als in der FtF, könnte in weiterführenden Untersuchungen nachgegangen werden. Die Analysen des Blickverhaltens deuten zwar zunächst darauf hin, dass in der FtS auf ähnliche Verhaltensrepertoires im Rahmen der kooperativen Herstellung einer konversationellen Erzählung als *embodied joint project* zurückgegriffen wird, wie sie in der Literatur für die FtF beschrieben werden. Es ist jedoch sicher nicht unproblematisch, dass es in der FtS nicht zur Etablierung „echten“ Blickkontakts kommen kann (umso bemerkenswerter scheint es daher, dass die Analyse zeigt, dass das Blickverhalten dennoch als Ressource der Handlungskoordination genutzt wird). Aufgrund der evidenten Unterschiede hinsichtlich der „ökologischen“ Bedingungen der Interaktionskonstitution ist in Zweifel zu stellen, dass sich die Mechanismen der interaktiven Herstellung von Partizipation bzw. Ko-Partizipation in der FtS nicht in irgendeiner Form von denen der in der FtF unterscheiden. Es scheint vielmehr wahrscheinlich, dass „technology that links one setting to another expand[s] our notion of human participation in a historically built social and material world“ (C. Goodwin/M.H. Goodwin 2004:239). Die Frage, inwiefern gerade der Umstand, dass ‚echter‘ Blickkontakt in der FtS nicht möglich ist, sich auf die Art der interaktiven Herstellung des Partizipationsrahmens auswirkt, kann letztlich nur in einer kontrastiven Studie nachgegangen werden, die bspw. narrative Sequenzen in der FtF mit narrativen Sequenzen in der FtS unter eben diesem Gesichtspunkt vergleicht. Die vorliegende Studie lässt bereits erkennen, dass der anvisierte sequenzielle Kontext einen fruchtbaren Boden für einen entsprechenden Vergleich bereitstellen könnte.

6. Literatur

- Arminen, I., Licoppe, C., & Spagnolli, A. (2016). Respecifying Mediated Interaction. *Research on Language and Social Interaction*, 49(4), 290–309.
- Auer, P. (2017). Anfang und Ende fokussierter Interaktion: Eine Einführung.
- Bavelas, J. B., Coates, L., & Johnson, T. (2002). Listener Responses as a Collaborative Process: the Role of eye Gaze. *Journal of Communication*, 52, 566–580.
- Betz, E. (2008). *Grammar and Interaction. Pivots in German Conversation*. Amsterdam; Philadelphia: Benjamins.
- Chafe, W. (1994). *Discourse, consciousness, and time. The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago; London: The University of Chicago Press.
- Clark, H. H. (1996). *Using language*. Cambridge; New York: Cambridge University Press.
- De Fornel, M. (1994). Le cadre interactionnel de l'échange visiophonique. *Réseaux*, 12(64), 107–132.
- Deppermann, A., & Schmitt, R. (2007). Koordination. Zur Begründung eines neuen Forschungsgegenstandes. In R. Schmitt (Ed.), *Koordination. Analysen zur multimodalen Interaktion*. (pp. 15–54). Tübingen: Narr.
- Deppermann, A., & Helmer, H. (2013). Zur Grammatik des Verstehens im Gespräch: Inferenzen anzeigen und Handlungskonsequenzen ziehen mit also und dann. *Zeitschrift für Sprachwissenschaft*, 32(1), 1–39.
- Du Bois, J. W. (2007). The stance triangle. In R. Englebreston (Ed.), *Stancetaking in Discourse. Subjectivity, evaluation, interaction* (pp. 139–182). Amsterdam/Philadelphia: Benjamins.
- Goffman, E. (1963). *Behavior in Public Places: Notes on the Social Organization of Gatherings*. New York: Free Press.
- Golato, A. (2000). An innovative German quotative for reporting on embodied actions: Und ich so/und er so 'and i'm like/and he's like.' *Journal of Pragmatics*, 32(1), 29–54.
- Goodwin, C. (1981). *Conversational Organization: Interaction Between Speakers and Hearers*. New York: Academic Press.
- Goodwin, C. (1984). Notes on story structure and the organization of participation. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action. Studies in Conversational Analysis* (pp. 225–246). New York u.a.: Cambridge University Press.
- Goodwin, C. (1986). Audience diversity, participation and interpretation. *Text*, 6(3), 283–316.
- Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics*, 32, 1489–1522.
- Goodwin, C. (2007). Participation, stance and affect in the organization of activities. *Discourse & Society*, 18(1), 53–73.
- Goodwin, C., & Goodwin, M. H. (2004). Participation. In *A Companion to Linguistic Anthropology* (pp. 222–244). Oxford: Blackwell.
- Goodwin, M. H. (1997). Byplay: Negotiating Evaluation in Storytelling. In G. R. Guy, J. Baugh, D. Schiffrin, & C. Feagin (Eds.), *Towards a Social Science of Language: Papers in Honour of William Labov* (Vol. 2, pp. 77–102). Philadelphia: Benjamins.
- Goodwin, M. H. (2017). Haptic Sociality. The Embodied Interactive Constitution of Intimacy through Touch. In C. Meyer, J. Streeck, & J. J. Scott (Eds.), *Intercorporeality: Emerging Socialities in Interaction* (pp. 73–102). Oxford: Oxford University Press.
- Harper, R., Watson, R., & Licoppe, C. (2017). Interpersonal video communication as a site of human sociality. *Pragmatics*, 27(3), 301–318.

- Hausendorf, H. (2010). Interaktion im Raum. Interaktionstheoretische Bemerkungen zu einem vernachlässigten Aspekt von Anwesenheit. In A. Deppermann & A. Linke (Eds.), *Sprache intermedial: Stimme und Schrift, Bild und Ton* (pp. 163–197). Tübingen: Narr.
- Hausendorf, H. (2013). On the interactive achievement of space – and its possible meanings. In P. Auer, M. Hilpert, A. Stukenbrock, & B. Szmrecsanyi (Eds.), *Space in Language and Linguistics: Geographical, Interactional and Cognitive Perspectives* (pp. 276–303). Berlin; Boston: De Gruyter.
- Hausendorf, H. (2015). Interaktionslinguistik. In L. Eichinger (Ed.), *Sprachwissenschaft im Fokus. Positionsbestimmungen und Perspektiven* (pp. 43–69). Berlin; München; Boston: De Gruyter.
- Heath, C. C. (1982). The display of reciprocity: An instance of a sequential relationship in speech and body movement. *Semiotica*, 42(2–4), 147–168.
- Heath, C. C., & Luff, P. (1993). Disembodied Conduct: Interactional Asymmetries in Video Mediated Communication. In Technology in Working Order. In G. Button (Ed.), *Technology in Working Order: Studies of Work, Interaction, and Technology* (pp. 35–54). New York: Routledge.
- Jefferson, G. (1972). Side Sequences. In D. N. Sudnow (Ed.), *Studies in social interaction* (pp. 294–333). New York: Free Press.
- Jefferson, G. (1978). Sequential Aspects of Storytelling in Conversation. *Studies in the Organization of Conversational Interaction*, 219–248.
- Jefferson, G. (1978). Sequential Aspects of Storytelling in Conversation. *Studies in the Organization of Conversational Interaction*.
- Keating, E., & Mirus, G. (2004). American Sign Language in virtual space: Interactions between deaf users of computer-mediated video communication and the impact of technology on language practices. *Language in Society*, 32(5), 693–714.
- Kendon, A. (1967). Some Functions of Gaze-Direction in Social Interaction. *Acta Psychologica*, 26, 22–63.
- König, K. (2017). Question tags als Diskursmarker? Ansätze zu einer systematischen Beschreibung von ne im gesprochenen Deutsch. In B. Hardarik, D. Arnulf, H. Henrike, & T. Spranz-Fogasy (Eds.), *Diskursmarker im Deutschen. Reflexionen und Analysen* (pp. 233–258). Mannheim: Verlag für Gesprächsforschung.
- Langacker, R. W. (2001). Discourse in Cognitive Grammar. *Cognitive Linguistics*, 12(2), 143–188.
- Lanwer, J. P. (2011). “Ick lieb dir wohl!” Dialektologische Untersuchungen zur Stilisierung regionaler Substandards in der Face-to-face-Interaktion. *Niederdeutsches Wort*, 51, 107–131.
- Lanwer, J. P. (i.Dr.). Blended joint attention in medial-vermittelter Interaktion. Beobachtungen zur Konstitution eines virtuellen Interaktionsraumes in Videotelefonaten mit Skype. In K. Marx & A. Schmidt (Eds.), *Interaktion und Medien*. Heidelberg: Winter.
- Lerner, G. H. (1992). Lerner - Assisted Storytelling. Deploying shared knowledge as a practical matter.pdf. *Qualitative Sociology*, 15(3), 247–271.
- Levinson, S. C. (1988). Putting linguistics on a proper footing: Explorations in Goffman’s concepts of participation. In P. Drew & A. Wootton (Eds.), *Erving Goffman: Exploring the interaction order* (pp. 161–227). Oxford: Polity Press.
- Licoppe, C., & Morel, J. (2012). Video-in-Interaction: “Talking Heads” and the Multimodal Organization of Mobile and Skype Video Calls. *Research on Language and Social Interaction*, 45(4), 399–429.

- Mondada, L. (2007). Interaktionsraum und Koordinierung. In R. Schmitt (Ed.), *Koordination. Analysen zur multimodalen Interaktion* (pp. 55–93). Tübingen: Narr.
- Mondada, L. (2010). Eröffnungen und Präeröffnungen in medienvermittelter Interaktion: Das Beispiel Videokonferenz. In L. Mondada & R. Schmitt (Eds.), *Situationseröffnungen. Zur multimodalen Herstellung fokussierter Interaktion* (pp. 277–334). Tübingen: Narr.
- Mondada, L. (2013). Interactional space and the study of embodied talk-in-interaction. In P. Auer, M. Hilpert, A. Stukenbrock, & B. Szmrecsanyi (Eds.), *Space in Language and Linguistics: Geographical, Interactional and Cognitive Perspectives* (pp. 247–275). Berlin; Boston: De Gruyter.
- Mondada, L. (2014). Conventions for multimodal transcription.
https://mainly.sciencesconf.org/conference/mainly/pages/Mondada2013_conv_multimodality_copie.pdf
- Quasthoff, U. (1981). Zuhöreraktivitäten beim konversationellen Erzählen. In *Dialogforschung* (pp. 287–313).
- Rossano, F. (2012). Gaze in Conversation. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (First Edit., pp. 308–329). Chichester: Blackwell.
- Rossano, F. (2013). *Gaze behavior in face-to-face interaction*. Nijmegen: Radboud University.
- Sacks, H. (1971). Das Erzählen von Geschichten innerhalb von Unterhaltungen. In R. Kjolseth & F. Sack (Eds.), *Zur Soziologie der Sprache. Ausgewählte Beiträge vom 7. Weltkongreß der Soziologie* (pp. 307–314). Opladen: Westdeutscher Verlag.
- Schegloff, E. A. (1982). Discourse as an interactional achievement: Some uses of “uh-huh” and other things that come between sentences. In D. Tannen (Ed.), *Analyzing discourse: Text and talk* (pp. 71–93). Washington D.C.: Georgetown University Press.
- Schegloff, E. A. (2007). *Sequence Organization in Interaction. A Primer in Conversation Analysis*. New York: Cambridge University Press.
- Schmitt, R. (2005). Zur multimodalen Struktur von turn-taking. *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* *Online-Zeitschrift zur verbalen Interaktion*, 6, 17–61.
- Schütz, A., & Luckmann, T. (1979). *Strukturen der Lebenswelt*. (Vol. 1). Frankfurt am Main: Suhrkamp.
- Schwitalla, J. (1993). Über einige Weisen des gemeinsamen Sprechens. Ein Beitrag zur Theorie der Beteiligungsrollen im Gespräch. *Zeitschrift für Sprachwissenschaft*, 11(1), 68–98.
- Selting, M., Auer, P., Barth-Weingarten, D., Bergmann, J., Bergmann, P., Birkner, K., et al. (2009). Gesprächsanalytisches Transkriptionssystem 2 (GAT 2) Margret. *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion*, 10, 353–402.
- Sidnell, J. (2006). Coordinating Gesture, Talk, and Gaze in Reenactments. *Research on Language & Social Interaction*, 39(4), 377–409.
- Stivers, T. (2008). Stance, alignment, and affiliation during storytelling: When nodding is a token of affiliation. *Research on Language & Social Interaction*, 41(1), 31–57.
- Stivers, T. (2013). Sequence Organization. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (First Edit., pp. 191–209). Chichester: Blackwell.
- Streeck, J. (2014). Mutual gaze and recognition. Revisiting Kendon’s “Gaze direction in two-person conversation.” In M. Seyfeddinipur & M. Gullberg (Eds.), *From Gesture in Conversation to Visible Action as Utterance: Essays in honor of Adam Kendon* (pp. 35–55). Amsterdam; Philadelphia: Benjamins.

- Stukenbrock, A. (2013). Commentary: What Difference Does Space Make for Interaction and Interaction for Space? In *Space in Language and Linguistics. Geographical, Interactional, and Cognitive Perspectives*. (pp. 401–408).
- Weiss, C., & Auer, P. (2016). Das Blickverhalten des Rezipienten bei Sprecherhäsitationen: eine explorative Studie. *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion*, 17(17), 132–167.
- Zima, E. (2017). Multimodale Mittel der Rederechtsaushandlung im gemeinsamen Erzählen in der Face-to-Face Interaktion Elisabeth Zima, 18(18), 241–273.