

MULTIMODALE KONVERSATIONSANALYSE

**Ko-Orientierung und Ko-Referenzierung
in Augmented Reality basierter Kooperation**

Dissertation

Zur Erlangung des akademischen Grades Dr. phil.

im Fachbereich Linguistik

aus dem Sonderforschungsbereich 673 “Alignment in Communication”

der Universität Bielefeld

vorgelegt von

Christian Schnier

aus Herford

- 2015 -

Gedruckt auf alterungsbeständigem Papier ISO 9706

Erste Gutachterin: Prof. Dr. Karola Pitsch

Zweiter Gutachter: Prof. Dr. Ulrich Dausendschön-Gay

Ko-Orientierung und Ko-Referenzierung in Augmented Reality basierter Kooperation

Christian Schnier

– 2015 –

Danksagung

Die vorliegende Arbeit ist im Rahmen meiner Tätigkeit als Wissenschaftlicher Mitarbeiter im C5-Projekt “Alignment in AR-based cooperation” des Sonderforschungsbereiches “Alignment in Communication” (SFB-673/DFG) der Universität Bielefeld entstanden. Bei der Umsetzung und Fertigstellung unterstützten mich viele Freunde, Kollegen und Wegbegleiter, denen ich an dieser Stelle herzlich danken möchte!

Prof. Dr. Karola Pitsch und Prof. Dr. Ulrich Dausendschön-Gay haben diese Arbeit angeregt und begleitet. In Theorie-, Methoden- und Analysefragen haben sie mir mit manchem Rat zur Seite gestanden und durch ihr motivierendes Interesse sehr zum Fortschreiten der Arbeit beigetragen. Ihnen gilt mein besonderer Dank.

Das C5-Projekt wurde von Prof. Dr. Karola Pitsch und Dr. Thomas Hermann geleitet. Ihnen ist es gelungen, eine Arbeitsatmosphäre zu schaffen, in der sich die Linguistik und die Informatik auf Augenhöhe begegnen, um ihr jeweiliges Potential synergetisch einzubringen. Ebenfalls bedanken möchte ich mich in diesem Zusammenhang bei meinem Projektkollegen Alexander Neumann, der zu dieser Arbeitsatmosphäre ganz wesentlich beigetragen hat. Ohne seine kompetente, kollegiale und freundschaftliche Unterstützung hätte diese Arbeit so nicht entstehen können. Für ihr unermüdliches Engagement im Videoschnitt und der Annotation des C5-Korpus danke ich Katharina Geretzky und Paul Diestelhorst.

Die AG „Angewandte Informatik“ der Universität Bielefeld hat mich finanziell unterstützt, als fachfremdes Mitglied in ihren Reihen mit viel Wohlwollen aufgenommen und mir ein Arbeitsumfeld geboten, in dem ich meine Ideen und Analyseergebnisse im interdisziplinären Rahmen diskutieren konnte. Dafür bedanke ich mich herzlich bei Prof. Dr. Britta Wrede und Prof. Dr. Franz Kummert. In den Dank eingeschlossen seien alle Mitglieder und Arbeitskollegen der AG, die mir über die Jahre ein gleichermaßen inspirierendes wie freundschaftlich und motivational unterstützendes Umfeld geschenkt haben.

Der DAAD hat meine Arbeit mit einem Forschungsstipendium an der University of California at Los Angeles finanziell unterstützt und mir dadurch die Möglichkeit eröffnet, mich in der ‚Keimzelle‘ der Konversationsanalyse bei Prof. Dr. Charles Goodwin fachlich weiterzubilden und einzelne Analyseergebnisse mit ihm zu diskutieren.

Für die finanzielle Förderung durch ein Promotionsabschlusstipendium bedanke ich mich ebenfalls beim Bielefelder Nachwuchsfonds.

Für zahlreiche Datensitzungen, konstruktive Diskussionen, wertvolle Literaturangaben und ermutigende Worte, wenn es einmal nicht so rund lief, danke ich meinen Arbeitskollegen Luise, Raphaela, Marcel und Malte aus Karolas Forschungsgruppe “Interactional Linguistics & Human-Robot-Interaction” und den Teilnehmern des Forschungskolloquiums „Sprache und Interaktion“.

Meinen Eltern danke ich nicht nur für ihre tatkräftige Unterstützung im Lektorat der Arbeit, sondern auch für den nötigen Rückhalt, ihr Interesse und unerschütterliches Vertrauen, mit denen sie meinen Weg begleitet haben. Für die professionelle Unterstützung in Statistikfragen, fruchtbare Diskussionen über ‚das große Ganze‘ und die Durchhaltungsmotivation danke ich meinem Bruder Fabian. Mein umfassendster Dank richtet sich an meine Freundin Sara, die auch die wohl unvermeidlichen Frustrationen im Verlauf einer solchen Arbeit verständnisvoll aufgefangen hat. Mit ihrer unermüdlichen Geduld und menschlichen Unterstützung hat sie ganz wesentlich zur Fertigstellung dieser Arbeit beigetragen.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Gegenstandskonstitution	4
1.2	Konzeptueller und methodischer Rahmen	7
1.2.1	Joint Attention und Ko-Orientierung: Unterschiedliche Paradigmen eines Phänomens....	7
1.2.2	Die ethnomethodologische Konversationsanalyse multimodaler Prägung	9
1.2.3	Augmented Reality als empirisches Werkzeug der Interaktionsforschung.....	13
1.2.4	Studiendesign.....	14
1.2.5	Korpus.....	15
2	Einfluss des ‘Mutual Monitoring’ auf die Verfahren zur Herstellung von Ko-Orientierung	19
2.1	Mutual Monitoring als Gegenstand der Interaktionsforschung	21
2.2	Mutual Monitoring als Grundbedingung der Face-to-Face Interaktion	23
2.2.1	Phänomenologische Differenzierung: Attrahierende und mediatisierende Aufmerksamkeitsbindung.....	23
2.2.2	Phänomenologische Ursachen: Multimodale Ablauformate der Präkontexte.....	25
2.2.3	Mediatorfunktion multimodaler Äußerungsgestalten auf spezifische Referenten.....	28
2.2.4	Zusammenfassung	33
2.3	Mutual Monitoring-basierte Verfahren zur situativen Desambiguierung	35
2.3.1	Definition von Simultaneität.....	35
2.3.2	Multimodale Verdichtungsräume	37
2.3.3	„Koordination“ und die Beziehungskonstellation von Sequentialität und Simultaneität....	38
2.3.3.1	„Koordination“ als Forschungsgegenstand	38
2.3.3.2	Beziehungskonstellation von Sequentialität und Simultaneität.....	41
2.3.3.3	Zentrale Ergebnisse zur organisationalen Relation der Strukturbildungsmechanismen.....	42
2.3.4	Verhältnis der overlap-Konzeption zur kinesischen Simultaneität.....	44
2.3.5	Ableitung des Forschungsdesiderats.....	49
2.3.6	Phänomenologische Differenzierung: Konfligierende & funktionale Simultaneität.....	50
2.3.7	Bearbeitung konfligierender Simultanaktivitäten	54
2.3.7.1	Formen gestischer Sequentialisierung	54
2.3.7.2	Sequentialisierung durch Direktiva.....	61
2.3.7.3	Sequentialisierung durch Blickkoordination	67
2.3.8	Funktionale Simultaneität.....	69
2.3.8.1	Formen situativer Desambiguierung.....	70
2.3.8.2	Simultaneität in Überprüfungs-, Planungs- und Kooperationsphasen.....	72
2.3.9	Zusammenfassung	77
2.4	Mutual Monitoring in Augmented Reality-basierter Interaktion	86
2.4.1	Emergierende Handlungsparallelität in der AR-Bedingung.....	86
2.4.2	Auflösung konkurrierender Aktivitäten	89
2.4.2.1	Die Unterstellung von Ko-Orientierung.....	89
2.4.2.2	Auflösung konkurrierender Aktivitäten unter der Anwesenheitsbedingung von Ko-Orientierung.....	93
2.4.3	Zusammenfassung	97

3	Robustheit interaktiver Koordinierungsressourcen.....	101
3.1	Deixis im Verweisraum – Räumliche Orientierung.....	103
3.2	Die Relevanz lokaldeiktischer Ausdrücke („hier“) und Rekonstruktion deiktischer Gesten bei der Herstellung von Ko-Orientierung.....	107
3.2.1	Erfolgreiche Ko-Orientierung durch verbalsprachliche Lokaldeixis.....	110
3.2.2	Reparatur von Ko-Orientierung: Wenn lokaldeiktische Ausdrücke insuffizient sind	111
3.2.3	Deiktische Gesten im analytischen Vergleich der Bedingungen (AR vs. F2F).....	113
3.2.4	Abweichende Fälle.....	116
3.2.4.1	Desambiguierungsprobleme – Divergenz von Bedeutungskonstitution und -interpretation einer Orientierungsgestalt	116
3.2.4.2	Verbalsprachliche Ausdrücke in multiplen Orientierungsangeboten	120
3.3	Die Herstellung von Normalität: “wait & see” als Handlungsprinzip	122
3.4	Weitere Koordinierungsressourcen zur Herstellung von Ko-Orientierung.....	124
3.4.1	Diskursstrukturierende Partikeln, Aufgabenratifikationen und -eröffnungen	125
3.4.1.1	Prospektive Orientierungsfunktion.....	126
3.4.1.2	Retrospektive Orientierungsfunktion	129
3.4.1.3	Abweichender Fall zur retrospektiven Orientierungsfunktion.....	131
3.4.2	Prestrukturierung der Interaktionsumgebung – Präparation des Referenzobjekts	134
3.4.2.1	Präparation des Referenzobjekts im Vorfeld der Aufgabenankündigung	134
3.4.2.2	Objektpräparation als Hinweis zur Ko-Konstruktion des Referenten.....	136
3.4.3	Steigerung der Objektsalienz.....	138
3.4.3.1	Anheben und Vorzeigen	138
3.4.3.2	Trajektorielle Modifikationen.....	140
3.4.3.3	Objektmanipulative Relevanzmarkierung bei konkurrierenden Orientierungsangeboten.....	142
3.5	Analytische Kontrastierung zum Exhibition-Design-Szenario.....	144
3.5.1	Interaktionsbedingungen im Exhibition-Design-Szenario.....	145
3.5.2	Aushandlung neuer Orientierungsroutinen	146
3.6	Zusammenfassung und Fazit: Robustheit von Interaktionssystemen	152
4	Ko-Orientierung in handlungsschematischen Zusammenhängen	159
4.1	Diskursstrukturelle Einheiten als Forschungsgegenstand	160
4.2	Ableitung des Forschungsdesiderates.....	161
4.3	Sequenzstrukturelle Auffälligkeiten der Verbalebene	163
4.3.1	Methodische Vorbemerkung zur Verbalannotation	163
4.3.2	Sequenzstrukturelle Auffälligkeiten und Hypothesenbildung	164
4.4	Hypothesenverifizierung.....	165
4.4.1	Präzisierung durch Visualisierung der Redebeteiligung im Interaktionsverlauf	165
4.4.1.1	Methodische Vorbemerkung zum statistischen Verfahren des “Binning”	166
4.4.1.2	Beobachtungen zur Redebeitragsverteilung.....	169
4.4.2	Präzisierung durch Visualisierung weiterer Beteiligungsressourcen: Objektmanipulative Handlungen	170
4.4.3	Abweichende Beteiligungsressourcen: Deiktische Gesten.....	172

4.4.4	Aktivitätsbasierte Klassifikation der Beteiligungsressourcen.....	172
4.4.4.1	Makrostruktureller Vergleich der Beteiligungsressourcen.....	173
4.4.4.2	Mesostruktureller Vergleich der einzelnen Handlungskategorien der Modalität „deiktische Gesten“.....	176
4.4.4.3	Mikrostrukturelle Analyse zur Bestimmung der interaktionalen Funktion strukturell divergenter Handlungskategorien.....	178
4.4.5	Zusammenfassung: Kernbeobachtungen und ‚Seiteneffekte‘.....	184
4.5	Bestimmung der kernaktivitätsspezifischen Abläufe des Präkontextes.....	185
4.5.1	Initiale Handlungskonfiguration.....	185
4.5.2	Unmittelbar vorausgehende Handlungskonfiguration.....	188
4.6	Zusammenfassung und Fazit.....	194
5	Schlussbetrachtung und Ausblick.....	199
5.1	Grundlagenerkenntnisse: Orientierungslogiken der F2F-Interaktion.....	199
5.2	Auswirkungen eines eingeschränkten “mutual monitoring”.....	201
5.3	Anschlussoptionen für die Folgeforschung.....	203
	Literaturverzeichnis.....	205
	Abbildungsverzeichnis.....	215
	Anhang.....	217
	Transkriptionskonventionen.....	217
	Eidesstattliche Erklärung.....	219

1 Einleitung

“Awareness reveals a phenomenal domain largely unexplored by the social sciences, a domain which directs our attention to a complex body of socially organised practice and reasoning.”

(Heath et al. 2002: 346)

Eine der zentralen Anforderungen sozialer Interaktion besteht in der Herstellung gemeinsamer Orientierungen. Auf Grundlage gegenseitiger Beobachtung (“mutual monitoring”, Goffman 1963; M. H. Goodwin 1980; Clark & Krych 2004) koordinieren Interaktanten ihre Wahrnehmungen auf materielle Umgebungsressourcen, um sie in den Fokus des kooperativen Handelns zu rücken. In sozialen Interaktionsformen reichen sich Menschen die Butter am Frühstückstisch, betrachten gemeinsam Exponate in Museumsausstellungen und tratschen am Gartenzaun über die zu hohen Obstbäume des Nachbarn. All dies sind Beispiele der menschlichen Sozialität, die ohne die Fähigkeit, die eigene sensorische Wahrnehmung mit der des Gegenüber koordinieren zu können, nicht denkbar wäre.

Im vorherrschenden Paradigma ist dieses Phänomen als “Joint Attention” konzeptualisiert und unter einer kognitivistischen Perspektive entwicklungspsychologischer Ansätze erforscht worden. In der Herstellung gemeinsamer Aufmerksamkeiten beginnen Kinder, dem situativen Handeln ihrer Bezugspersonen intentionale Absichten zuzuschreiben (“mind-reading”, Baron-Cohen 1997). Dabei lernen sie, dass die intersubjektiven Bewusstseinsvorgänge über die im Aufmerksamkeitsfokus liegende Entität für ein gemeinsames Handlungsziel aufeinander abgestimmt werden können (“shared intentionality”, Tomasello 1995). Als zentraler Entwicklungsmechanismus gilt die Fähigkeit des Kindes, den Blicken und Gesten seiner Bezugsperson zu folgen (“gaze-following”, Scaife & Bruner 1975), sowie selbst – ontogenetisch nachfolgend – die Aufmerksamkeitsorientierung des Gegenüber durch Blickadressierungen und Gesten zu manipulieren. Innerhalb des kognitionspsychologischen Paradigmas ist festzustellen, dass die ontogenetische Relevanz des “gaze-following” zu einer Fokussierung auf visuell-perzeptive Teilnehmungsweisen in der frühkindlichen Kognitionsentwicklung geführt hat (Bigelow 2003; Akthar & Gernsbacher 2007; Yu & Smith 2013; Langton et al. 2000). Obgleich in jüngerer Zeit ein wachsendes Interesse an der Relation zwischen “Joint Attention” und dem Signalisierungssystem „Sprache“ bei adoleszenter Ausprägung erkennbar wird (z.B. Tomasello 2000), ist aus dem Primat des Visuellen eine epistemische Lücke entstanden: Das intra- und intersubjektive Zusammenwirken multimodaler Ausdrucksressourcen (Blick, Gestik, Körperdisplay, etc.) ist bei der Herstellung von “Joint Attention” bislang kaum systematisch erforscht worden.

Demgegenüber haben Studien aus dem Bereich der ethnomethodologischen Konversationsanalyse (KA) in ihrer Erweiterung der multimodalen Bedeutungskonstitution (Goodwin 2000a, 2003a, 2013; Dausendschön-Gay & Krafft 2002; Deppermann & Schmitt 2006) die sequentiellen Organisationsprinzipien der gemeinsamen Orientierung auf empirischer Datengrundlage von Videoaufzeichnungen natürlicher Interaktionen untersucht. Im analytischen Fokus auf die alltagspraktischen Herstellungsverfahren der Interaktanten ist das Phänomen als „Ko-Orientierung“ (z.B. Hausendorf 2010) konzeptualisiert worden. Um interaktional relevante Wahrnehmungen im Raum zu enactieren, verwenden Interaktanten unterschiedliche semiotische Körper- und Kontextressourcen (Verbalsprache, Syntax, Prosodie, Blick, Gestik, Körperdisplay, materielle Umgebungsressourcen, etc.) und konfigurieren sie in feingradueller Abstimmung zum situativ emergenten Adressatenverhalten in wahrnehmbare Äußerungs- und Orientierungsgestalten. Die aktive Aufmerksamkeitsmanipulation des Ko-Partizipanten basiert auf der Fähigkeit, die eigene Wahrnehmung durch multimodale Ausdrucksressourcen wahrnehmbar zu „beschreiben“ (“accounting practices”, Garfinkel 1967), um sie der Wahrnehmung des Gegenüber zuzuführen.

Obgleich die multimodal ausgerichtete Konversationsanalyse ab Mitte der 70er Jahre grundlegende Konstitutionsprinzipien der interaktiven Herstellung gemeinsamer Orientierungen sequenzanalytisch erforscht hat, lassen sich – äquivalent zur kognitivistischen Ausrichtung auf das Phänomen – epistemische Lücken erkennen: Ein zentrales Problem liegt in der methodischen Schwierigkeit begründet, die jeweiligen Teilnehmerbeobachtungen der analytischen Rekonstruktion “from within” (Garfinkel 1967) zugänglich zu machen. Während der KA in ihrer anfänglichen Fokussierung auf die sequentiellen Ordnungsstrukturen der Verbalsprache mit der Konzeption der “next turn proof procedure” (Sacks et al. 1974) ein dem Gegenstand angemessenes analytisches Rekonstruktionsverfahren zur Verfügung stand, lassen sich für dieses

Validierungsinstrument unter einer multimodalen Perspektive auf Interaktion methodische Schwächen erkennen. In der Gleichzeitigkeit intra- und interpersoneller Aktivitäten („Simultaneität“) manifestiert sich die Verständnissicherung nicht ausschließlich in den Anschlussreaktionen, sondern erfolgt bereits simultan zur situativ emergenten Äußerungsgestalt. Vor diesem Hintergrund operiert die KA mit zwei Grundannahmen, die per se eigenständige empirische Fragen konstituieren. Sie geht einerseits von einer Persistenzbedingung des “mutual monitoring” in face-to-face Interaktionen aus (“Situations begin when mutual monitoring occurs, and lapse when the second-last person has left.” (Goffman 1963: 18)). Andererseits – und dies kann als handlungstheoretische Konsequenz der ersten Annahme verstanden werden – nimmt sie an, dass multimodale Äußerungsgestalten im Sinne gestaltpsychologischer Prinzipien als ein kommunikatives Ganzes hervorgebracht und rezipiert werden (vgl. “contextual configuration”, Goodwin 2000a, 2003a, 2009; s.a. Dausendschön-Gay & Krafft 2002). Während die multimodale Assemblierungsqualität kommunikativer Äußerungsgestalten ein empirisch beobachtbares Konstitutionsmerkmal bildet, lässt sich die holistische Rezeption in der multimodalen Komplexität sequentiell- und simultan-strukturierter Austauschprozesse von Orientierungshinweisen kaum überprüfen. Dabei stellt sich die Frage, welche interaktionalen Effekte beobachtbar werden, wenn den Interaktanten nur Teilinformationen der holistischen Äußerungsgestalt zur Verfügung stehen.

An dieser Stelle setzt die vorliegende Arbeit an, indem sie die aufgezeigten epistemischen Lücken der kognitivistischen und praxeologischen Perspektive aus einem genuin multimodalen Interaktionsverständnis adressiert. Dabei wird ein methodischer Ansatz verfolgt, der aus heuristischen Gründen mit dem konversationsanalytischen Aufzeichnungspostulat natürlicher Interaktionen bricht. Die empirische Datengrundlage bilden semi-experimentelle Studien aus dem C5-Projekt “Alignment in AR-based cooperation” (SFB-673/DFG), in denen wir Interaktanten paarweise baten, die Neugestaltung des Naherholungsgebietes „Bielefelder Obersee“ aus der Position konträrer Rollen (Investor vs. Naturschützer) auszuhandeln. Neben einer topographischen Karte der Region, hatten sie 18 Objekte zur Verfügung, die potentielle Planungskonzepte darstellten (Kletterpark, Hotel, Naturschutzgebiet, etc.) und frei auf der Karte platziert werden konnten. Auch wenn die aufgezeichneten Interaktionen erkennen lassen, dass die Teilnehmer das aufoktroyierte Rollenverhältnis zum Teil in spielerischer und selbstinszenierender Weise reproduzieren, sind sie auf der koordinativen Ebene gefordert, gemeinsame Orientierungen auf interaktional relevante Objekte herzustellen. Das heißt, um koordinative Vorgänge wie die Herstellung von Ko-Orientierung zu untersuchen, stellt das Aufzeichnungspostulat natürlicher Interaktionen unserer Ansicht nach keine epistemische Notwendigkeit dar. Im Gegenteil, semi-experimentelle Settings eröffnen die Möglichkeit, schwer greifbare Phänomene, die oftmals zu axiomatischen Grundbedingungen degradiert werden, mit neuen Analysemethoden zu überprüfen und ggf. zu präzisieren.

Diesem Vorhaben widmet sich die nachfolgende Arbeit, indem sie einen methodischen Ansatz verfolgt, der strukturell ähnliche Interaktionsbedingungen in den analytischen Fokus rückt, die sich im Kern darin unterscheiden, unterschiedliche Möglichkeiten zum “mutual monitoring” zu generieren: Während wir in einer ersten Studie Interaktanten baten, die oben geschilderte Aufgabe unter Face-to-Face-Bedingungen bei ganzheitlichem “mutual monitoring” zu bearbeiten, wurde dieses Koordinierungspotential in einer separaten Vergleichsstudie unter den Bedingungen von Augmented Reality (AR) eingeschränkt. Das von Dierker et al. (2009) und Neumann (2011) zur kooperativen Nutzung entwickelte „ARbInI System“ stellt eine Interaktionsumgebung bereit, in der die ko-präsenten Akteure durch das Tragen von Head-Mounted-Displays ihre orientierungsrelevanten Aktivitäten unter eingeschränktem “mutual monitoring” koordinieren müssen (sie können nicht sehen, wo ihr Ko-Partizipant hinschaut). Die in der kognitionspsychologischen Forschung prädominante Untersuchungsressource des “gaze-following” (Scaife & Bruner 1975) wird folglich unterbunden, so dass ein Lupeneffekt auf diejenigen interaktiven Verfahren hervortritt, die abseits reiner Blickkoordination bei der Herstellung von Ko-Orientierung beteiligt sind. Das verwendete AR-System bietet dabei zusätzlich eine wertvolle analytische Ressource, mit der die situativen Orientierungen und Reaktionen der Interaktanten in einem höheren Auflösungslevel rekonstruiert werden können: Die Aufzeichnung der Teilnehmerperspektiven ermöglicht zum einen die Bestimmung der zum jeweiligen Zeitpunkt visuell wahrgenommenen semiotischen Ressourcen und zum anderen – dem Garfinkelschen Paradigma folgend – die analytische Rekonstruktion sozialer Praktiken „from within“.

Im Rahmen des aufgezeigten Forschungsansatzes werden in der vorliegenden Arbeit drei Themenblöcke bearbeitet. Gemäß des konversationsanalytischen Postulats, die Untersuchungsgegenstände aus den

Daten selbst zu entwickeln (“from the data themselves”, Schegloff & Sacks 1973), werden die einzelnen Themenblöcke anhand unterschiedlicher phänomenologischer Schwerpunkte ausdifferenziert, die sich sukzessive aus den analytischen Resultaten der einzelnen empirischen Beobachtungen begründen.

Das **erste Analysekapitel (Kapitel 2)** widmet sich dem Einfluss des “mutual monitoring” auf die Verfahren zur Herstellung von Ko-Orientierung. Unter der Fragestellung „Wie wird aus der gegenseitigen Beobachtung eine objektfokussierte Interaktion?“ setzen wir “mutual monitoring” nicht als Grundbedingung voraus, sondern betrachten “mutual monitoring”-basierte Verfahren als analytisch rekonstruierbare Phänomene in unterschiedlichen Handlungskontexten. Im Rahmen der datengetriebenen Entwicklung der einzelnen Themenschwerpunkte werden hierbei u.a. die folgenden Desiderate adressiert: (1) Welche qualitativen Unterschiede lassen sich in der Ausdrucksassemblierung kommunikativer Orientierungsgestalten und damit verbundener Anforderungsspezifikationen im “mutual monitoring” sich orientierender Interaktanten erkennen? (2) Mit welchen multimodalen Verfahren werden situative Aufmerksamkeitskonkurrenzen in interpersonellen Simultanaktivitäten bearbeitet? (3) Wie unterscheidet sich die Bearbeitung interpersoneller Simultanaktivitäten unter vollumfänglichen (F2F) und eingeschränkten Bedingungen (AR) des “mutual monitoring”? Aus den empirischen Beobachtungen dieses Analysekapitels leiten sich die beiden nachfolgenden Analysekapitel ab.

Das **zweite Analysekapitel (Kapitel 3)** greift die empirischen Beobachtungen zur interaktiven Bewältigung lokal anstehender Aufgaben und konversationeller Probleme in der AR-Bedingung auf und fragt nach der Robustheit interaktiver Koordinierungsressourcen. (1) Mit welchen multimodalen Verfahren bearbeiten die Interaktanten die gesteigerten Koordinierungs- und Orientierungsanforderungen der AR-Bedingung? (2) Wie tragen die verwendeten Verfahren dazu bei, Distorsions- und Störeffekte zu minimieren? (3) Welche Rückschlüsse lassen sich daraus auf die Robustheit interaktiver Orientierungs- und Koordinierungsressourcen/-verfahren im Speziellen und interaktionaler Systeme im Allgemeinen ziehen? Das Ziel ist es somit, zentrale Stabilitätskonstituenten der sozialen Interaktion ausfindig zu machen und zu ergünden, unter welchen spezifischen Bedingungen diese Stabilität gefährdet ist.

Das **dritte Analysekapitel (Kapitel 4)** unterscheidet sich insofern von den beiden vorigen, als es Ko-Orientierung nicht aus einer mikrophenomenologischen Perspektive lokal anstehender Interaktionsaufgaben und konversationeller Probleme analysiert, sondern im Kontext handlungsschematischer Zusammenhänge betrachtet. Auf Grundlage einer empirischen Beobachtung der F2F-Bedingung, in der ein Teilnehmer seinem Ko-Partizipanten über einen längeren Zeitraum strategisch das Handlungsrecht zur Vorschlagskonstitution entzieht, fragen wir nach den übergeordneten Konstitutions- und Orientierungslogiken, aus denen sich Transformationen der interaktionalen Dynamik wie die eruptive Veränderung des intersubjektiven Beteiligungsniveaus begründen. Welchen Konstitutionsbedingungen folgt die Herausbildung größerer Struktureinheiten und woran orientieren sich die Interaktanten dabei?

Aus der Bearbeitung dieser drei Themenkomplexe ergibt sich eine zweifache Zielsetzung: Einerseits soll auf phänomenologischer Ebene ein umfassendes Verständnis für soziale Wahrnehmungspraktiken am Beispiel der Ko-Orientierung und Ko-Referenzierung gewonnen werden, das die im Eingangszitat von Heath et al. formulierte epistemische Lücke adressiert.¹ Andererseits soll auf methodologischer Ebene eine analytische Präzisierung der konversationsanalytischen Grundannahmen des “mutual monitoring” und der holistischen Rezeption kommunikativer Gestalten erfolgen, in der deutlich wird, worauf die Interaktanten in spezifischen Interaktionssituationen reagieren und welche kommunikativen Ressourcen sie innerhalb der Gestaltformen nutzen, wenn ihnen diese nicht ganzheitlich zur Verfügung stehen. Hier schließt sich die Frage an, inwieweit diese Grundannahmen eine Voraussetzung zum Orientierungserfolg und zur kooperativen Bearbeitung lokal anstehender Interaktionsaufgabe bilden.

Bevor wir in die einzelnen Themenkomplexe einsteigen, beginnen wir nachfolgend mit einem konkreten Fallbeispiel, an dem sich das übergeordnete Erkenntnisziel dieser Arbeit praktisch skizzieren lässt. Im Anschluss daran wird in den erforderlichen konzeptuellen und methodologischen Hintergrund eingeführt sowie das dem empirischen Analyseteil zugrundeliegende Studiendesign und Korpus näher erläutert.

¹ Heath et al. (2002: 318) verwenden den Begriff “awareness” nicht um auf Bewusstseinsvorgänge zu fokussieren, sondern auf die alltagspraktischen Herstellungsverfahren der Wahrnehmung und Orientierung: “[...] we wish to suggest that awareness is not simply a ‘state of mind’ or a ‘cognitive ability’, but rather a feature of practical action which is systematically accomplished within developing course of everyday activities.”

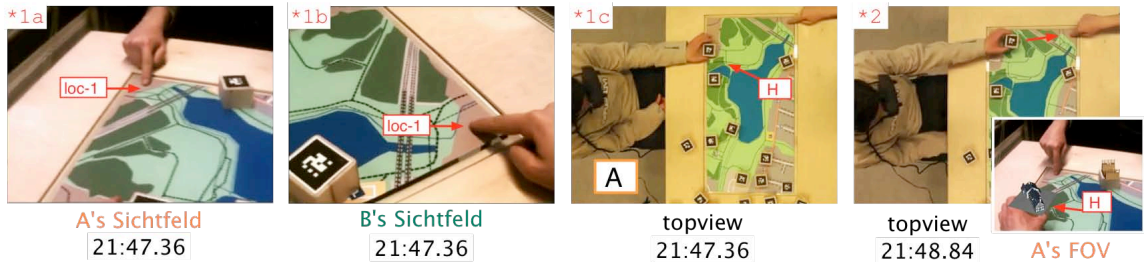
1.1 Gegenstandskonstitution

Zum Einstieg in die Thematik der interaktiven Herstellungsprinzipien von Ko-Orientierung wird nachfolgend ein Fallbeispiel aus unserem Korpus präsentiert, an dem sich die epistemische Relevanz des Forschungsgegenstandes aufzeigen und das übergeordnete Forschungsinteresse dieser Arbeit praktisch skizzieren lässt. Das Beispiel soll zugleich einen ersten analytischen Eindruck von der Spezifik des Datenmaterials ‚Augmented Reality basierte Interaktion‘ vermitteln, das neben ‚herkömmlichen‘ face-to-face Interaktionen die empirische Datengrundlage dieser Arbeit bildet. Der methodische Ansatz, ein AR-System als empirisches Werkzeug der Interaktionsforschung zu nutzen, wird in Abschnitt 1.2.3 erläutert. Für eine erste Annäherung sei an dieser Stelle daher nur darauf verwiesen, dass das System die eingeschränkten Sichtfelder (ca. 42° in der Horizontalen) der Interaktanten aufzeichnet und für postexperimentelle Analysen zur Verfügung stellt. In der multimodalen Transkription der nachfolgenden Orientierungsepisode finden sich infolgedessen – neben der externen Kameraperspektive – auch extrahierte Standbilder aus dem Videostream der Interaktanten, der als analytische Ressource zur Rekonstruktion der situativen Teilnehmerorientierungen genutzt wird (nähere Informationen, auch zur Bewertung dieser Analyseressource, folgen in 1.2.3). Während die Transkription der verbalsprachlichen Ebene den von Selting et al. (1998; 2009) vorgeschlagenen GAT-Zeichenkonventionen folgt, wurden für die sichtbare Ebene der Interaktion eigene Konventionen entwickelt. Die verwendeten Transkriptionskonventionen finden sich im Anhang. Da die Aufzeichnung der Teilnehmerperspektiven durch Head-Mounted-Displays (HMDs) realisiert wird, haben die Interaktanten keinen Zugriff auf die gegenwärtige Orientierung ihres Ko-Partizipanten (die Augen der Interaktanten sind durch die HMDs verdeckt). Insofern stellt das Setting eine herausfordernde Interaktionsbedingung dar, in der die Herstellung von Ko-Orientierung nicht auf Grundlage von Blickkoordination erfolgen kann, sondern mit den übrigen semiotischen Ressourcen bewältigt werden muss (Verbalsprache, Gestik, Objektmanipulationen, sequentielle Einbettung, etc.).

Das nachfolgende Fragment zeigt eine Orientierungsepisode von acht Sekunden, in der die interaktive Herstellung von Ko-Orientierung fehlschlägt. Dabei ist eine Divergenz zwischen dem intendierten Orientierungsangebot des gegenwärtigen Sprechers und der Interpretation des Rezipienten festzustellen, in deren Folge beide Interaktanten auf unterschiedliche Entitäten orientiert sind. Epistemisch interessant ist hierbei, dass die ‚missglückte‘ Ko-Orientierung aus einer etischen Beobachterperspektive klar als solche zu erkennen ist, aus der emischen Perspektive der Interaktanten jedoch unbemerkt und darüber hinaus interaktional folgenlos bleibt. Im Rahmen der interaktiven Konzeption des Untersuchungsgegenstandes „Ko-Orientierung“ manifestieren sich in dieser empirischen Beobachtung strukturelle Unterschiede zur kognitions-psychologischen Perspektive auf das Phänomen „Joint Attention“, aus denen sich ein erkenntnistheoretischer Nutzen der praxeologischen Erforschung des Phänomens in interaktionalen Kontexten ableiten lässt, der im Anschluss an die Analyse expliziert werden soll.

AR-Baseline:	VP-Gruppe 10 (VP-A: Investor / VP-B: Naturschützer)
Zeit:	21:46 - 21:54

```
01 B-ver: |kann man DAS |nich HIER platzieren, so nen bisschen außerhalb;
B-act:          |dG>>@loc-1.....
A-lH:  |grasp>>>>>>>|@H.....lift.....frwd>>>>>>>>>>loc-1.....
                                 *1a-c                                           *2
```



Wir steigen in die Analyse der Interaktion ein, kurz nachdem VP-A (in der Rolle des Investors) das Objekt „Hotel“ (H) als bearbeitungsrelevante Interaktionsaufgabe statuiert und einen initialen Vorschlag zur Objektplatzierung verbalisiert hat, der von seinem Ko-Partizipanten VP-B (in der Rolle des Naturschützers) abgelehnt wurde. Der Beginn des Transkripts setzt in der unmittelbar folgenden sequentiellen Anschlussbehandlung ein, in der B seiner initialen Vorschlagsablehnung die Unterbreitung eines Gegenvorschlags anschließt: Die multimodale Vorschlagskonstitution wird auf der verbalsprachlichen Ebene durch die Frage „kann man DAS nich HIER platzieren, so nen bisschen außerhalb;“ (01) explizit als aushandlungswürdiges Alternativangebot formuliert, das er in koordinativer Selbstabstimmung zwischen dem verbalsprachlichen Lokaldeiktikum „HIER“ und der Ausführung einer deiktischen Geste mit dem Verweisort „loc-1“ indexikalisiert. In den Sichtfeldern der Interaktanten ist zu erkennen, dass sie zu diesem Zeitpunkt auf den vorgeschlagenen Lokalisationsort „loc-1“ ko-orientiert sind (*1a+b). Noch im Vollzug der Vorschlagskonstitution, kurz nach der holistischen Gestaltformation des Orientierungsangebots aus prosodisch markiertem Lokaladverb „HIER“ und deiktischer Geste in der Strokephase, ist eine Reaktion A’s zu erkennen, in der er das linkshändig manipulierte Objekt „Hotel“ (H) in Richtung des von B ausgewiesenen Lokalisationsortes „loc-1“ führt (vgl. *1c und *2).

```

02 B-ver: |HIER, (.) !HIER! wär ich doch dafür;
B-act: |.....|circle..|tap.....@loc-1.....
A-ver: |ja: genau |entweder HIER |außerhalb |hier is SCHWIERig
A-lH: |bwrdr>>>>>|place(loudly)
A-rH: |dG>>>>>>@loc-2.....|tap
      *3          *4          *5

```



Kurz darauf bricht A die projizierte Ratifizierungsaktivität (Hinführung des Objektes „Hotel“ (H) auf den von B indizierten Lokalisationsort „loc-1“) ab und verbalisiert mit „ja: genau entweder HIER außerhalb“ (2) selbst einen Vorschlag zur Platzierung: Dabei zieht er das linkshändig geführte Objekt „H“ leicht zurück (*3), stellt es akustisch wahrnehmbar auf der Karte ab (*4) und weist deiktisch mit seiner rechten Hand einen alternativen Lokalisationsort aus („loc-2“), der auf der gegenüberliegenden Seite der Karte liegt (*5). Betrachten wir B’s Sichtfeld während dieser Vorschlagsepisode, wird deutlich, dass er in Fokussierung auf die kontemporär relevante Interaktionsaufgabe „Hotel“ den rechtshändig indizierten Alternativvorschlag „loc-2“ seines Ko-Partizipanten A nicht wahrnehmen kann. Stattdessen interpretiert er A’s „provisorische Platzierung“ als intendiertes Angebot und orientiert sich zum vermeintlich vorgeschlagenen Lokalisationsort (*3+*4). In der sequentiellen Folgehandlung B’s bestätigt sich diese Teilnehmerinterpretation retrospektiv: Auf den irrtümlich inferierten Platzierungsvorschlag reagierend, enaktiert B eine Relevanzsteigerung seines eigenen Vorschlags durch ein in prosodischer Nachdrücklichkeit markiertes „HIER, (.) !HIER! wär ich doch dafür;“ (02) in Kombination mit einer ostentativen Markierung aus „kreisender Zeigegeste“ („circle“) und „Fingertippen“ („tap“) auf seinen vorgeschlagenen Lokalisationsort „loc-1“, den er als Standort des „Hotels“ präferiert. Analog zur Fehldeutung B’s ist für A nicht erkennbar, dass sich B’s Reaktion auf den falschen Referenten – die „provisorische Platzierung“ (*4) des Objektes bezieht. Infolgedessen muss A davon ausgehen, dass sein Alternativvorschlag zur Objektplatzierung von B wahrgenommen und geprüft wurde.

03 A-ver: |HIER is es gut (.) genau;|
 A-lH: |place>>>>>>>>>>>>>>>>>>@loc-1|
 *6a-c



Unter der Annahme, dass sein Alternativvorschlag „loc-2“ wahrgenommen und geprüft wurde, reagiert A auf B's Insistieren, den vorgeschlagenen Standort „loc-1“ zu ratifizieren: Dabei wertet A einerseits die interaktionale Relevanz des eigenen Vorschlags mit „hier is SCHWIERig“ ab (und zwar noch im Handlungsvollzug der vehementen Fürsprache B's (02)), andererseits lenkt er verbalsprachlich mit „HIER is es gut (.) genau;“ (03) ein und platziert das „Hotel“ auf den fortwährend ostentativ ausgewiesenen Lokalisationsort „loc-1“ seines Ko-Partizipanten (*6a-c).

In der analytischen Rekonstruktion des Falles manifestiert sich eine zentrale empirische Beobachtung zum Untersuchungsgegenstand, die für die vorliegende Arbeit leitmotivisch ist. Innerhalb psychologischer Ansätze wird das Phänomen „Joint Attention“ als triadische Konstellation beschrieben, in der (mindestens) zwei Individuen ihre Aufmerksamkeit auf eine Entität koordinieren und ihre intentionalen Absichten teilen („shared intentionality“, Tomasello 1995). Unter einer kognitivistischen Perspektive wird auf die in dieser Form einzigartige menschliche Fähigkeit verwiesen, dem Gegenüber mentale Zustände zu attribuieren, diese mit den eigenen abzugleichen und im Sinne einer gemeinsamen mentalen Repräsentation für das kooperative Handlungsziel zu nutzen („mind-reading“, Baron-Cohen 1997; Lee et al. 1998 (detaillierte Ausführungen folgen in 1.2.1)). Ohne die Validität dieser kognitiven Fähigkeit in Frage zu stellen, zeigt sich im empirischen Beispiel, dass die interaktionale Realität zuerst anderen Konstitutions- und Orientierungslogiken folgt, bevor letztendlich so etwas wie ‚geteilte Intentionalität‘ entstehen kann: Die Interaktanten orientieren sich ‚ganz pragmatisch‘ an den Angebotsstrukturen der jeweiligen Handlungsaktivitäten ihrer Ko-Partizipanten – seien sie intentional oder wie im Beispiel akzidentell durch die provisorische Platzierung eines Objektes bedingt – und deuten sie sinnhaft unter Bezug auf das situative Kontextgeschehen. Interaktanten operieren somit auf Grundlage von Arbeitshypothesen, die sie aus den jeweiligen Simultan- und Anschlussreaktionen ihrer Ko-Partizipanten entwickeln. Eine intentionale Deckungsgleichheit erscheint dabei keine zwingende Voraussetzung zum kooperativen Handlungsvollzug zu bilden. So zeigt sich im Beispiel, dass die divergenten Teilnehmerinterpretationen für die Bearbeitung der interaktionalen Aufgabe, einen geeigneten Platzierungsort für das „Hotel“ zu finden, unproblematisch sind. Hierfür lassen sich im Wesentlichen zwei Gründe aufführen: Durch B's vehementes Insistieren, das Objekt auf den von ihm vorgeschlagenen Lokalisationsort „loc-1“ zu platzieren, erlangt der Alternativvorschlag seines Ko-Partizipanten keine interaktive Relevanz. Darüber hinaus zeigt sich in A's eigener Abwahl des Alternativvorschlags „loc-2“ jedoch auch, dass er nicht auf diese interaktive Relevanz drängt, sondern die Anschlussreaktion seines Ko-Partizipanten zur interaktionalen Fortsetzung nutzt und das Objekt auf den ostentativ ausgewiesenen Lokalisationsort „loc-1“ platziert.

In dieser empirischen Beobachtung manifestiert sich die epistemische Bedeutung einer praxeologischen Erforschung des Phänomens, da sie darauf hindeutet, dass Ereignisse der gemeinsamen Herstellung von Orientierungen nicht losgelöst aus ihrem situativen Entstehungskontext betrachtet werden können (vgl. Kidwell & Zimmerman 2007). Bei der Frage „Wie stellen Interaktanten Ko-Orientierung (Joint Attention) her?“, geht es in erster Linie darum, mit welchen multimodalen Verfahren sie die situative und interaktive Bearbeitungsrelevanz einer interaktionalen Aufgabe herstellen und etablieren, für die eine gemeinsame Orientierung erforderlich ist. Dieses Erkenntnisziel – für das wir die kognitions-psychologische Konzeption des „mind-reading“ durch eine praxeologische Konzeption des „action-reading“ ersetzen – soll der vorliegenden Arbeit als übergeordneter Bezugsrahmen dienen, unter den sich die in den jeweiligen Kapiteln fokussierten Phänomene zur Ko-Orientierung subsumieren lassen.

1.2 Konzeptueller und methodischer Rahmen

1.2.1 Joint Attention und Ko-Orientierung: Unterschiedliche Paradigmen eines Phänomens

Die Mechanismen der Aufmerksamkeitsmanipulation bilden als zentrales Charakteristikum der menschlichen Interaktionsfähigkeit einen Untersuchungsgegenstand, der im Interessenfokus unterschiedlicher Forschungsdisziplinen steht. Im Folgenden werden die psychologischen Ansätze im Rahmen der "Joint Attention" Forschung skizziert, um die konzeptuellen Gemeinsamkeiten und Unterschiede zum interaktionslinguistischen Paradigma der „Ko-Orientierung“ aufzuzeigen.

Innerhalb der kognitions-psychologischen Forschung beschreibt das Phänomen "Joint Attention" eine triadische Konstellation (vgl. Abbildung 1, links), in der zwei Individuen ihre Aufmerksamkeit auf eine Entität koordinieren, die im beiderseitigen Interesse liegt ("two individuals coordinate attention to an object of mutual interest", Bakeman & Anderson 1984).

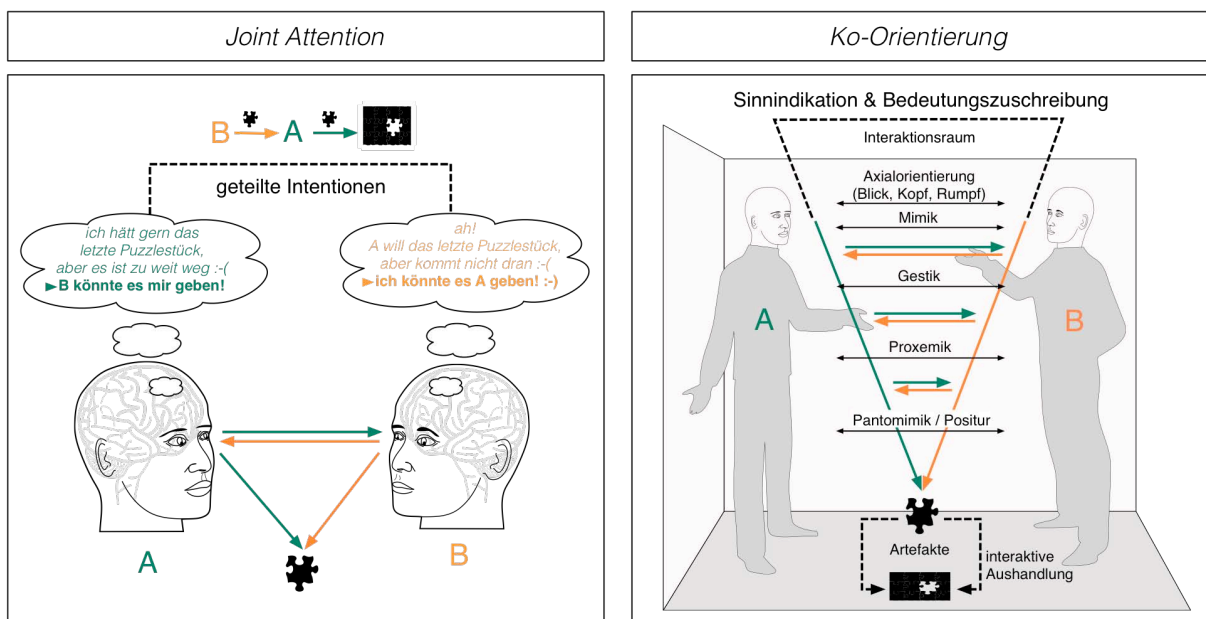


Abbildung 1: Links: Triadische Joint Attention Konzeption ("mind-reading", z.B. Baron-Cohen 1997; Lee et al. 1998) – Rechts: Ko-Orientierung als sequenzstrukturelles und multimodales Oberflächenphänomen im interaktiven Handlungsvollzug ("action-reading")

Dabei wird eine wechselseitige Angleichung postuliert, in der die Akteure ihr intentionales Verhalten aufeinander abstimmen und so die hinter der Koordinierung auf gemeinsame Entitäten liegende Absicht gegenseitig transparent machen. Dieser Sachverhalt wird bei Tomasello (1995) mit dem Begriff der "shared intentionality" gefasst. Kaplan & Hafner (2006) liefern mit Bezug auf Hobson (2002) eine griffige Definition:

"To reach joint attention an agent must understand, monitor and direct the intentions underlying the attentional behaviour of the other agent. Attention can only be reached if both agents are aware of this coordination of "perspectives" towards the world (Hobson 2002)." (Kaplan & Hafner 2006: 136)

Speziell in entwicklungspsychologischen Ansätzen wird die Indikatorfunktion von Joint Attention für mentale Prozesse betont (Camaioni 1992; Tomasello 1995) und das Phänomen im Rahmen der frühkindlichen Ontogenese als "precursor to a theory of mind" untersucht (Baron-Cohen 1991; Charman et al. 2000). Vor diesem Hintergrund bildet Joint Attention eine grundlegende Voraussetzung zur Attribuierung und Dechiffrierung mentaler Zustände und der Erkenntnis, dass eigene mentale Zustände von denen anderer abweichen können (vgl. "mind-reading", Baron-Cohen 1997; Lee et al. 1998). Die Herausbildung

dieser Fähigkeit wird als prozessuale Entwicklung skizziert, in der das frühkindliche, ab ca. sechs Monaten beginnende “gaze following” (Scaife & Bruner 1975) eine Entwicklungsstufe zum “social referencing” (Campos & Stenberg 1981; Stenberg 2003) bildet, das die Orientierung der Kleinkinder am affektiven Verhalten ihrer Bezugspersonen beschreibt. Für Kinder mit autistischen Syndromen konnten zahlreiche Studien belegen, dass die Fähigkeit zur Joint-Attention-Bildung gestört ist. Als Indiz gilt dabei u.a. ein verringerter Blickwechsel zwischen Mutter/Vater und Referenzobjekt (z.B. Charman et al. 1997). Je nach Ausprägung der Spektrumskrankheit kann die Blickrichtung anderer Personen nicht mit einem sich darin äußernden intentionalen Verhalten in Verbindung gebracht werden (vgl. “mindblindness”, Baron-Cohen 1995).

In den Studien zur Erforschung des Phänomens „Joint Attention“ fällt auf, dass visuell-perzeptive Beteiligungsweisen den prädominanten Untersuchungsfokus bilden (Bigelow 2003; Akhtar & Gernsbacher 2007; Yu & Smith 2013; Langton et al. 2000). Auch wenn “gaze following” für die ontogenetische Entwicklung von Joint Attention Fähigkeiten einen übergeordneten Stellenwert einnimmt – Brooks und Meltzoff (2005) sprechen von einer “‘front end’ ability” – ist festzustellen, dass der Untersuchungsgegenstand zwar durchaus in jüngster Zeit als multimodales Phänomen beschrieben wird, in der einschlägigen Forschungsliteratur hierzu jedoch kaum empirisch fundierte Erkenntnisse vorliegen. Gerade für eine Joint Attention Forschung, die nicht den Entwicklungsverlauf, sondern die holistischen Organisationsprinzipien des Phänomens bei ganzheitlich adoleszenter Ausprägung im Blick hat, ist eine Betrachtung, in der das Herstellen von Joint Attention in seiner multidimensionalen Komplexität erfasst wird, unabdingbar. Unter dieser perspektivischen Erweiterung stellt Tomasello heraus:

“The reason that linguistic skills are so highly correlated with joint attentional skills is that language is nothing more than another type – albeit a very special type – of joint attentional skill; people use language to influence and manipulate one another’s attention.” (Tomasello 2000: 406)

Die Beziehungskonstellation von “linguistic skills” und “joint attentional skills” ist dabei auch aus einer soziologischen Perspektive nachvollziehbar. So findet soziale Interaktion nach Goffman (1983a) statt, wenn (a) eine wechselseitige Wahrnehmung von mindestens zwei Individuen vorliegt, (b) sie die Wahrnehmung des jeweils anderen wahrnehmen, (c) einen gemeinsamen Aufmerksamkeitsfokus teilen und (d) ihr Handeln am jeweiligen Verhalten des anderen orientieren. Auch wenn die Verschränkung dieser Fähigkeiten evident erscheint, ist festzustellen, dass bei Tomasello keine mikroanalytische Beschreibung der Verfahren angestrebt wird, die empirische Hinweise über die Art der Relation zwischen “linguistic skills” und “joint attentional skills” liefern könnten. Die postulierte Zweiteilung dieser Fähigkeiten, bzw. der Klassifikation von “linguistic skills” als Spezialfall basaler Joint Attention Fähigkeiten ist insofern als theoretischer Ansatz zu verstehen, den es als eigenständige empirische Frage zu überprüfen gilt.

Unter einer praxeologischen Perspektive auf Interaktion ergänzen Kidwell & Zimmerman (2007) diesen Ansatz, indem sie an Videodaten aus einer Kindertagesstätte Joint Attention als “action” analysieren und damit auf einen Interaktionsprozess verweisen, der nicht losgelöst von der sozialen Aktivität betrachtet werden kann, in der Ereignisse von Joint Attention eingebettet sind. Ihre Analysen zeigen, dass die Interaktanten systematisch und unter Verwendung multimodaler Ressourcen gemeinsame Aufmerksamkeit auf Entitäten ihrer Umgebung in sequentiellen Ordnungsstrukturen etablieren. Im Rahmen dieses Untersuchungsparadigmas hat die konversationsanalytische Forschung in ihrer Erweiterung der multimodalen Bedeutungskonstitution (Goodwin 2000a, 2003, 2011) zahlreiche empirische Erkenntnisse hervorgebracht, in der die organisatorischen Prinzipien zur Herstellung gemeinsamer Aufmerksamkeit herausgearbeitet wurden. Interaktionsteilnehmer verwenden verschiedene semiotische Ressourcen (Text, Prosodie, Gestik, Körperdisplay), die nach Gestaltprinzipien zu holistischen Einheiten in kleinschrittiger Koordinierung zum Verhalten des Interaktionspartners konfiguriert werden, um Orientierungen und Orientierungswechsel in situ anzuzeigen und interpretierbar zu machen (Goodwin 2000a; Heath et al. 2002; Hindmarsh & Heath 2000a; Dausendschön-Gay & Krafft 2002; Detailausführung folgen in 2.1).

Der in dieser Arbeit eingenommene Standpunkt eines konversationsanalytisch-multimodalen Forschungsprogramms unterstellt a priori keine Dominanz spezifischer kommunikativer Ressourcen, sondern betrachtet sie im Zusammenhang „komplexer Äußerungsgestalten“, die holistisch wahrgenommen und rezipiert werden (Dausendschön-Gay & Krafft 2002; Goodwin 2000a, 2003). Mit Verweis auf das analytische Einstiegsbeispiel (vgl. 1.1) stellen wir daher dem traditionellen Ansatz des “mind-reading” (z.B. Baron-Cohen 1997; Lee et al. 1998) die Konzeption des “action-reading” gegenüber (Abbildung 1, rechts).

Dabei ist der kommunikative Erfolg nicht an eine "shared intentionality" gebunden, sondern an Sinnzuschreibung von Handlungen und wechselseitigen Deutungsprozessen (vgl. Weber 1984), indem das eigene Handeln sinnhaft konstituiert, als eine sinnhafte Handlung projiziert und auf Rezipientenseite sinnhaft gedeutet wird. Ob dabei letztendlich geteilte Intentionen vorliegen, ist für den eigentlichen Zuschreibungsprozess erstmal sekundär und bedarf bei etwaigen Abweichungen weiterer Aushandlungen der Interaktionspartner oder einer retrospektiven Anpassung der initialen Sinnzuschreibung. Interaktionspartner sind danach in der Lage, mit Arbeitshypothesen zu operieren, die nicht zwangsläufig eine intentionale Deckungsgleichheit beider Interaktionspartner abbilden müssen. Aus diesem Grund verwenden wir den Begriff der „Ko-Orientierung“, der im Vergleich zu Joint Attention nicht auf intrinsische Zustände verweist, sondern das Phänomen an der kommunikativen Oberfläche sequenzanalytisch untersuchbar macht:

„Wenn Raum stets ein Wahrnehmungspotential für die Sinnesorgane der Anwesenden darstellt, ergibt sich daraus eine Daueranforderung für die Interaktion: Die für die Interaktion gerade relevanten Wahrnehmungen im Raum müssen ausgewählt und etabliert werden. Ich nenne das „Ko-Orientierung“. Ko-Orientierung vollzieht sich in Relation zu den Sinnesorganen der Anwesenden: oben und unten, vorne und hinten, rechts und links – Ausdrücke wie diese liefern Beispiele für Orientierungen, wie sie für die Navigation im Nahbereich der Interaktion unerlässlich sind. Die Ko-Orientierung greift als Aspekt der Situierung vor allem auf die Sensorik der Anwesenden zurück.“ (Hausendorf 2010: 178)

Innerhalb dieses Paradigmas stellt Joint Attention kein temporäres Ereignis dar, wie es insbesondere der klassisch-triadische Ansatz ("two individuals coordinate attention to an object of mutual interest" (Bakeman & Adamson 1984)) im Rahmen des prototypischen "gaze-following" postuliert (Scaife & Bruner 1975), sondern einen sequentiell geordneten Prozess, der von den Teilnehmern interaktiv, multimodal und unter kleinschrittiger Koordinierung orientierungsrelevanter Verfahren hervorgebracht und ‚gesteuert‘ wird.

1.2.2 Die ethnomethodologische Konversationsanalyse multimodaler Prägung

Die Analytik der vorliegenden Arbeit greift auf die theoretischen und methodischen Prinzipien der ethnomethodologischen Konversationsanalyse in ihrer Erweiterung der multimodalen Bedeutungskonstitution zurück. Die Grundannahmen dieses Forschungsprogramms werden nachfolgend expliziert, um die in dieser Arbeit eingenommene Perspektive auf den Untersuchungsgegenstand ‚Interaktion‘ und das analytische Vorgehen einordnen zu können. Für eine ausführliche Darstellung sei an dieser Stelle insbesondere auf die folgenden Aufsätze und Monografien verwiesen, denen die nachfolgende Übersichtsdarstellung zugrunde liegt: Bergmann (1981, 1994); Goodwin & Heritage (1990); ten Have (2004, 2007); Hutchby & Wooffitt (2008).

Die Konversationsanalyse (KA) bezeichnet einen Untersuchungsansatz, der sich aus der von Harold Garfinkel (1967) begründeten Ethnomethodologie abgespalten und mit einem Fokus auf ‚soziale Interaktion‘ als eigenständige Forschungsrichtung ab Mitte der 60er Jahre entwickelt hat. Die soziologischen Leitmotive der Ethnomethodologie sind für das Theorie- und Methodenverständnis der Konversationsanalyse grundlegend. In seinem zentralen Werk "Studies in Ethnomethodology" (1967) zeigt Garfinkel, dass die soziale Wirklichkeit und sinnhafte Ordnung der Alltagspraxis („gelebte Geordnetheit“) kein rein exogenes und statisches Konstrukt darstellt, das durch normative Strukturen und Zwänge die Handlungsentwürfe der Interaktanten prägt (vgl. Pearsons' Strukturfunktionalismus²), sondern maßgeblich von den Interaktanten selbst – im lokalen und situativen Handeln – endogen erzeugt und fortlaufend bestätigt wird. Garfinkel rückt damit die „Vollzugswirklichkeit“ (Bergmann 1994, bei Garfinkel: "ongoing accomplishment") ins definitorische Zentrum seines handlungstheoretischen Ansatzes. Im Interessenfokus stehen die methodischen Prinzipien der Akteure (sog. "Ethno-Methoden"), mit denen sie ihre Handlungen sinnhaft gestalten. Aufgrund ihrer Methodizität ist die Sinnindikation des sozialen Handelns durch ver-

² Goodwin & Heritage (1990: 284): "Central to the Parsonian perspective was a tripartite analytic conception of cultural, personality, and social systems in which cultural values, once internalized as personality dispositions, were conceived as the causal drivers of social behavior". Garfinkels Kritik an der Pearson'schen Perspektive des regelgesteuerten Akteurs kommt besonders in seiner Porträtierung als "cultural dope" (Trottel, Idiot) zum Ausdruck: "By 'cultural dope' I refer to the man-in-the-sociologist's-society who produces the stable features of the society by acting in compliance with preestablished and legitimate alternatives of action that common culture provides" (Garfinkel 1967: 68).

schiedene Strukturmerkmale beschreibbar. Das grundlegende Konstitutionsprinzip liegt in Garfinkels Beobachtung begründet, dass sich Interaktanten die Sinnhaftigkeit ihrer Handlungen durch entsprechende Markierungsleistungen im Vollzug anzeigen und dabei kontinuierlich Hinweise auf die Interpretation ihrer Handlungen liefern:

“Ethnomethodological studies analyze everyday activities as members’ methods for making those same activities visibly-rational-and-reportable-for-all-practical-purposes, i.e., “accountable”, as organizations of commonplace everyday activities.” (Garfinkel 1967: vii)

Die “accounting practices” der Sinnkonstitution haben nach Garfinkel einen „reflexiven“ und „indexikalen“ Charakter. Die Reflexivität verweist hierbei auf die selbst-explizierenden Eigenschaften sozialer Handlungen: Durch “accounts” liefern die Interaktanten Verstehenshinweise auf den Sinn ihrer Handlungen und die “accounts” werden wiederum selbst im Vollzug der Handlungen bestätigt. Die Indexikalität hingegen referiert auf die Kontextsensitivität sozialer Handlungen: Die Sinnhaftigkeit der Handlungsentwürfe kann nur unter den spezifischen kontextuellen Bedingungen der situativen Vollzugswirklichkeit erfasst werden. Die alltagspraktischen Ordnungsleistungen (“Ethno-Methoden”), mit denen die Handelnden ihre „Wirklichkeitsdefinitionen in der aktuellen Interaktion aufeinander abstimmen“ (Bergmann 1994: 6), bilden den Untersuchungsfokus ethnomethodologischer Forschung.

Das ethnomethodologische Paradigma wird Mitte der 60er Jahre von Harvey Sacks und Emanuel Schegloff aufgegriffen und in der von ihnen begründeten Konversationsanalyse (KA) am Untersuchungsgegenstand der ‚sozialen Interaktion‘ ausdifferenziert. Die KA kann somit als praktische Realisierung des ethnomethodologischen Forschungsprogramms verstanden werden, die einen definierbaren Teilbereich der sozialen Realität fokussiert: “[...] Sack’s approach, in its focus on the analysis of naturally occurring talk-in-interaction, represents the most fruitful means of doing ethnomethodological study” (Hutchby & Wooffitt 2008: 27; s.a. ten Have 2004: 25f.). Im Untersuchungsfokus der KA steht die analytische Rekonstruktion der generativen Verfahren, mit denen die Interaktionsteilnehmer das interaktive Geschehen sinnhaft strukturieren, um ihre Handlungsvollzüge wechselseitig wahrnehmbar und interpretierbar zu gestalten. Bergmann (1981) formuliert das Erkenntnisinteresse der KA folgendermaßen:

„Ihr Ziel ist es, diejenigen Verfahren empirisch zu bestimmen, mittels derer die Teilnehmer an einem Gespräch im Vollzug ihrer (sprachlichen) Handlungen die Geordnetheit der (sprachlichen) Interaktion herstellen, das Verhalten ihrer Handlungspartner auf die in ihm zum Ausdruck kommende Geordnetheit hin analysieren und die Resultate dieser Analysen wiederum in ihren Äußerungen manifest werden lassen.“ (Bergmann 1981: 15f.)

Im methodischen Selbstverständnis der KA bilden Aufzeichnungen von ‚natürlichen‘ Interaktionen die Datengrundlage der strikt empirischen Analyse, um das authentische Teilnehmerverhalten und die „gelebte Geordnetheit“ in kommunikativen Vorgängen möglichst unverfälscht untersuchen zu können. Analyseleitendes Prinzip bildet hierbei die emische Rekonstruktion sequentieller Ordnungsleistungen “from within” (Garfinkel 1967). Das heißt, die empirische Analyse erfolgt nicht aus einer etischen Beobachterperspektive, sondern aus der vollzugsrekonstruktiven Teilnehmerperspektive der Interaktanten selbst, um die inhärente Logik der situativen und sequentiellen Emergenz interaktionaler Ereignisse angemessen erfassen zu können.

Dieser Anspruch bringt einige methodische Implikationen mit sich, die sich einer starren methodologischen Fixierung entziehen. Als kontextsensitives Geschehen in der Zeit haben interaktionale Ereignisse einen transformatorischen Charakter. Spezifische Interaktionsituationen sind durch eine nicht-replizierbare Einzigartigkeit gekennzeichnet, die durch die strikte Anwendung methodischer Regeln in ihrer genuinen Gestalt nicht erfasst werden kann. In Anlehnung an Garfinkels Postulat des “unique adequacy requirement”³ wird die Festlegung einer Methodologie aus diesem Grund in der KA abgelehnt. Der bewusste Verzicht auf eine Methodologie ist jedoch nicht mit einer methodischen Willkür gleichzusetzen. Auf Grundlage der einschlägigen Forschungsliteratur und etablierter Arbeitspraktiken hat sich eine „analytische Mentalität“ (Schenkein 1978) herausgebildet, in der auf exemplarische Formen der Datenaufberei-

³ “For methods of investigation Garfinkel postulated a ‘unique adequacy requirement’, which means that methods must be so fashioned that they are uniquely suited to their object – but this can only be decided after information about the object of investigation has successfully been obtained, which therefore makes any formalization impossible” (Bergmann 2004 :32).

tung und -analyse zurückgegriffen werden kann, die sich für die konversationsanalytische Zielsetzung als besonders fruchtbar erwiesen haben:

Den Ausgangspunkt der empirischen Analyse bildet die methodische Fixierung interaktionaler Ereignisse. Während zu Beginn der konversationsanalytischen Forschung in der Regel an verbalsprachlichen Daten gearbeitet wurde (Sacks und Schegloff arbeiteten an Telefongesprächen), geriet mit dem Aufkommen der Videotechnik und der Möglichkeit, Interaktionen audio-visuell fixieren zu können, zunehmend eine multimodale Perspektive auf Interaktion in den analytischen Fokus. Der Datenaufzeichnung folgt in einem nächsten Arbeitsschritt ihre Verschriftlichung (Transkription):

„Im Vorgang der Transkription gilt es, das aufgezeichnete Rohmaterial nicht von scheinbar irrelevanten Bestandteilen zu reinigen, sondern in seinen Details zu bewahren, d.h., mit allen Dialektismen, Intonationskonturen, Versprechern, Pausen, Unterbrechungen etc. zu erhalten.“ (Bergmann 1994: 10)

Der Transkription kommt hierbei sowohl eine dokumentarische als auch analytische Funktion zu. In ihrer dokumentarischen Funktion fungieren Transkripte als empirische Belegstellen, um die rekonstruierten Phänomene einer wissenschaftlichen Community zugänglich zu machen. In ihrer analytischen Funktion hingegen bieten sie die Möglichkeit, ein komplexes Interaktionsgeschehen in seine funktionalen Komponenten zu zergliedern, um es in seiner sequentiellen (verbalsprachliche Fokussierung), bzw. sequentiell-simultanen Verlaufsform (multimodale Fokussierung) rekonstruieren zu können. Insofern kann im Vorgang des Transkribierens ein Detailverständnis der Orientierungslogiken und Ordnungsprinzipien erlangt werden, mit denen die Interaktanten ihre Handlungen sinnhaft aufeinander beziehen.⁴

Für die angestrebte Strukturexplication der Teilnehmerverfahren sind einige analyseleitende Arbeitsprinzipien von zentraler Bedeutung. Da die Konversationsanalyse nicht darauf abzielt, das Interaktionsgeschehen extern zu deuten, Motivlagen zu eruieren oder das Teilnehmerverhalten durch methodologisch fixierte Analysekategorien zu bewerten, werden keine vorgefertigten Hypothesen an die Analyse herangebracht, sondern relevante Fragestellungen, Analyse- und Beschreibungskategorien aus dem Datenmaterial selbst entwickelt. Das Datenmaterial wird im Zuge dessen auf rekurrent auftretende strukturelle Regelmäßigkeiten hin untersucht. Einen analytischen Ansatzpunkt bilden interaktionale Phänomene, in denen die Interaktanten spezifische konversationelle Probleme und Aufgaben bearbeiten („guided doings“, Goffman 1974: 22ff.; Sacks & Schegloff 1973). In diesen Bearbeitungsverfahren (z.B. die interaktive Bearbeitung der Frage „What to do next?“ (Garfinkel 1967)) manifestieren sich die organisationalen Prinzipien, mit denen die Teilnehmer das interaktionale Problem, bzw. die Aufgabe interaktiv bewältigen und so das Interaktionsgeschehen sinnhaft und nachvollziehbar strukturieren („order at all points“, Sacks 1984: 22). Das Phänomen wird daraufhin rekonstruiert, methodisch expliziert und auf sein rekurrentes Auftreten im Datenkorpus hin überprüft. Ausgehend vom Einzelfall werden Fallkollektionen erstellt, in denen sich das Phänomen in verschiedenen Fallalternationen manifestiert. Der Erstellung von Fallkollektionen zu spezifischen Bearbeitungsstrategien anstehender Interaktionsaufgaben/konversationeller Probleme kommt hierbei eine zweifache Funktion zu: Einerseits geben Fallkollektionen Aufschluss über die Variationsbreite und kontextuelle Abhängigkeit der Bearbeitungsstrategien, andererseits ermöglichen sie eine vom spezifischen Einzelfall abstrahierende Beschreibung der alltagspraktischen Verfahren, mit denen die Teilnehmer die strukturelle Geordnetheit der Interaktion herstellen.

Der in dieser Arbeit zugrundeliegende multimodale Ansatz stellt eine konzeptionelle Weiterentwicklung des konversationsanalytischen Forschungsprogramms dar, der auf die durch audio-visuelle Medien sichtbar gewordene multimodale Komplexität interaktionaler Ereignisse reagiert:

„Die Notwendigkeit, sich mit dem für Kommunikation konstitutiven Zusammenspiel aller Ausdrucksebenen zu beschäftigen, entsteht [...] nicht zuletzt durch den technologischen Wandel und dessen Folgen für die konversationsanalytische Forschungspraxis. In den Anfängen der Konversationsanalyse wurden für die Untersu-

⁴ Unter einer multimodalen Perspektive auf Interaktion ist hierbei jedoch anzumerken, dass der analytische Nutzen der Transkription mit dem zeitlichen Aufwand ihrer Erstellung konfligiert. Insbesondere multimodale Transkriptionen, in denen die kommunikativen Ressourcen der Interaktanten holistisch erfasst werden (Sprache, Körperorientierung, Gestik, Blick, etc.), sind extrem zeitaufwändig und für Korpora mit mehreren Stunden Aufnahmematerial kaum zu realisieren. Der multimodale Ansatz erfordert daher einen stärkeren preanalytischen Zuschnitt auf das im Forschungsfokus liegende Phänomen und eine Sensibilisierung, dieses Phänomen auch im Rohmaterial aufspüren zu können, um es dann sequenzanalytisch in der Transkription zu zerlegen und anschließend zu entscheiden, ob es sich tatsächlich um eine Fallalternation des Phänomens handelt.

chung der sozialen Organisation von Gesprächen in der Regel nur Tonaufnahmen herangezogen. Die technischen Bedingungen (Restriktion auf die auditiven Informationen, das Hörbare) führten dabei zu einer – zwar nicht theoretisch motivierten, aber doch analysefaktisch folgenreichen – Priorisierung des Gesprochenen gegenüber anderen interaktionsrelevanten Formen körperlichen Ausdrucks.“ (Schmitt 2005: 21)

Die multimodal ausgerichtete Konversationsanalyse wendet sich somit explizit gegen die Priorisierung des Gesprochenen, indem sie interaktionale Phänomene als einen „ganzheitlichen und letztendlich von der Körperlichkeit der Beteiligten nicht zu trennenden Prozess begreift“ (ebd.: 18f.). Die Fokussierung auf multimodale Phänomene liegt jedoch nicht im ‘common sense’ koexistenter Ausdrucksweisen begründet, sondern in der empirischen Erkenntnis, dass die in situ verwendeten semiotischen Körper- und Umgebungsressourcen ein kommunikatives Ganzes bilden (“contextual configuration”, Goodwin 2000a, 2003a), dessen Bedeutungskonstitution sich erst aus dem Zusammenwirken der Einzelkomponenten begründet:

“[...] actions are both assembled and understood through a process in which different kinds of sign phenomena instantiated in diverse media, what I call semiotic fields, are juxtaposed in a way that enables them to mutually elaborate each other.“ (Goodwin 2009: 21)

Im Postulat der multimodalen Bedeutungskonstitution manifestiert sich gleichermaßen eine theoretische wie analytisch-praktische Zuspitzung des indexikalen Charakters interaktionaler Ereignisse: Die verschiedenen Ausdrucksmodi werden aus epistemischen Gründen vollständig egalisiert (daher ihre generalisierende Bezeichnung als “semiotic fields”), da ihre Signalisierungsfunktion nicht per se gegeben ist, sondern auf die kontextuellen Anforderungen der spezifischen Situation reagiert. Diese strukturelle Gleichwertigkeit und funktionale Adaptivität semiotischer Ressourcen (verbal, nonverbal, materiell) ist in Goodwins Konzeption der “ecology of sign systems”⁵ (z.B. Goodwin 2003a: 35ff.) verankert und wird an unterschiedlichen Beispielen empirisch ausdifferenziert: Am Beispiel ‚hopscotch‘-spielender Kinder (im deutschsprachigen Raum auch als Hüpfspiel „Himmel-und-Hölle“ bekannt) zeigt Goodwin (2000a), dass der Verweis auf einzelne Zahlenfelder in Abhängigkeit zur situativen Orientierung des Ko-Partizipienten erfolgt und daher unterschiedliche Möglichkeiten der Realisierung präferiert werden (Zeigegesten mit dem Arm vs. Zeigegesten mit dem Fuß, wenn der Ko-Partizipient bereits auf die Markierungslinien am Boden orientiert ist). In diesem Zusammenhang sei vor allem auch auf Goodwins Arbeiten zur multimodalen Bedeutungskonstitution in der Interaktion mit einem Aphasiker (z.B. 1995, 2000b, 2003c) verwiesen, in denen das interaktionale Potential der wechselseitigen Substitution von Signalisierungssystemen deutlich wird: Der partielle Ausfall der sprachlichen Modalität führt nicht zu einer ‚Unmündigkeit‘ des Aphasikers, sondern dazu, dass andere Signalisierungssysteme (prosodische und gestische Markierungen, situative Relevantsetzungen der materiellen Raumstruktur, etc.) die verbalsprachliche Bedeutungskonstitution zum Teil kompensieren und zu einem funktionalen Sprachsystem herausgebildet werden.

Die empirische Erkenntnis, dass die einzelnen Ausdrucksmodi in den lokal wechselnden “contextual configurations” unterschiedliche Funktionen übernehmen können, ist für die multimodal ausgerichtete Konversationsanalyse von zentraler Bedeutung. Goodwins konzeptuelles Framework der “ecology of sign systems” bietet dabei einen theoretisch wie analytisch relevanten Ansatzpunkt, der den Analytiker dazu zwingt, die Bedeutungskonstitution in der Wechselseitigkeit aller zum Ausdruck kommenden semiotischen Ressourcen innerhalb einer spezifischen Situation zu rekonstruieren und auf ihre Signalisierungsfunktion – im und nach dem Handlungsvollzug – beim Ko-Partizipienten hin zu überprüfen.

Für eine erste Einordnung soll diese Darstellung der multimodalen Perspektive auf Interaktion genügen, da sie im Verlauf der Arbeit – mit Bezug auf die jeweiligen Forschungsfragen – sukzessiv durch zentrale Konzeptionen der multimodal ausgerichteten Konversationsanalyse erweitert wird (“mutual monitoring”, „Koordination“, „multimodale Verdichtungsräume“, „Sequentialität und Simultaneität“, „Deixis und räumliche Orientierung“).

⁵ Goodwin (2003a: 36): “The term ecology is used to note the way in which these separate systems function as differentiated, interdependent components of a larger whole that can adapt to changing circumstances.”

1.2.3 Augmented Reality als empirisches Werkzeug der Interaktionsforschung

In konversationsanalytischen Studien aus den Forschungsbereichen der “Workplace Studies”, “Computer Supported Cooperative Work” (CSCW), Fremdspracheninteraktion und der Interaktion bei aphasischen Störungen wurde die interaktive Relevantsetzung und Bedeutungskonstitution materieller Umgebungsressourcen als integraler Bestandteil alltäglicher Kommunikationssituationen untersucht. Basierend auf authentischem Datenmaterial natürlichsprachlicher Interaktionen zeigen die in diesen Disziplinen gewonnenen Erkenntnisse, wie es Interaktanten unter erschwerten Interaktionsbedingungen gelingt, kommunikative Ressourcen dynamisch anzupassen, um interaktionale Aufgaben bearbeiten zu können (vgl. 2.1).

Während natürliche Interaktionsbedingungen nicht die Möglichkeit bieten, valide zu bestimmen, welche visuell wahrnehmbaren Ressourcen für die Interaktanten eine Orientierungsfunktion bilden, untersuchen wir das Phänomen der Ko-Orientierung in einem semi-experimentellen Setting unter den Bedingungen von Augmented Reality (AR). Nach Azuma (1997) ist AR als eine Technologie definiert, in der die physische Realität mit virtuellen 3D-Komponenten (Objekte, Entitäten, etc.) in Echtzeit kombiniert wird. Das von Dierker et al. (2009) und Neumann (2011) entwickelte „ARbInI System“ für Echtzeitinteraktionen kopräsender Akteure (vgl. Abbildung 2) bietet die Möglichkeit, Objekte, auf die im Versuch Ko-Orientierung hergestellt wird, in ihrer Bewegung und räumlichen Konfiguration exakt bestimmen zu können. Die Entwicklung des Systems als “interception and manipulation device” (Dierker et al. 2009; Schnier et al. 2011a; Pitsch et al. 2013) ermöglicht zudem, audio-visuelle Perzeptionen der Teilnehmer online zu manipulieren und sensorische Informationen der Interaktanten abzufangen und für postexperimentelle Analysen aufzuzeichnen. Die Manipulationsfunktion ist dabei bidirektional konzipiert: Neben der perzeptorischen Manipulation besteht auch die Möglichkeit, materielle Umgebungsressourcen der Interaktanten zu verändern, um die Auswirkungen einer herausfordernden “ecology of sign systems” (z.B. instabile Referenzobjekte) auf die multimodalen Verfahren der Teilnehmer bei spezifischen interaktionalen Aufgaben (z.B. der Herstellung gemeinsamer Orientierung) untersuchen zu können. Das verwendete AR System kann damit gemäß der psycholinguistischen Tradition systematischer Manipulation von Interaktionsbedingungen genutzt werden, ist jedoch durch seine Fokussierung auf freie dyadische Interaktion weniger strikt an ein experimentelles Untersuchungsdesign gebunden. Infolgedessen ist unser Korpus aus semi-experimentellen Studien entstanden, in denen das System als empirisches Werkzeug für interaktionslinguistische Forschung genutzt wurde (Pitsch et al. 2013).

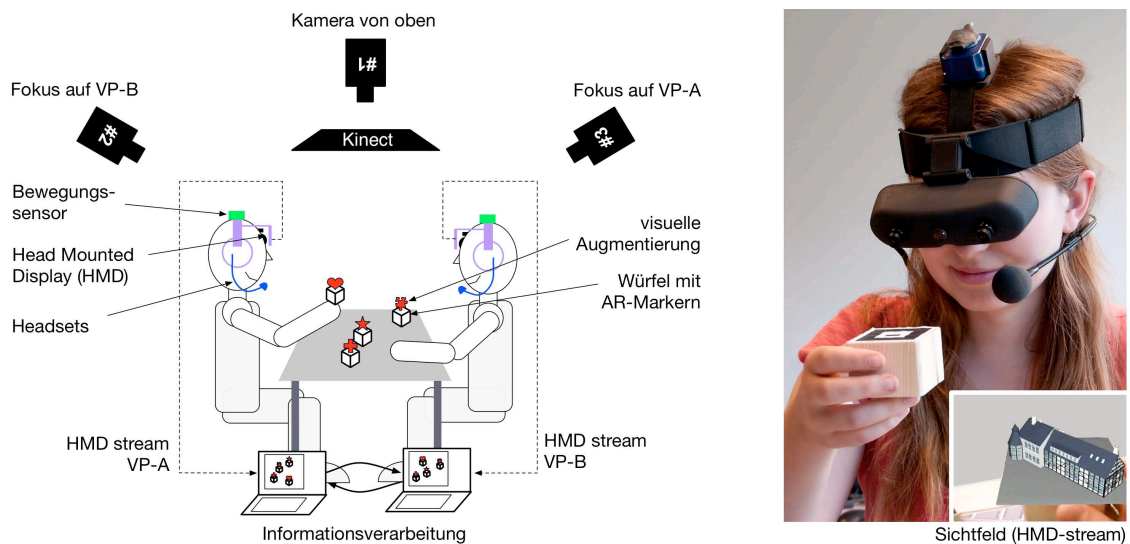


Abbildung 2: AR-System (links) & Versuchsperson mit technischem Equipment (rechts)

Die für unsere Forschungszwecke zentrale Komponente des Systems besteht aus zwei gekoppelten Head Mounted Displays (Trivision ARvision 3D) mit einem horizontalen Sichtfeld von je $42,2^\circ$ (in der Diagonale 52°). Die Sichtfeldeinschränkung von annähernd 180° in natürlichen face-to-face Situationen auf $42,2^\circ$

stellt dabei eine Interaktionsbedingung dar, unter der die Fähigkeit des “mutual monitoring” bei aufgabenfokussierter Interaktion stark eingeschränkt ist (Schnier et al. 2011a,b). Unser Setting fokussiert somit genau diejenigen Ressourcen zur Herstellung von Ko-Orientierung, die abseits reiner Blickkoordination (“gaze following”, Scaife & Bruner 1975) beobachtbar sind. Visuell-perzeptive Beteiligungsweisen, die den prädominanten Untersuchungsfokus der kognitions- und entwicklungspsychologischen Forschung zu “joint attention” bilden (vgl. 1.2.1), werden somit aus heuristischen Gründen unterbunden.

Die verwendeten Head Mounted Displays bieten dabei zugleich eine wertvolle analytische Ressource: Die Aufzeichnung der Teilnehmerperspektive ermöglicht zum einen die Bestimmung der zum jeweiligen Zeitpunkt wahrgenommenen semiotischen Ressourcen und eröffnet zum anderen – dem Garfinkelschen Paradigma folgend – die analytische Rekonstruktion sozialer Praktiken “from within”. Dabei ist darauf hinzuweisen, dass die Bestimmung semiotischer Ressourcen im jeweiligen Sichtfeld der Teilnehmer nicht zwingend Informationen über den gegenwärtigen Aufmerksamkeitsfokus des Interaktanten liefert, da Fixationen während sakkadischer Augenbewegungen durch ein Kamerabild nicht repräsentiert werden können. In diesem Zusammenhang konnten Kollenberg et al. (2010) jedoch nachweisen, dass das Tragen von Head Mounted Displays mit Sichtfeldeinschränkung zu verminderten Augenrotationen führt, so dass Aufmerksamkeitsverlagerungen primär durch Modifikationen der Kopforientierung vollzogen werden. Blickrichtungsänderungen zu salienten Objekten/Artefakten der Interaktionsumgebung bilden somit einen guten Indikator für eine sich verlagernde Aufmerksamkeit der Interaktanten.

1.2.4 Studiendesign

In unserem semi-experimentellen Setting baten wir Interaktanten in dyadischer Konstellation, die Neugestaltung des Bielefelder Naherholungsgebiets “Obersee” auszuhandeln. Um Aushandlungsprozesse anzuregen, sollten sie im Versuch aus der Position zwei sich adversativ zueinander verhaltender Gruppen agieren (Naturschützer vs. Investor) und die der jeweiligen Gruppe spezifischen Interessen durchsetzen.

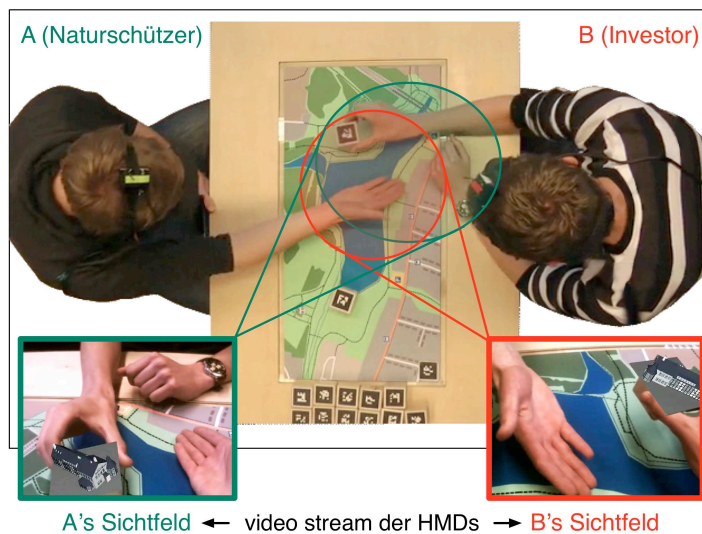


Abbildung 3: AR-System im Gebrauch mit dem Obersee II Szenario (2011/2012)

Neben einer topographischen Karte des Naherholungsgebiets, stellten 18 Objekte verschiedene potentielle Planungsvorhaben zur Attraktivitätssteigerung des Areals dar (z.B. Kletterpark, Naturlehrpfad, Quadpark, Hotel etc.). Die Auswahl der Objekte war dabei hinsichtlich ihrer inhärenten Interessenspezifität (“Profitsteigerung” (Investor) und “Naturerhaltung” (Naturschützer)) ausgewogen. In der AR Bedingung wurden die Planungskonzepte durch virtuelle Augmentierungen auf physisch greifbare Holzwürfel realisiert. In einer Vergleichsbedingung ohne Augmented Reality (Face-to-Face) wurden 3D-Prints der virtuellen Objekte aus der AR-Bedingung erstellt oder Objekte mit strukturell ähnlichen Merkmalspezifikationen aus dem Miniatureisenbahnbereich ausgewählt und auf den Holzwürfeln befestigt. Für die Aushandlung zur Neugestaltung des Obersees hatten die Interaktanten ca. 20 Minuten Zeit. Um einen Anreiz zur

fokussierten Bearbeitung der Aufgabe zu schaffen, wurde die Versuchspersoneninstruktion mit der Anmerkung abgeschlossen, dass die wartende Lokalpresse in ca. 20 Minuten Auskunft über die Neugestaltung des Bielefelder Oberseereals erwartet. Dieser fiktive Hinweis sollte die implizite Annahme generieren, dass die fokussierte Bearbeitung der Aufgabe notwendig ist, da etwaige postexperimentelle Tests, Befragungen o.Ä. nach Abschluss der bevorstehenden Experimentphase folgen.

Die aus den Studien gewonnenen Versuchsreihen weisen eine Anhäufung an Interaktionssituationen auf, die für das Forschungsvorhaben der vorliegenden Arbeit relevant sind: Interaktanten konstituieren multimodale Vorschlagshandlungen, in denen Objekte physisch manipuliert werden, um den Beginn eines vorgeschlagenen Objekts/Planungskonzepts gegenseitig wahrnehmbar zu gestalten. Da der Prozess des "where to look next?" in starker Abhängigkeit zur jeweiligen Aufgabe steht (Yarbus (1967)), bilden Vorschlagshandlungen eine Untersuchungsdomäne, in der das Phänomen der Ko-Orientierung durch seine zeitliche Expansion analytisch gut zugänglich ist. So wurden Vorschläge in der Literatur als Aktivitäten beschrieben, die vom Vorschlaggebenden durch vorbereitende Handlungen eingeführt werden, um dem Ko-Partizipanten zu signalisieren "[...] the forthcoming action will be a 'delicate' action [...] which the recipient may not like very much accept [...]" (Houtkoop-Steenstra 1990). Vor diesem Hintergrund bildet die kommunikative Gattung von Vorschlagshandlungen einen Lupeneffekt auf die Verfahren zur Herstellung von Ko-Orientierung, da die gesellschaftlich fossilisierten und formalisierten Lösungen des kommunikativen Problems ‚Vorschlagsunterbreitung‘ (s.a. Luckmann 1986) in ihrer Form eine zeitliche Expansion aufweisen, die den sequentiellen Charakter von Ko-Orientierungsereignissen offen legt.

1.2.5 Korpus

Die im C5 Projekt "Alignment in AR-based cooperation" (SFB-673/DFG) durchgeführten Versuchsreihen bilden den "Obersee II"⁶ Korpus, in dem in verschiedenen Studien die Interaktionsbedingungen der Interaktanten manipuliert wurden (Abbildung 4; vgl. Pitsch et al. 2013).

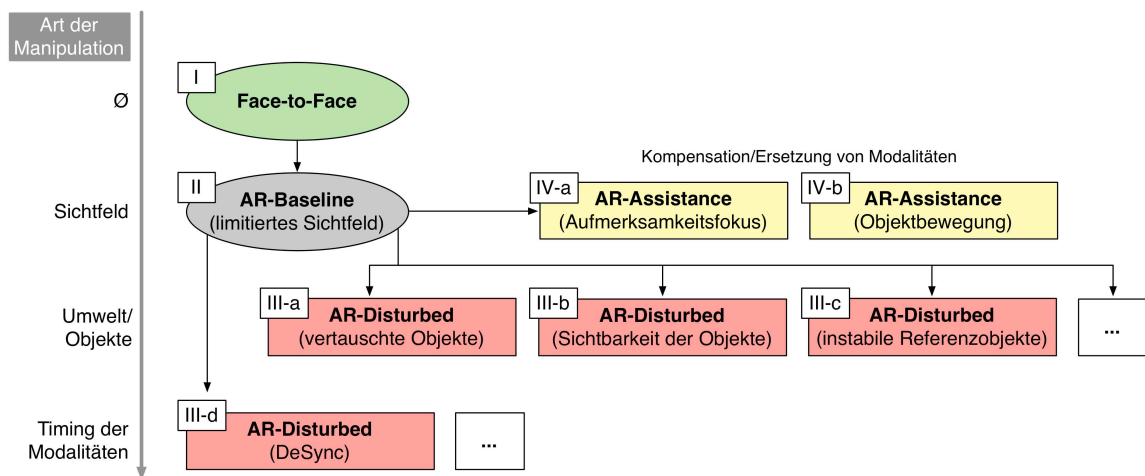


Abbildung 4: Korpusstruktur

- I. **Face-to-Face** (12 Dyaden): Baseline-Studie, in der die Teilnehmer in einer konventionellen face-to-face Situation agieren und lediglich mit Bewegungssensoren auf dem Kopf ausgestattet sind. Die im Versuch ausgehandelten Objekte/Planungskonzepte sind homomorphe, physische Modelle der virtuellen Objekte der AR-Baseline-Studie.

⁶ Das Obersee Szenario ist in Kollaboration mit dem C1 Projekt des SFB und den Central Lab Facilities (CLF) des CITEC entwickelt worden und wurde im Rahmen von Studien zur Kombination konversationsanalytischer Methodik und kinesischer Bewegungserfassung ("MotionCapture") exploriert (Pitsch et al. 2010; Brüning et al. 2011, 2012). Die positiven Erfahrungen mit dem Setting bzgl. Teilnehmerbeteiligung und Potential für interaktive Aushandlungsprozesse haben zu einem Wechsel des initialen Prototypenszenarios "Exhibition Design" auf eine AR-adaptierte Form des Obersee Szenarios geführt („Obersee II“).

- II. **AR-Baseline** (10 Dyaden, 1 Triade): Baseline-Studie, in der die Teilnehmer durch das Tragen von Head Mounted Displays (HMDs) unter einer Sichtfeldeinschränkung von 42,2° (anstelle von ca. 180° bei ungestörter visueller Perzeption) agieren. Die HMDs verhindern gegenseitige Blickkoordination und schränken das “mutual monitoring” der Interaktanten ein.
- III. **AR-Disturbed:** Auf der AR-Baseline basierende (Pre-)Studien, in denen die materielle Umwelt der Teilnehmer nach verschiedenen Kriterien manipuliert wurde:
- a) Vertauschte Objekte (2 Dyaden): Pre-Studie, in der den Teilnehmern auf zwei oder drei materiellen Trägerwürfeln vertauschte Objekte angezeigt werden.
 - b) Partiiell sichtbare Objekte (2 Dyaden): Pre-Studie, in der jedem der Teilnehmer nur eine bestimmte Anzahl von augmentierten Objekten angezeigt wird, die beim jeweiligen Ko-Partizipanten nicht angezeigt werden.
 - c) Instabilität der Referenzobjekte [7 AR-Dyaden (6 F2F-Dyaden als Vergleichsfolie) im “Exhibition Desing” Setting]: Studie, in der die Teilnehmer in einer Interaktionsumgebung aus Stabilität (physisch manipulierbare Holzwürfel) und Instabilität (augmentierte Objekte) agieren. Das “Exhibition Desing” Setting bildet einen anfänglichen Prototyp im Entwicklungsprozess des gegenwärtigen „ARbInI Systems“. Die technisch bedingten Restriktionen konstituieren eine herausfordernde “ecology of sign systems” (e.g. Goodwin 2003a: 35ff.), da die semiotischen Umgebungsressourcen nicht kontinuierlich als kommunikative Ressourcen verfügbar sind. Die Interaktionsumgebung stellt damit besondere Anforderungen an die Koordinierungsaktivitäten der Interaktanten.
 - d) DeSync (in Planung): Auditive und visuelle Informationen werden den beiden Teilnehmern asynchron präsentiert.
- IV. **AR-Assistance:**
- a) Visuell augmentierter Aufmerksamkeitsfokus des Ko-Partizipanten (8 Dyaden): Studie, in der den Teilnehmern eine als Kompensation entwickelte Assistenz für die fehlende Ressource der Blickkoordination angeboten wird. Die Interaktanten erhalten visuelle Hinweise (rote Pfeile im HMD) auf den gegenwärtigen Aufmerksamkeitsfokus des jeweiligen Ko-Partizipanten.
 - b) Sonifikation der Objektbewegung (7 Dyaden): Analysen aus I. und II. haben gezeigt, dass Aktivitäten des jeweiligen Ko-Partizipanten aufgrund der Sichtfeldeinschränkung z.T. nicht wahrgenommen werden. In der Studie erhalten die Teilnehmer auditive Hinweise auf die physische Manipulation, wenn Objekte vom Ko-Partizipanten bewegt werden.

Studienauswahl

Aus dem umfangreichen Forschungsprogramm des C5-Projekts wurden für die zugrundeliegenden Fragestellungen dieser Arbeit die folgenden Studien I, II und III-c als empirische Datengrundlage ausgewählt:

	Obersee II		Exhibition Design
Bedingung	F2F-Baseline (Face-to-Face)	AR-Baseline (Augmented Reality)	AR-Disturbed (Augmented Reality)
Entstehungsdatum	04/2012	12/2011	08/2010
Dyaden	12	10	7
Teilnehmer	24	20	14
Stichprobe			
Altersspanne	19-62 Jahre		18-75 Jahre
Ø Alter	23.8	28.0	30.9
Geschlecht (w/m)	(10/14)	(12/8)	(5/9)
Beruf/Tätigkeit	24 Studierende	18 Studierende, 2 Sekretärinnen	k.A.
Versuchsangaben			
Ø Aushandlungsphase	24:20 min	19:37 min	10:25 min
Anzahl an Objekten	18	18	16
Ø ausgehandelte Objektplatzierungen	15,45	14,2	16 (festgelegt)

Der analytische Fokus ist vor allem auf die F2F- und AR-Baselinestudie gerichtet, da sie als vergleichbare Studien konzipiert wurden, deren grundlegendes Unterscheidungskriterium in der Einschränkung des “mutual monitoring” innerhalb der AR-Bedingung besteht. Die Studien können somit unter einer komparativen Analytik aufeinander bezogen werden, um den Einfluss des “mutual monitoring” auf die Verfahren zur Herstellung von Ko-Orientierung zu ergründen. Darüber hinaus ermöglichen die HMD-basierten Restriktionen im “mutual monitoring” der AR-Bedingung einen Lupeneffekt auf diejenigen Ressourcen zu erzeugen, die abseits reiner Blickkoordination (“gaze following”, Scaife & Bruner 1975) an der Herstellung von Ko-Orientierung beteiligt sind. Die Einschränkung des “mutual monitoring” impliziert die Frage nach der Robustheit von Interaktionssystemen: Welche interaktionalen Probleme sind in der Herstellung von Ko-Orientierung beobachtbar, wenn die zentrale Ressource der Blickkoordination nicht genutzt werden kann? Lassen sich multimodale Verfahren identifizieren, mit denen die Interaktanten ihre Orientierungsbemühungen stabil halten? Wie verändern sich diese Verfahren, wenn die Interaktionsumgebung der Interaktanten keine stabilen Referenzobjekte bereitstellt, auf die sie sich zuverlässig beziehen können? Im Kontext dieser Fragen verwenden wir dann auch die “AR-Disturbed”-Bedingung des Exhibition-Design-Szenarios als analytische Vergleichsfolie zur AR-Baselinestudie, um die interaktionale Relevanz stabiler und instabiler Umgebungsressourcen bei der Herstellung von Ko-Orientierung zu eruieren.

2 Einfluss des ‘Mutual Monitoring’ auf die Verfahren zur Herstellung von Ko-Orientierung

“Situations begin when mutual monitoring occurs, and lapse when the second-last person has left.”
(Goffman 1963: 18)

Die gegenseitigen Beobachtung von Interaktanten stellt nach Goffman das zentrale Konstitutionsmerkmal sozialer Situationen dar und bildet eine grundlegende Voraussetzung zur Herstellung von Ko-Orientierung in natürlichen face-to-face Interaktionen. Um gemeinsame Orientierung auf interaktional relevante Entitäten herstellen zu können, verwenden Gesprächsteilnehmer unterschiedliche semiotische Körper- und Kontextressourcen und konfigurieren sie – in feingradueller Abstimmung zum situativ emergenten Adressatenverhalten – in wahrnehmbare Äußerungs- und Orientierungsgestalten. Die aktive Aufmerksamkeitsmanipulation eines Ko-Partizipanten basiert auf der Fähigkeit, die eigene Wahrnehmung durch multimodale Ausdrucksressourcen wahrnehmbar zu „beschreiben“, um sie der Wahrnehmung des Gegenüber zuzuführen.

Auch wenn “mutual monitoring” als zentraler Mechanismus zur Koordinierung interaktiver Handlungsweisen bekannt ist, verweist die Forschungslage zum Phänomen auf epistemische Lücken, die vor allem in der methodischen Schwierigkeit begründet liegen, die Beobachtungen der Teilnehmer und ihre reziproke Beziehung zueinander analytisch zugänglich zu machen. Eine konzeptionelle Weiterentwicklung des Gegenstandes von einer axiomatischen Grundbedingung zur analytisch- ausdifferenzierten Rekonstruktion in unterschiedlichen Kontexten situativer Koordinierung ist daher schwierig. Die multimodal ausgerichtete Konversationsanalyse steht hier gewissermaßen vor dem entgegengesetzten Problem quantitativer Forschungsansätze, in denen der Phänomenkomplex “awareness” durch technische Verfahren des Eyetracking untersucht wird. Während Eyetracking-Methoden eine analytische Selektion auf das foveale Sehen bedingen und somit die Wahrnehmungsleistung des peripheren Sehens (und übriger Perzeptionsmöglichkeiten) ausblenden, geht die Konversationsanalyse im Sinne gestaltpsychologischer Prinzipien von einer holistischen Produktion und Rezeption multimodaler Äußerungsgestalten aus (vgl. “contextual configuration” bei Goodwin 2000a, 2003). Obgleich der holistische Ansatz eine konsequente und notwendige Erweiterung der logozentrischen Perspektive auf unimodale Phänomene der Verbalsprache darstellt, lässt er eine Problematik erkennen, die sich in der Übertragbarkeit des gestaltpsychologischen Paradigmas auf kontextuell- spezifische Einzelfälle manifestiert. Die kontextuelle Spezifik des Einzelfalls ist durch situativ variierende Ausprägungen endogener und exogener Aufmerksamkeitskonkurrenzen gekennzeichnet. Bei Luhmann heißt es:

„Trotz reflexiver Selbstregulierung bleiben Interaktionssysteme auf der Ebene des Wahrnehmens in hohem Maße störanfällig. Was der Wahrnehmung auffällt, hat möglicherweise soziale Relevanz, kann in die laufende Kommunikation einbrechen, sie stören, sie stoppen. Die Wahrnehmung des Wahrnehmens reicht nicht aus, um dies zu verhindern; sie sortiert nur die Ereignisse unter dem Kriterium, ob auch andere sie wahrnehmen (was ihnen gesteigerte Bedeutung gibt) oder nicht.“ (Luhmann 1984: 562)

Sowohl die theoretische Reflexion als auch empirische Beobachtung zur Interaktionscharakteristik situativer Aufmerksamkeitskonkurrenzen verweist auf die epistemische Notwendigkeit, das Postulat einer holistischen Produktion und insbesondere der Rezeption multimodaler Äußerungsgestalten als eigenständige empirische Frage zu behandeln.

Mit dem Konzept der “next turn proof procedure” (Sacks et al. 1974) steht der Konversationsanalyse zwar ein analytisches Rekonstruktionsverfahren zur Verfügung, das es erlaubt, die jeweiligen Teilnehmerinterpretationen vorausgehender Aktivitäten (“first turn”) aus den Anschlussreaktionen (“second turn”) zu ermitteln, aber hieraus lassen sich bei genauerer Betrachtung nur bedingt Erkenntnisse zur Assemblierungsqualität der Äußerungsgestalt sowie ihrer Wahrnehmungsspezifika gewinnen. Grund dafür ist die dem Forschungskontext geschuldete konzeptuelle Nähe zur Verbalsprache: Als teilnehmerrelevante Kategorie bezeichnet der “next turn proof” eine auf verbalsprachliche Redezüge bezogene Verständigungssicherung (z.B. Frage – Gegenfrage) und dient damit gleichermaßen, in seiner Funktion als analytische Kategorie, der semantischen Rekonstruktion von Teilnehmerinterpretationen. Idealtypisch ergibt sich daraus eine „Se-

quenz von dreizügigem Elementarformat“ (Schneider 2005: 64), in der sich – sofern keine Abstimmungsprobleme auftreten – in jedem dritten Redezug empirische Hinweise auf eine im ersten Redezug situierte Orientierungsgestalt manifestieren. Die Assemblierungsqualität und Wahrnehmungsspezifität einer Orientierungsgestalt kann im Zuge dessen nur binär im Sinne guter und schlechter Gestaltformen (bzw. im Sinne des Gestaltgebers verstandener oder nicht verstandener Gestaltformen) erfasst werden.

Betrachten wir hingegen das Interaktionsgeschehen aus einer multimodalen Perspektive, wird deutlich, dass der “second turn” nur einen geringen Teilbereich an Reaktionen abbildet und sich die Interaktanten fortwährend, jedoch in situativ abweichender Ausprägung, in sequentiell- und simultan-strukturierten Austauschprozessen von Orientierungshinweisen zur Bedeutungskonstitution ihrer Handlungsaktivitäten befinden. Reaktionen und vor allem „Mikro-Reaktionen“ (der Beginn einer Kopfdrehung, eine einsetzende Blickrichtungs- oder Posituränderungen etc.) manifestieren sich bereits während der Ausdrucksassemblierung einer Orientierungsgestalt oder im Falle antizipatorischer Inferenzprozesse zum Teil bereits davor. Wenn wir uns für die situative Emergenz qualitativer Unterschiede im “mutual monitoring“-Verhalten sich orientierender Akteure und ihrer Auswirkung auf die Verfahren zur Herstellung von Ko-Orientierung interessieren, greift die “next turn proof procedure” folglich zu kurz.

Aufgrund der phänomenologischen Konsistenz des Untersuchungsgegenstandes “mutual monitoring” und den sich daraus ableitenden methodischen Schwierigkeiten, wird im vorliegenden Kapitel eine mikroperspektivische Analytik vertreten, in der die reziproken Minimalreaktionen der Teilnehmer in ihrer sequentiellen und simultanen Verlaufsform rekonstruiert werden. Im Rahmen dieser „Atomisierungstendenz“ (Bergmann 1985: 316) liegt der Untersuchungsfokus auf einzelnen Orientierungsgestalten, die wir als morphologisch kleinste Einheit multimodaler Ausdrucksassemblierungen betrachten. Als Datengrundlage verwenden wir in diesem Kapitel vorrangig die face-to-face Bedingung, um sie als empirisch durchdrungene Vergleichsfolie auf die “mutual monitoring“-einschränkende AR-Bedingung anwenden zu können.

Das Kapitel ist in drei thematische Schwerpunkte aufgeteilt, die sich einem konzisen Überblick zum konversationsanalytischen Forschungsstand des Phänomens (2.1) anschließen. Der erste Teil (2.2) widmet sich der empirischen Beobachtung unterschiedlicher Assemblierungsqualitäten von Orientierungsgestalten und damit verbundener Anforderungsspezifitäten im “mutual monitoring” sich orientierender Interaktanten. Dabei stellen wir fest, dass Orientierungsgestalten in ihrer aufmerksamkeitslenkenden Funktion unterschiedliche Qualitäten besitzen, die sich auf einem Kontinuum zwischen aufmerksamkeitsbindend („Attraktorfunktion“) und aufmerksamkeitsvermittelnd („Mediatorfunktion“) einordnen lassen. Die Kategorien unterscheiden sich in ihrem Potential, die Summe an Fortsetzungsoptionen (situative Ambiguität), die interaktional begriffen ein Fortsetzungsproblem erhöhter koordinativer Anforderung und Aushandlungsrelevanz konstituieren, hinsichtlich einer faktischen interaktionalen Fortsetzung zu desambiguieren (situative Desambiguierung). Die Herstellung von Ko-Orientierung ist somit an den Erfolg situativer Desambiguierung gekoppelt. In diesem Sinne manifestiert sich Ko-Orientierung in einem interaktionalen Bestreben, situative Aufmerksamkeitskonkurrenzen abzubauen und eine lineare Orientierung auf eine spezifische Entität herzustellen.

Aus dieser empirischen Beobachtung leiten wir den zweiten thematischen Schwerpunkt des Kapitels ab und betrachten “mutual monitoring“-basierte Verfahren zur situativen Desambiguierung am Beispiel interaktionaler Fortsetzungsprobleme (2.3). Die Bearbeitung interaktionaler Fortsetzungsprobleme erweist sich als epistemisch besonders ergiebige Forschungsfeld, da hier erhöhte Koordinierungsanforderungen im “mutual monitoring” der Interaktanten zu beobachten sind. Aufgrund der interdependenten Abhängigkeit der Selektionsprinzipien „What to do next?“ (Garfinkel 1967) und „Where to look next?“ (vgl. Yarus 1967) manifestieren sich Unklarheiten der interaktionalen Fortsetzung oftmals auch in einem Orientierungsproblem aufmerksamkeitskonkurrierender Fortsetzungsoptionen, das von den Interaktanten bearbeitet werden muss. Vor diesem Hintergrund stellen wir den Untersuchungsgegenstand systematisch in den Kontext simultaner Handlungsaktivitäten, in denen sich Unklarheiten der interaktionalen Fortsetzung abzeichnen oder als faktisches Problem konkurrierender Handlungsoptionen existent sind.

Den theoretischen Hintergrund des Kapitels bildet der aktuelle konversationsanalytische Diskurs zum Forschungsgegenstand „Koordination“ (vgl. Deppermann & Schmitt 2007), aus dem wir offene Forschungsfragen ableiten und auf empirischer Basis einer umfangreichen Fallkollektion phänomenologisch unterschiedlicher Formen simultaner Aktivität adressieren (insgesamt 13 Fälle, aufgeteilt in „funktionale“ und „konfligierende“ Formen simultaner Handlungsaktivitäten). Das Forschungsinteresse gilt dabei einer-

seits den offenen Fragen zur organisationalen Relation der Strukturbildungsmechanismen ‚Sequentialität‘ und ‚Simultaneität‘: Bildet die beobachtbare Persistenz simultaner Aktivitäten auch aus einer emischen Perspektive eine teilnehmerrelevante Kategorie? Warum bereitet Interaktanten die etisch wahrnehmbare Permanenz und Gleichzeitigkeit visueller Informationen im Interaktionsgeschehen zumeist keine Probleme? In welchem strukturellen Verhältnis stehen ‚Sequentialität‘ und ‚Simultaneität‘ zueinander?

Andererseits – und dies steht unmittelbar im thematischen Zusammenhang zur organisationalen Relation der Strukturbildungsmechanismen – richtet sich unser Forschungsvorhaben auf eine Systematisierung kinesischer Simultanaktivitäten, die ein phänomenologisches Äquivalent zur verbalsprachlichen “overlap“-Klassifikation nach Schegloff (2000) bildet.

Die gewonnenen empirischen Erkenntnisse zu den Handlungspotentialen des “mutual monitoring” in der face-to-face-Bedingung werden abschließend, im dritten thematischen Schwerpunkt des Kapitels (2.4), analytisch mit Simultanaktivitäten in der AR-Bedingung kontrastiert. Dabei stellen wir fest, dass die auf permanentem “mutual monitoring” beruhende Fähigkeit, konfligierende Simultanaktivitäten in sequentielle Ordnungsstrukturen aufzulösen, in der AR Bedingung gestört ist. Simultanaktivitäten, in denen die Interaktanten aufgabendifferente Handlungen vollziehen, können in ihrer initialen Entstehung nicht erkannt und bearbeitet werden. Hier sei jedoch angemerkt, dass ein „Zusammenbrechen“ des Interaktionssystems, im Sinne einer sequentiellen Aufspaltung in individuelle Handlungsstränge, kein konstitutives Merkmal der AR-Bedingung im Oberseeszenario darstellt. Warum das Interaktionssystem stabil bleibt und die Interaktanten im Wesentlichen kooperativ der Bearbeitung ihrer Aufgaben nachgehen, bildet jedoch die forschungsleitende Fragestellung im nachfolgenden Kapitel (Kapitel 3).

2.1 Mutual Monitoring als Gegenstand der Interaktionsforschung

“Mutual monitoring” verweist auf eine Fähigkeit, die es Interaktanten in sozialer Interaktion ermöglicht, ihr jeweiliges Verhalten online zu beobachten, in situ zu interpretieren und durch feingraduelle Verfahren aufeinander abzustimmen (M. H. Goodwin 1980; Clark & Krych 2004).

“[...] speakers [...] monitor addressees for understanding and, when necessary, alter their utterances in progress. Addressees cooperate by displaying and signaling their understanding in progress.” (Clark & Krych 2004: 62)

In Bezug zur Effektivität des Phänomens haben Heath et al. (1995) am Beispiel kooperativer Arbeit in Kontrollräumen der Londoner U-Bahn gezeigt, dass auch “peripheral’ monitoring” eine funktionale und wirkungsmächtige Ressource zur Koordinierung gemeinsamer Handlungen bietet. Die Fähigkeit, Parallelhandlungen außerhalb der zentralen Aufmerksamkeitsorientierung verfolgen zu können, wird unter dem Begriff “peripheral awareness” als zentrales Schlüsselprinzip kooperativer Handlungen diskutiert (Knoblauch & Heath 1999: 176). Als Voraussetzung zur interaktiven Koordinierung kommunikativer Prozesse bildet “mutual monitoring” ein konstitutives Merkmal sozialer Interaktion:

“I would define a social situation as an environment of mutual monitoring possibilities, anywhere within which an individual will find himself accessible to the naked sense of all others who are ‘present’, and similarly find them accessible to him.” (Goffman 1964: 135)

Goffmans Definition sozialer Handlungen verdeutlicht zusätzlich den genuin multimodalen Charakter koordinativer Kommunikationsprozesse. Interaktanten verwenden unterschiedlichste semiotische Körper- und Umgebungsressourcen, die als holistische Einheiten (“contextual configurations” (Goodwin 2000a, 2003) konfiguriert werden und in ihrer multimodalen Gestalt deutliche Projektions- und Bedeutungsdivergenzen zu den Einzelkomponenten der Äußerungsgestalt aufweisen (Goodwin 2011). Um die Komposition semiotischer Ressourcen möglichst ihrer intendierten Bedeutung und Projektionsleistung entsprechend wahrnehmen und bearbeiten zu können, werden Äußerungsgestalten mit Orientierungsangeboten versehen, die Hinweise darauf liefern, wie die folgende Aktivität zu verstehen und zu behandeln ist („Kontextualisierungshinweise“: Auer 1986; 1992; Gumperz 1992; “Rendering activities visible”: Heath & Luff 1992a; Heath et al. 2002; „On-Line-Hilfen“: Dausendschön-Gay & Krafft 2002).

Orientierungsangebote in multimodalen Äußerungsgestalten sind dabei nicht im Sinne einer alleinigen „Orientierungshilfe“ der sich gerade äußernden Person aufzufassen, sondern werden in kooperativ-

dynamischen Prozessen hervorgebracht. Der Rezipient fungiert als aktiver „Mitwirkender“ (Dausendschön-Gay 2000), der mit Verfahren der Verständnissicherung anzeigt, ob Orientierungsangebote eine anschlussfähige Interpretation der Äußerungsgestalt ermöglichen. Im „Hilfe zur Selbsthilfe Prinzip“ kann er im Falle von Verständnisproblemen weitere Orientierungsangebote „einfordern“⁷. Während diese Verfahren anfänglich in der Literatur unter dem Begriff “back-channel-behavior” (Yngve 1970) als parallel zum Äußerungsbeitrag (“out of turn”) artikuliert Verfahren untersucht wurden, weisen spätere Konzeptionen eine Berücksichtigung ihrer Relevanz zur Bedeutungskonstitution der emergierenden Äußerungsgestalt im Sinne der Ko-Konstruktion auf (vgl. “continuers” bei Schegloff (1982) und “reactive tokens” bei Clancy et al. (1996)).

Im Rahmen der *workplace studies* zeigen verschiedene Studien, wie Interaktanten unter Bedingungen agieren, in denen das Potential des mutual monitoring gehemmt ist. Prominentes Beispiel sind Szenarien aus Operationssälen, bei denen die Akteure in den Möglichkeiten gegenseitiger Beobachtung einerseits durch eine strikte Aufgabenfokussierung, andererseits durch das Tragen von Mundschutz oder der Verwendung von Operationsmikroskopen etc.⁸ eingeschränkt sind (Heath et al. 2003; Nardi et al. 1995; Moreira 2004; D. Goodwin et al. 2005; Koschman et al. 2006; Schubert 2006; Svensson et al. 2007; Hindmarsh & Pilnick 2007; Mondada 2007a, 2011). In derartigen Szenarien wurden routinisierte Verfahren beobachtet, die sich auf verbalsprachlicher Ebene in der Formulierung prägnanter, atomarer Anweisungen äußern. Diese werden oftmals zur Auflösung formaler Machtverhältnisse in Frageform konstituiert (Bsp.: „könn’ wir Netzanschluss?“ als illokutionäre Aufforderung zum Anschluss eines Perfusors an die Netzspannung (vgl. Schubert 2006: 240).

Mit einem Fokus auf Multimodalität wurden organisatorische Prinzipien gefunden, die spezifische Anforderungen der Interaktionssituation adressieren: z.B. die körperliche Konfiguration eines “shared work-places”, die zeitliche Koordinierung von Verbalität und kinesischer Handlungsaktivität oder die deiktische Verwendung von Operationsbesteck (Mondada 2007a; Koschman et al. 2006). Die Routinehaftigkeit der empirisch auffindbaren Verfahren wird vor allem in den Forschungsergebnissen von Hindmarsh & Pilnick (2007) deutlich, die unter der Bezeichnung “intercorporeal knowing” Verfahren subsumieren, in denen eine auf Expertise (“highly contingent professional vision”⁹) beruhende Fähigkeit zur Antizipation nächster Handlungen sichtbar wird.

“So in contrast to many studies of coordination, even those concerned with improvisation, our study reveals how the teams featured in our data are intimately sensitive to delicate and subtle shifts in the embodied conduct of colleagues. This sensitivity, what we term *intercorporeal knowing*, underpins the team’s abilities to seamlessly coordinate emerging activities.” (Hindmarsh & Pilnick 2007: 1414)

Eine auf Routinisierung beruhende Projektion und Antizipation emergierender Handlungsaktivitäten stellt dabei besonders im Arbeitskontext invasiver Eingriffe eine zeitliche Optimierung relevanter Handlungsschritte dar. Damit liefert das Ergebnis im Bezug zur Temporalität eine auf natürlichen Daten basierende Bestätigung für Clarks “principle of least joint effort” (Clark 1996; Clark & Brennan 1991). Nach diesem Prinzip handeln Interaktanten opportunistisch (Clark & Krych 2004: 64): “they try to select from the available methods the ones they think take the least effort for the two of them jointly – the least cost in time, resources, errors etc. [...] If people are opportunistic, they should generally opt for the grounding methods that are most efficient” (Clark & Brennan 1991).

Im Forschungsfeld der *computer-supported cooperative work* (CSCW) haben eine Reihe von Studien zu mediatisierter Kommunikation und “collaborative virtual environments” (CVEs) gezeigt, dass die Möglich-

⁷ Auf Sprecherseite ist dieses Prinzip gleichermaßen gültig: Blickadressierung und/oder gestische Handlungsweisen können als funktionale Praktiken genutzt werden, um Hilfestellungen (z.B. bei Formulierungs- oder Erinnerungsschwierigkeiten) vom Rezipienten einzufordern (Bavelas & Chovil 2000; Goodwin 1986).

⁸ Heath et al. (2003: 91ff.) weisen auch auf die technologischen Entwicklungen hin (Visualisierung und Augmentierung von relevanten Informationen), die zunehmend als unterstützende und überwachende Maßnahmen bei chirurgischen Eingriffen genutzt werden: “Recent technologies in use within surgery involve innovations such as computer-assisted systems for image-guided surgery and intelligent warning systems for the monitoring of patient conditions and machine failures.”

⁹ Die Unterschied zwischen einer „alltäglichen“ und „professionalisierten“ Perzeption kommt bei Hindmarsh & Pilnick (2007: 1414) besonders im folgenden Zitat zum Ausdruck: “[...] within the anaesthetic room, participants do not notice the bodies of their colleagues in terms of an arm moving from left to right, or a head turning this way or that – rather they see the offer of help, the beginning of an intubation sequence or whatever. In this way the body remains ‘seen but unnoticed’ (Garfinkel 1967).”

keit der Anpassung multimodaler Verfahren unter stark restringierten Interaktionsbedingungen (z.B. “disembodiment in remote collaboration”) begrenzt ist (Luff et al. 2003; Heath & Luff 1991, 1992b; Hindmarsh et al. 1998). In solchen Szenarien, in denen die gegenseitige Beobachtung von Handlungsaktivitäten gravierend gestört ist (z.B. Sichtfeldeinschränkungen, zeitliche Verzögerungen der Signalübertragung, körperliche Entgrenzung etc. (vgl. Fraser et al. 2000)), lassen sich Asymmetrien der interpersonellen Kommunikation beobachten: Partizipanten interagieren in einer “fractured ecology” – einer Entkopplung zwischen situativer Handlung und relevantem Handlungskontext, in der körperliche Ressourcen weniger effektiv sind als in face-to-face Situationen (Luff et al. 2003; Heath & Luff 1992b).

“[...] participants are unable to design their own conduct in such a way that it is sensible and recognizable to a co-participant who has only limited access to the environment in which the action is produced. In this sense, conduct in fractured – fractured from the environment in which it is produced and from the environment in which it is received.” (Luff et al. 2003: 55)

Die meisten der im Forschungsfeld der CSCW studierten Systeme fokussieren technische Entwicklungen der Telekollaboration¹⁰ und unterscheiden sich somit grundlegend von unserem methodischen Ansatz, der die Möglichkeit ko-präsenter Interaktanten zur gemeinsamen physischen Manipulation ihrer Interaktionsumgebung erhält.

2.2 Mutual Monitoring als Grundbedingung der Face-to-Face Interaktion

In natürlichen “face engagements” (Goffman 1963: 83ff.) bildet mutual monitoring eine Grundbedingung zur Herstellung gemeinsamer Orientierung. Im Hinblick auf Goffmans Definition sozialer Interaktionssituationen als “environment of mutual monitoring possibilities” (Goffman 1964; vgl. 2.1) stellt sich die Frage, welche unterschiedlichen Handlungspotentiale des mutual monitoring in unseren Daten analytisch rekonstruiert werden können und ob mutual monitoring in eine funktionale und sinnvoll kategorisierbare Relation zur Bearbeitung interaktionaler Aufgaben gestellt werden kann.

Zum Einstieg in die Thematik differenzieren wir zwei in unseren face-to-face Daten analytisch rekonstruierbare Grundbedingungen des mutual monitoring: Die „attrahierende“ und „mediatisierende“ Aufmerksamkeitsbindung des Ko-Partizipanten am Beispiel gestischer Handlungen.

Aus heuristischen Gründen analysieren wir die Phänomene zuerst aus ihrer kontextuellen Emergenz losgelöst. Im Anschluss daran werden wir die multimodalen Ablaufformate der Präkontexte hinzuziehen und damit die konstituierenden Ursachen der Ausdifferenzierung analytisch nachholen. Die phänomenologische Differenzierung gründet methodisch auf einem semi-automatischen Verfahren, bei dem die Greifhand des orientierenden Akteurs und die Nasenspitze des Ko-Partizipanten als Indikator für eine sich verlagernde Orientierung getrackt¹¹ wurden. Auch wenn sagittale Richtungsänderungen der Nase in der Literatur als zuverlässiger Indikator der Aufmerksamkeitsorientierung beschrieben wurden¹² (vgl. Emery 2000: 582), stellt das Tracken der Nasenspitze lediglich eine Operationalisierung der visuellen Aufmerksamkeitsverlagerung dar. Aus diesem Grund erfolgte ein zusätzlicher Abgleich mit der Videoaufnahme der auf den Akteur gerichteten Kamera. Die hochauflösenden Kamerabilder ermöglichen eine fundierte Beurteilung, ob die Pupillenausrichtung des Akteurs isoliert oder in simultaner Verlagerung zur Kopfpositur erfolgt.

2.2.1 Phänomenologische Differenzierung: Attrahierende und mediatisierende Aufmerksamkeitsbindung

Abbildung 5 zeigt eine Sequenz von 840ms, in der VP-A seine rechte Hand aus ihrer “home position” (Sacks & Schegloff 2002: 133ff. ; Kendon 1980: 212: “rest position”) zu einem Objekt am Kartenrand führt. Dabei wurden semi-automatisch die Greifhand des orientierenden Akteurs und die Nasenspitze des

¹⁰ Ein konziser Überblick zu Telekollaborationssystemen, in denen Distorsionseffekte am Beispiel gestischer Projektion adressiert werden, findet sich bei Kirk et al. (2005): “Ways of the Hands”.

¹¹ Programm „Elasty“ (Version 2.0.2)

¹² “The whole head, in particular the orientation in which it is directed (using the nose, for example) is a sufficient indicator of attention direction (and therefore interest)” (Emery 2000: 582).

Ko-Partizipienten getrackt (*1a, die Fixpunkte sind mit einem roten Kreis markiert). Auch wenn die getrackten Einzelpunkte der Trajektorie im folgenden Beispiel mit den einzeln aufeinanderfolgenden frames korrespondieren (beispielsweise liegen zwischen f5 und f8 zwei getrackte Punkte), ist der Trackingalgorithmus nicht frame-gebunden, sondern dient der Visualisierung der Geschwindigkeit einer zurückgelegten Trajektorie durch Farbintensität und Punktabstand. Im Vergleich zum nachfolgend präsentierten Beispiel (Abbildung 5), ist die besagte Framekorrespondenz hier folglich der partikularen und relativ homogenen Ausführungsgeschwindigkeit der Greifhandlung geschuldet.

Betrachten wir die Zuordnung der getrackten Punkte im Trajektorienverlauf des orientierenden Akteurs (VP-A) und seines Ko-Partizipienten (VP-B) wird deutlich, dass bereits der Gestenonset (“preparation”, McNeill 1992) der ausführenden Greifaktivität in den ersten 200 ms die Orientierung von VP-B bindet (*1b). A’s zum Objekt geführte Greifhand fungiert als Attraktor, dessen Lokalisationsänderung durch eine simultan koordinierte Aufmerksamkeitsverlagerung des Ko-Partizipienten nachvollzogen wird. Infolgedessen markiert erst die stroke-Phase¹³ der Greifaktivität einen Zeitpunkt (frame 21), in dem beide Interaktanten auf das Objekt ko-orientiert sind.

F2F-Baseline:	VP-Gruppe 08 (VP-A: Naturschützer / VP-B: Investor)
Zeit:	24:55.52 - 24:56.36 (840ms / 40ms pro frame)

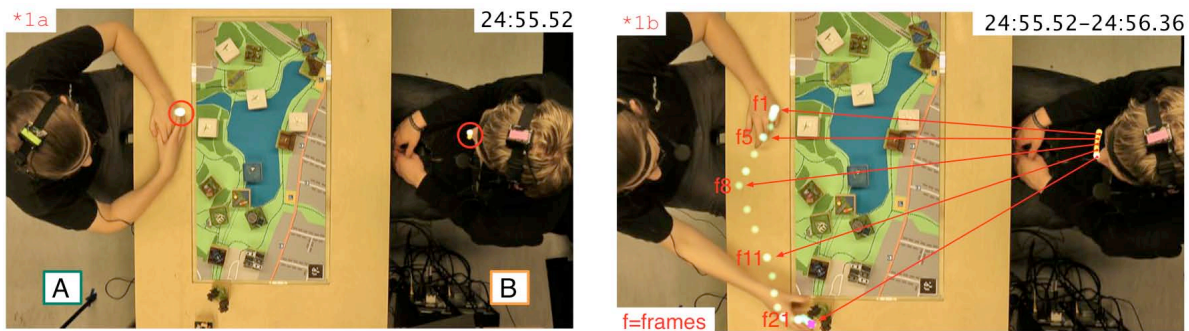


Abbildung 5: Attraktorfunktion der Gestenphase [preparation]

Kontrastierend zeigt Abbildung 6 ein Orientierungsformat, bei dem die zum Objekt geführte Greifhand nicht die Funktion fokussierter Aufmerksamkeitsbindung einnimmt. In der hier dargestellten Sequenz von 960 ms reagiert der Ko-Partizipient (VP-B) zwar auch auf die Onsetbewegung des orientierenden Akteurs (VP-A) mit dem Beginn der Aufmerksamkeitsverlagerung, nutzt die zum Greifen geführte Hand jedoch nicht als Attraktor, sondern als Mediator, der den Handlungsbeginn einer am Kartenrand antizipierbaren Aktivität projiziert. Infolgedessen sind beide Interaktanten bereits nach ~400 ms (frame 10) auf den Handlungsraum ko-orientiert, obwohl sich A’s gestische Handlungsaktivität währenddessen erst auf halber Strecke der Preparationsphase befindet. Aus der Ko-Orientierung auf den Handlungsraum erfolgt dann eine Orientierungsanpassung des Ko-Partizipienten, für den A’s konkrete Objektauswahl in der stroke-Phase der Greifaktivität erkennbar wird (*1b, frame 24).

¹³ Aus einer praxeologischen Perspektive definieren wir den “stroke” einer gestischen Handlung nach McNeill (1992) als bedeutungstragendes Element der Geste. Infolgedessen inkorporiert die stroke-Phase eine semiotische Signalisierungsfunktion. Komplementäre Ansätze der Gestenliteratur fokussieren die Kinetik einer Gestenhandlung und stellen den “peak of effort” (ebd.: S. 83) in das definitorische Zentrum einer stroke-Phase (s.a. Kendon 1980).

F2F-Baseline:	VP-Gruppe 04 (VP-A: Investor / VP-B: Naturschützer)
Zeit:	17:06.00 - 17:06.96 (960ms / 40ms pro frame)

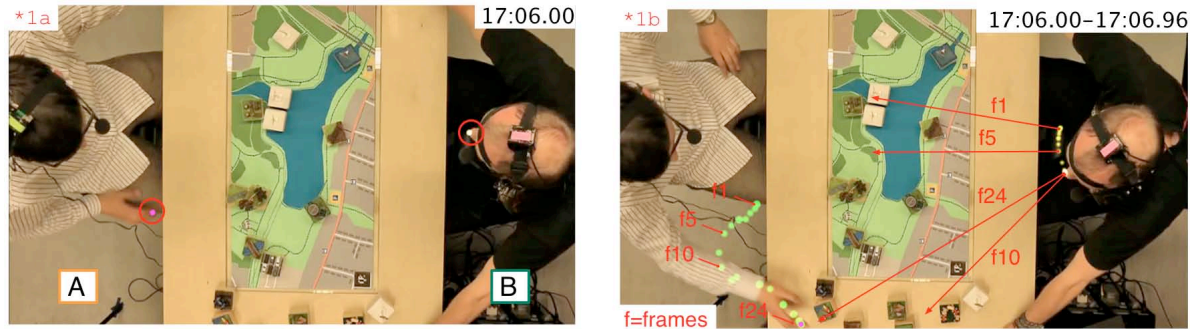


Abbildung 6: Mediatorfunktion der Gestenphase [preparation] (Handlungsantizipation)

2.2.2 Phänomenologische Ursachen: Multimodale Ablauformate der Präkontexte

Betrachten wir das multimodale Ablauformat des Präkontextes des in Abbildung 5 dargestellten Orientierungsformats, lassen sich die phänomenologischen Ursachen der gestischen Attraktorfunktion rekonstruieren. Noch vor der nachfolgend transkribierten Sequenz äußert VP-A ab 23:56 Bedenken bezüglich der von VP-B vorgeschlagenen Platzierung eines Hotels in unmittelbarer Nähe der Hauptstraße. Daran anschließend greift B das Argument der Lärmbelästigung auf und monologisiert 33 Sekunden (24:09-24:42) über etwaige Alternativplatzierungen. Die währenddessen geäußerte Skepsis seinerseits verdeutlicht, dass weder eine alternative Objektplatzierung noch der initial geäußerte Platzierungsvorschlag für ihn optimale Lösungen darstellen. B's „innerer Disput“ endet mit nachfolgendem Transkriptausschnitt:

FRAGMENT 1 (BEZUGNEHMEND AUF ABBILDUNG 5):

F2F-Baseline:	VP-Gruppe 08 (VP-A: Naturschützer / VP-B: Investor)
Zeit:	24:42 - 24:57

```

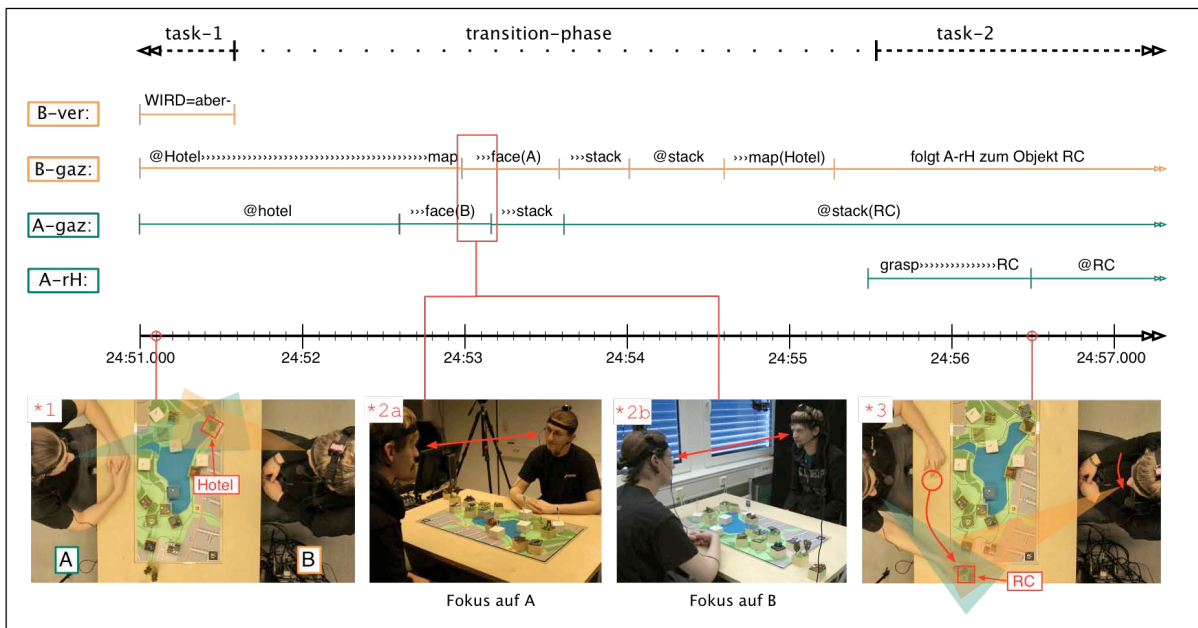
01 B-ver: NE: ich glaube äh=w(xxx)=|das das muss schon-|
                                                | (1.7) |

02 B-ver: wenn=wa vielleicht nen LÄRMschutzwall drum runziehen=oder=so

03 B-ver: wens wirklich SO schlimm WIRD aber-|
B-gaz: @hotel.....>>>>>>>>>>>>>>>>map>>>face (A)
A-gaz:                                     | >>>face (B)

```

B beendet seinen Redebeitrag mit einer abschließenden Ratifikation seines initialen Platzierungsvorschlags (01), der nach einer Pause von 1.7 Sekunden argumentativ mit der Möglichkeit einer „Lärmschutzwand“ expandiert wird (02+03). Auffällig ist hier, dass die turn-Expansion nicht die Formgebung einer kompletten Turnkonstruktionseinheit (TCU) aufweist, sondern durch den Diskursmarker „aber“, in seiner prosodisch markierten Kontinuität, eine potentielle turn-Fortführung ankündigt. Die Bearbeitung des Fortsetzungsproblems (grau markierter Transkriptausschnitt) ist im Folgenden dargestellt:



Der Diskursmarker „aber-“ signalisiert üblicherweise eine funktionale Relation zu einem sequentiell nachfolgenden Textelement (Millis et al. 1995), das hier jedoch suspendiert oder getilgt wird. Das multimodale Äußerungsformat lässt erkennen, dass B die Ko-Orientierung auf das Objekt „Hotel“ (*1) aufgibt und seinen Blick über die Karte führt („»map“). Diese kontextuelle Konfiguration wird von A als situative Ambiguität behandelt: Entweder hat B die Bearbeitung der Aufgabe „Hotelplatzierung“ abgeschlossen oder seine sich verlagernde Aufmerksamkeit ist als Such- oder Planungsaktivität der durch den Diskursmarker projizierten Fortsetzung zu interpretieren. Auf dieses Unklarheitsproblem reagiert A, indem er seinen Blick zum Ko-Partizipanten wendet (*2a), um das Informationsdefizit „what’s next?“ (Schegloff & Sacks 1973; „What to do next?“ (Garfinkel 1967: 12)) zu bearbeiten. Im Gegenzug reagiert B ebenfalls mit einer Blickrichtungsänderung zum Ko-Partizipanten A und stellt Blickkontakt her (*2b).

Die gegenseitige Blickwahrnehmung („mutual gaze“) geht dabei über die Signalisierungsfunktion komplementärer Aufmerksamkeitszuwendung (Argyle & Cook 1976, Kendon 1967, Goodwin 1981) hinaus, ohne die klassische Funktionalität der turn-Übergabe zu erfüllen (z.B. C. Goodwin 1980, Bavelas et al. 2002). Der Blickkontakt im mutual monitoring der Gesprächspartner ist eine Lösungsstrategie zur situativen Desambiguierung und konstituiert einen Typus von Handlungspotential, in dem nicht nur B legitimer Sprecher im Sinne einer projizierten turn-Fortsetzung wäre, sondern nun auch A den Status eines legitimen nächsten turn-Inhabers durch die Aufmerksamkeitsbindung des Ko-Partizipanten erlangt. In Goffman’s Terminologie hat er das Rollenverhältnis im „participation framework“ in ein „production format“ transformiert („change of footing“¹⁴, Goffman 1981: 128). In der Tat fährt A aus der Position des nächsten legitimen Sprechers fort und richtet seinen Blick auf die übrigen Objekte im „stack“. B’s Reaktion zeigt, dass er zwar der Orientierungsverlagerung A’s folgt, diese jedoch nicht kontinuierlich aufrechterhält und den Blick anschließend wiederum auf die anfängliche Aufgabenbearbeitung richtet. Erst die Onsetphase der Greifaktivität A’s bindet aufgrund der definiten Ankündigung einer neuen Aufgabenbearbeitung (task-2) die Aufmerksamkeit B’s im Sinne der in Abbildung 5 dargestellten Attraktorenfunktion und führt zur Ko-Orientierung auf das Objekt „RopesCourse“ (RC)(*3).

Zusammenfassend bildet die Unklarheit bezüglich der Aufgabenbearbeitung in der Transitionsphase zwischen „task-1“ und „task-2“ eine kontextuelle Bedingung, in der Handlungspotentiale zur geeigneten Fortführung der Interaktion durch mutual monitoring-Aktivitäten ausgelotet werden müssen. Das Resultat dieses feingraduellen Aushandlungsprozesses bedingt dabei letztendlich die in Abbildung 5 veranschaulichte Attraktorenfunktion der Greifhandlung A’s.

¹⁴ Bei Goffman definiert als: „change in the alignment we take up to ourselves and the others present as expressed in the way we manage the production and reception of an utterance“ (Goffman 1981: 128)

Betrachten wir nun kontrastierend das multimodale Ablaufformat des Präkontextes des in Abbildung 6 dargestellten Orientierungsverfahrens, wird deutlich, dass die Transitionsphase zwischen zwei interaktionalen Aufgaben die monitoring-Aktivität des gegenwärtigen Rezipienten der Orientierungshandlung beeinflusst. Im nachfolgenden Fragment 2 markiert die multimodale Äußerungsgestalt des orientierenden Akteurs eine ostentative Aufmerksamkeitsverlagerung zur Folgeaufgabe, so dass etwaige Fortsetzungsprobleme nicht entstehen, sondern durch instantane situative Desambiguierung verhindert werden.

FRAGMENT 2 (BEZUGNEHMEND AUF ABBILDUNG 6):

F2F-Baseline:	VP-Gruppe 04 (VP-A: Investor / VP-B: Naturschützer)
Zeit:	16:54 - 17:07

01 A-ver: |autoMAtisch DAdurch das die ZU:Gstrecke hier is °hh=äh:
 A-gaz: |@map(Zugstrecke)
 B-gaz: |@map(Zugstrecke)

02 A-ver: schätz ich mal NICH das sich HIER so viele vo=VÖgel einnisten;

03 A-ver: |die dann so schützenswert sind;| |also das:=mh:|
 B-ver: |n=joa:::=das glaub ich nicht; | (1.1) | | (1.2)

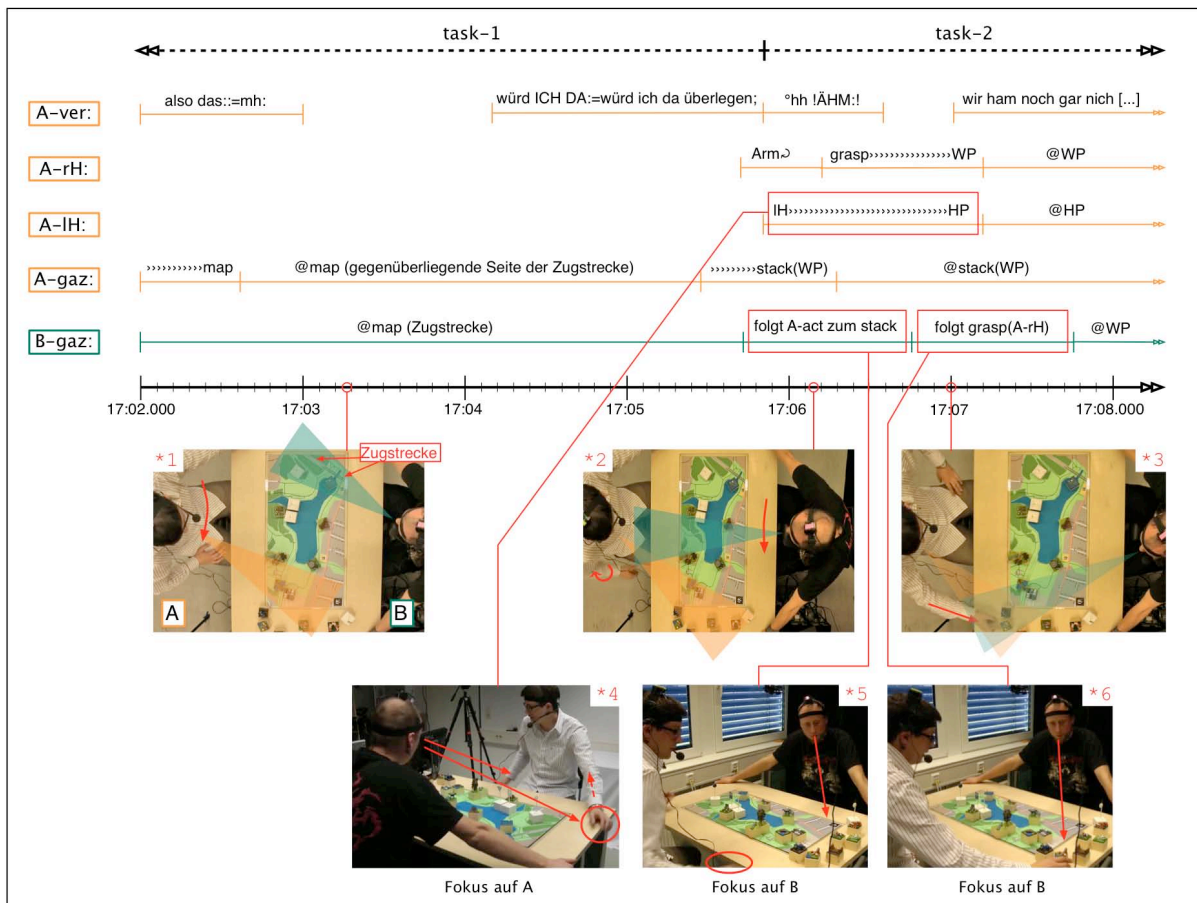
04 A-ver: würd ICH DA:=würd ich da überlegen; °hh !ÄHM:! (-) wir ham

05 A-ver: noch gar nicht an DIE:=also wir ham jetzt BISher an die

06 A-ver: freizeitbeschäftigung gedacht für (-) ich sag mal JUgendliche

07 A-ver: aufwärts=aber für kleine KINder haben wir noch gar nicht gedacht;

Der Beginn des Transkriptausschnitts zeigt eine von VP-A initiierte Beendigung der gegenwärtigen Aufgabe, in der die Partizipanten seit Minute 15:21 die Platzierung eines Vogelschutzgebietes aushandeln. Argumentativ verweist B dabei auf eine in der Nähe befindliche Zugstrecke, die ein Nisten schützenswerter Vögel verhindere (01-03). Das multimodale Ablaufformat (siehe nachfolgende Grafik) der Übergangsphase (grau markiert) zeigt, mit welchen ostentativ- multimodalen Verfahren VP-A eine Aufmerksamkeitsverlagerung des Ko-Partizipanten erzielt und eine Voraussetzung zur antizipatorischen Fokussierung des Handlungsraums schafft:



A verlagert seine Aufmerksamkeit aus der Ko-Orientierung auf die Zugstrecke zum gegenüberliegenden Kartenabschnitt (*1) und verbalisiert mit „also das::=mh: (1.1) würde ICH DA:=würde ich da überlegen;“ einen abschließenden Kommentar zur Aufgabenbearbeitung von task-1, dessen fallende prosodische Markierung eine geschlossene Turnkonstruktionseinheit (TCU) bildet. B's fortwährende Fokussierung der Zugstrecke weist darauf hin, dass A's Orientierungsverlagerung keine aufmerksamkeitsbindende Funktion ausübt. Erst die multimodale Gestalt aus i) "in-breath" + Hesitationssignal „!ÄHM:!“ als "pre-beginning" (Schegloff 1996a), ii) der Retraktion des linken Arms (*2) und iii) des in Handlungsbereitschaft geführten rechten Arms (*2+*3) bilden eine Orientierungsgestalt, auf die B mit einer Blickrichtungsänderung reagiert (*3).

In der sequentiellen Ordnung der Orientierungsgestalt signalisiert A's Retraktion der linkshändigen Handlungsbereitschaft, als "frame" der Aufgabenmarkierung „Vogelschutzgebiet“, eine Relevanzminderung von task-1 (*2+*6). A's rechte Hand spiegelt dieses Ablaufformat: Aus der "home position" (*1) markiert die Armdrehung zum "stack" (*2+*3) die Bereitschaft zum Handlungsvollzug in dem von ihm rechts gelegenen Handlungsraum. Vor diesem Hintergrund fungiert A's emergente Greifaktivität als Mediator, der den Handlungsbeginn einer am Kartenrand antizipierbaren Aktivität projiziert (*4). Erst sequentiell nachfolgend nutzt B die stroke-Phase zur Orientierungskorrektur auf den Referenten (*5), in Folge derer beide Interaktanten auf die neue Interaktionsaufgabe "task-2" ko-orientiert sind (*6).

2.2.3 Mediatorfunktion multimodaler Äußerungsgestalten auf spezifische Referenten

Antizipatorische Orientierungshandlungen, die auf einer Mediatorfunktion komplex- multimodaler Äußerungsgestalten beruhen, lassen sich in unseren Daten nicht nur für unspezifische Verweiskräume rekonstruieren (vgl. Fragment 2), sondern auch für konkrete, physikalisch abgrenzbare Referenten. Im nachfolgenden Fragment übt eine Kombination aus präkontextueller Handlung und ostentativ ausgeführter Äußerungsgestalt eine mediatisierende Orientierungsfunktion aus, die von der Ko-Partizipantin durch eine

Aufmerksamkeitsorientierung auf das intendierte Verweisobjekt bearbeitet wird, noch bevor die Sprecherin das Referenzobjekt verbalsprachlich identifiziert. Die Konfiguration beschreibt ein Potential des mutual monitoring, in dem die kontemporäre Äußerungsgestalt reflexiv sinnstiftend auf eine Aushandlungsepisode des Präkontextes bezogen wird. Mit Verweis auf Goodwin (2013) können Ereignisse der Interaktionsgeschichte das sinnhafte Substrat nachfolgender Handlungen bilden:

“Actions exist as rich, temporally unfolding process. Individual actions emerge from, and use, a consequential past shaped through chains of prior action, providing current participants with a dense, present environment, a rich now, containing many different kinds of resources that can be selectively decomposed, reused and transformed to build a next action, a proposal for how the future will be organized.” (Goodwin 2013: 21)

Unsere Analyse legt dabei im Besonderen dar, wie die präkontextuelle Einbettung einer Verweishandlung nicht nur als Fundament der Folgehandlung konstituiert wird, sondern auch das Orientierungsverhalten des Ko-Partizipanten aktiv beeinflusst.

FRAGMENT 3:

F2F-Baseline:	VP-Gruppe 12 (VP-A: Naturschützerin / VP-B: Investorin)
Zeit:	18:01 - 18:41

01 A-ver: |also ICH würde das hier auf !KEI!NEN fall hin bauen;
 A-act: |((stellt Objekt WS an den Kartenrand))

*1

*2



Topview
18:03.60



Topview
18:04.92

Der Beginn des Transkriptausschnitts zeigt eine vorschlagsabweisende Handlung, die im Zuge einer nachfolgenden Orientierungsaktivität als Präkontext eine reflexiv sinnstiftende Funktion einnimmt. Nachdem VP-A, in der Rolle der Naturschützerin, sich in einer längeren Argumentationsphase gegen den Bau einer „Wasserskianlage“ (WS) gewendet hat, weist sie VP-B’s Vorschlag mit „also ICH würde das hier auf !KEI!NEN fall hin bauen;“ zurück (01), greift nach dem Objekt „WS“ (*1) und platziert es am Kartenrand (*2). Die multimodale Äußerungsgestalt aus prosodisch markiertem Negationsformat („auf !KEI!NEN fall“) und physischer Änderung des Objektstatus (Lokalisationswechsel von einem „aktiven“ in einen „passiven“ Handlungsraum), führen zur Beendigung der kontemporären Interaktionsaufgabe.

(beide VPs orientieren auf die noch zu verhandelnden Objekte am Kartenrand)

18:13

02 B-ver: |um: mit der NATUR was mehr=(Lachen) in verbindung zu kommen
 B-act: |((greift das Objekt RC aus dem stack und führt es zur

*3

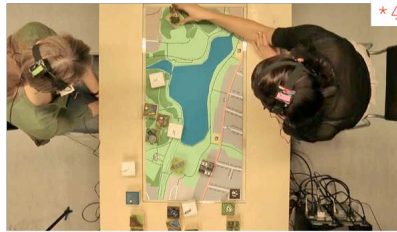
03 B-ver: würd ich IHNEN vorschlagen das wir HIER ein WÄLDchen von gegenüberliegenden Kartenseite).....

04 B-ver: ihnen nutzen um da einen KLETTERwald (-) HIN zu bauen- (-)
 B-act:

05 B-ver: WAS: die NATUR natürlich be!LÄ::SST! und nichts ABBAUT-
 B-act:
 06 B-ver: SO bekommen wir natURLiebhaber auch in die gegend- (-) DIE
 B-act: ((platziert RC))
*4



Topview
18:13.96

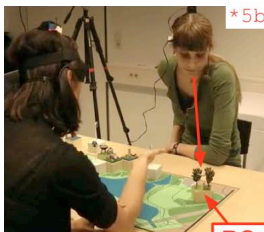


Topview
18:31.52

07 B-ver: |wahrscheinlich das natURschutzgebiet auch SEHR zu schätzen wissen-
 B-act: |dG>>>>>>>@NPA>>>HP.....
 A-gaz: |>>>NPA.....
*5a+b *6



Topview
18:35.28



Fokus auf A
18:35.28



Fokus auf A
18:35.64

Nach Beendigung der Interaktionsaufgabe „Wasserski“ (01), eröffnet B eine neue Interaktionsaufgabe. Ihre verbalsprachliche Vorschlagsankündigung „um: mit der NATUR was mehr=(Lachen) in verbindung zu kommen“ (02) fungiert dabei als Display einer nun nachfolgend weniger kontroversen Vorschlagshandlung (vgl. „pre-proposal-account“, Houtkoop-Steenstra 1990). Während der Initiierung des „pre-proposal-accounts“ greift sie das Objekt „Kletterpark“ (RC) aus der Gruppe der übrigen Objekte am Kartenrand (*3), führt es zur rechts von ihre liegenden Kartenseite und verbalisiert einen Vorschlag zur Platzierung (03-04).

Daran anschließend expandiert sie diesen Vorschlag (05+06) („post-proposal-account“, ebd.) und platziert das Objekt im Zuge dessen auf einem als Waldstück ausgewiesenen Areal (*4). Verbalsprachlich schließt B die Vorschlagshandlung durch eine argumentativ ausgeführte Verbindung zu einem bereits ausgehandelten „Naturschutzgebiet“, auf das sie mittels deiktischer Geste verweist (NPA) (06+07). Dabei folgt ihre Ko-Partizipantin A der deiktischen Verweishandlung (*5a), indem sie ihre Aufmerksamkeit von der vorgeschlagenen Objektplatzierung „RC“ (*5b) auf das argumentativ relationierte Objekt „NPA“ verlagert (*6). In Ko-Orientierung auf das Objekt „NPA“ markiert B den Abschluss der kontemporären Vorschlagskonstitution durch die Retraction ihrer Verweishand (*6).

08 B-ver: !TROTZ! ALLE!DEM! (-) BIN ich FÜR (.) WASSERSki;|
 B-act: frwd (A)>>>>>frwd>>>WS>>>HP.....|
 A-gaz:>>>B (face)>>WS..@WS.....|
*7 *8 *9a+b *10a+b *11



Aus der Konfiguration des terminierten, aber bislang nicht ratifizierten Vorschlagsangebots „NPA“, reorientiert B ihre Ko-Partizipantin A durch ein multimodal-indexikalisches Verweisformat auf das eingangs zurückgewiesene Objekt „Wasserski“ (WS). In der verbalsprachlichen Initiierung wird der adverbiale Konzessivkonkretor „!TROTZ! ALLE!DEM!“ in ein multimodales Äußerungsformat integriert, das auf pragmatischer Handlungsebene polyfunktional wirksam ist:

In der verbalsprachlichen Verwendung fungiert die Adverbialform einerseits konnektierend in Form eines Gliederungspartikels und verweist sowohl auf den (vorläufigen) Abschluss des Vorschlagsangebots „NPA“ (02-07) als auch auf den Beginn einer neuen Handlungsaktivität. Andererseits stellt die Adverbialform konzessiv einen anaphorischen Bezug her und deutet reflexiv auf etwas „Vorhergehendes [...], das mit dem Nachfolgenden nicht kompatibel ist“ (vgl. Fahrländer 2013: 3; s.a. Grimm’sche Wörterbuch 1889: 1114f.: „*ob armuth euer loos auch sei, hebt hoch die Stirn, trotz alle dem!*“). Vor diesem Hintergrund wird die Vorschlagshandlung „NPA“ retrospektiv mit einer Konnotation versehen, die sie als „Interimslösung“ zur interaktionalen Fortsetzung ausweist. Als parenthetischer Einschub dient sie der Überbrückung zum eigentlichen, in der sequentiellen Nachfolge stehenden Handlungsinteresse, dessen Relevanz von B mittels prosodischer Akzentuierung markiert wird („!TROTZ! ALLE!DEM!“).

Über die verbalsprachliche Verwendungsform hinaus ist der adverbiale Konzessivkonkretor an eine als Beat geschlagene Verweisform gekoppelt: B führt ihre Hand schlagartig – simultan zur akzentuierten Verbalisierung !TROTZ! – in Richtung des vorigen Vorschlagsangebots „NPA“, friert die Beatgeste im Moment des verbalisierten Lexems „!ALLE!DEM“ ein (*7) und retrahiert sie anschließend in ihre Ausgangsposition (*8). Auf pragmatischer Handlungsebene übt die ostentative Ausführung der multimodalen Äußerungsgestalt eine aufmerksamkeitssteuernde Funktion aus, auf die A durch eine Blickrichtungsänderung vom initialen Vorschlagsangebot „NPA“ (*7) auf ihre Ko-Partizipantin B reagiert (*8). Ihre Aufmerksamkeitszuwendung signalisiert die Erwartungsbereitschaft („what’s next?“) der von B durch Wortwahl und Prosodie relevant gesetzten Folgehandlung.

Aus der gemeinsamen „Wahrnehmungswahrnehmung“¹⁵ kennzeichnet B sequentiell nachfolgend die emergente Aktivität als Vorschlagshandlung („BIN ich FÜR“) und führt die Hand erneut in Richtung ihrer Ko-Partizipantin – diesmal in akuter Handlungsbereitschaft mit ausgestreckten Fingern zur indexikalischen Verweisfunktion (*9a+b). Die Verbalisierung der Vorschlagsankündigung, in Verbindung zur emergenten Verweishandlung, bedingen eine retrospektive Anpassung der Signalisierungsfunktion des Konzessivkonkretors: Auf pragmatischer Handlungsebene verweist die Adverbialform nicht mehr einzig

¹⁵ vgl. Luhmann (1984: 561f.): „Wenn Alter wahrnimmt, daß er wahrgenommen wird und das auch sein Wahrnehmen des Wahrgenommenwerdens wahrgenommen wird, muß er davon ausgehen, daß sein Verhalten als darauf eingestellt interpretiert wird; es wird dann, ob es ihm paßt oder nicht, als Kommunikation aufgefaßt, und das zwingt ihn fast unausweichlich dazu, es auch als Kommunikation zu kontrollieren.“

auf eine inkompatible Relation zum unmittelbaren Vorfeld – dem „Interimsvorschlags NPA“ –, sondern darüber hinaus auf einen reflexiven Sinnbezug zu einem Ereignis der Interaktionsgeschichte, das vor dem Interimsvorschlag „NPA“ liegt. Die Reflexivität des Projektionsformats „!TROTZ! !ALLE!DEM (-) BIN ich FÜR (.)“ konstituiert sich im Zuge dessen aus einem (impliziten) Handlungswissen, dass ein in Relation konstituierter Vorschlag, ohne vorbereitende Handlungen (vgl. „pre-proposal-account“, Houtkoop-Steenstra 1990) und bei Einhaltung grundlegender Kohärenzprinzipien, nicht auf ein unbekanntes Denotat verweisen sollte. Ein neues Referenzobjekt wäre folglich weder plausibel noch erwartbar, da die durch das Konzessivadverbial hergestellte Kohärenz ein interaktional etabliertes Wissen über das Referenzobjekt voraussetzt, d.h. der reflexive Sinngebungsprozess ist an die Reflexivität des Konzessivadverbials gekoppelt, obgleich sich beide Rückbezugformen auf unterschiedliche Denotate beziehen.

Für die nachfolgend beobachtbare antizipatorische Erschließung des Referenzobjekts ist dabei entscheidend, dass die Vorschlagseinheit „NPA“ durch die verbalisierte Einschränkung („!TROTZ! !ALLE!DEM“) als Bezugsrahmen des Referenten ausfällt. Ein Rückbezug auf den Interimsvorschlag bildete folglich ein Paradoxon („Trotz alledem bin ich für das Naturschutzgebiet“ (NPA)), das im Zuge wechselseitiger Unterstellung einer sinnhaften interaktionalen Fortsetzung nicht erwartbar ist (vgl. „Konversationsmaximen“, Grice 1975; „Sinnindikation“ und „Sinnoffenbarung“, Bergmann 1994). Infolgedessen ist der sprecherseitig intendierte Referent im Vorfeld der Vorschlagkonstitution „NPA“ zu suchen.

Aus der Konfiguration einer verdichteten Konvergenz präkontextueller Orientierungshinweise übt der nachfolgende deiktische Verweis B's nicht die Funktion fokussierter Aufmerksamkeitsbindung aus, sondern fungiert als Mediator auf das intendierte Referenzobjekt. In simultanem Vollzug zur Verweishandlung verlagert A ihre Aufmerksamkeit auf den Verweisraum, ohne B's deiktische Geste zu fokussieren (*10a+b). Vor diesem Hintergrund dient die Mediatorfunktion deiktischer Verweise primär nicht der Identifikation des Referenten, sondern konstituiert ein Display zeitlicher Orientierungsrelevanz, das Hinweise auf den Zeitpunkt liefert, wann eine Orientierungsverlagerung konditionell relevant wird.

Betrachten wir die Strokephase der Verweishandlung B's (*10a+b), zeigt sich, dass ihre distale Orientierung einen multireferenziellen Angebotscharakter aufweist, da sich gleich mehrere Objekte im vektoriellem Suchfeld der Zeigegeste befinden. In Anlehnung an Untersuchungen neuerer Deixisforschung (vgl. 3.1) bilden Zeigegesten keinen „punktgenauen Vektor [...], dessen imaginäre lineare Verlängerung automatisch zum Zeigeziel führt“, sondern konstituieren einen Suchraum, in dem das Zeigeziel vom Ko-Partizipanten erst aufgefunden werden muss (vgl. Stukenbrock 2012: 282: „Vektor-Ziel-Problem“; s.a. Goodwin 2000b: „domain of scrutiny“; Kranstedt et al. 2006: „pointing cone“; Rieser 2004: „region-pointing“). Die Besonderheit im vorliegenden Fall besteht darin, dass Such- und Verweisraum zeitlich zusammenfallen: VP-A ist nicht auf die verbalsprachliche Referenzierung im Suchraum angewiesen, sondern kann den Referenten instantan und reflexiv sinnhaft anhand der interaktionalen Vorgeschichte inferieren. Infolgedessen orientiert sie sich multimodal (Positur und Blick) auf den antizipierten Referenten (vgl. *10b vs. *11), noch bevor B das Objekt verbalsprachlich als „WASSERski;“ identifiziert (08).

In Abbildung 7 sind die drei relevanten Phasen der Orientierungshandlung in konziser Form zusammengefasst und in ihrer temporalen Abfolge illustriert.

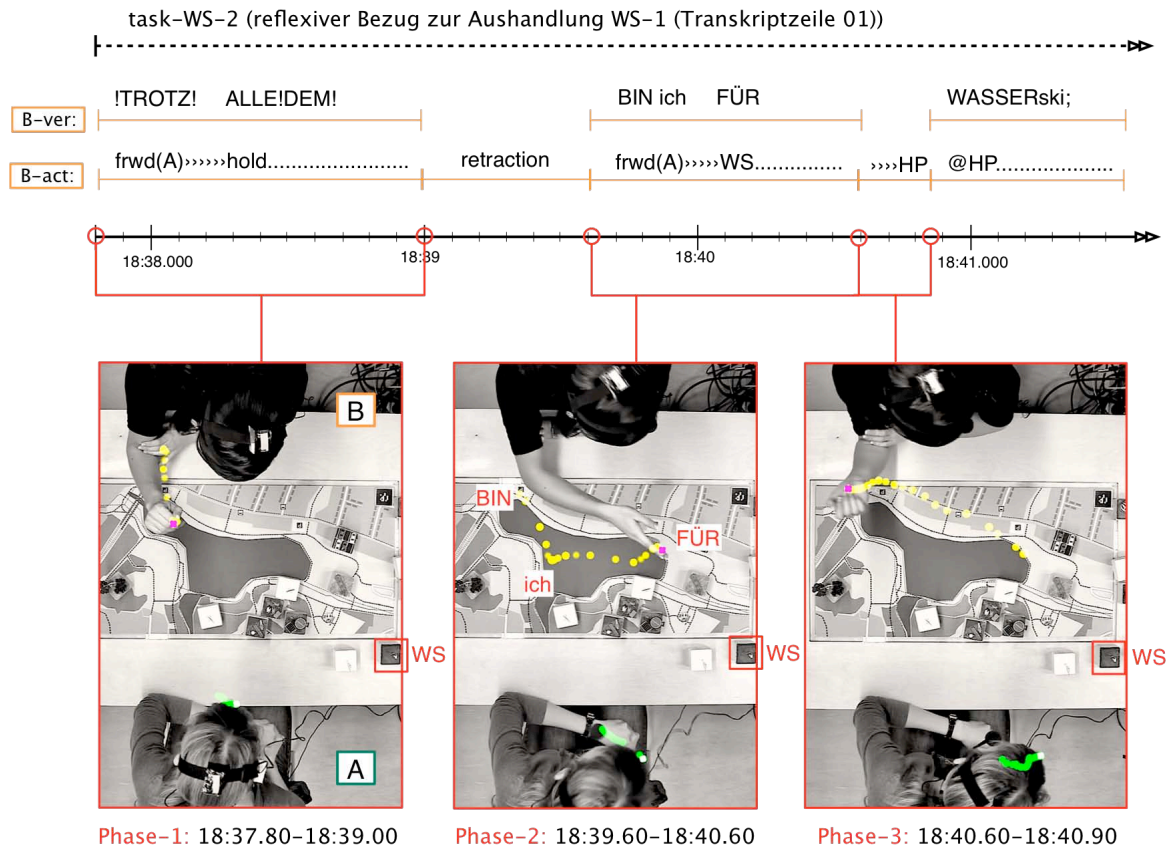


Abbildung 7: Mediatorfunktion auf einen spezifischen Referenten durch reflexiven Sinnbezug auf ein präkontextuelles Ereignis: Das Ablaufformat der antizipatorischen Orientierungshandlung lässt sich in drei Phasen skizzieren. Phase 1: Aufmerksamkeitsattraktion durch ein ostentativ-multimodales Äußerungsformat (akzentuierter Konzessivkonjektor + als Beat geschlagene Verweisform) – Phase 2: Projektion der Vorschlagskonstitution durch eine deiktische Verweishandlung in mediatisierender Funktion – Phase 3: Gestenretraktion und proximale Orientierungsverlagerung der Ko-Partizipantin auf den inferierten Referenten, bevor dieser verbalsprachlich desambiguiert wird.

2.2.4 Zusammenfassung

Die kontrastierende Analyse zwischen den Orientierungsformaten und ihrer Präkontexte (Abbildung 5 vs. Abbildung 6 und Abbildung 7) verweist auf die Notwendigkeit der Ausdifferenzierung von „mutual monitoring“-Aktivitäten als Verfahren zur situativen Desambiguierung, speziell in Transitionsphasen interaktionaler Aufgaben. Potentielle Fortsetzungsmöglichkeiten werden dabei hinsichtlich der emergenten faktischen Fortsetzung in „account-Verdichtungen“¹⁶ (Schmitt 2005: 49) desambiguiert. D.h., in Desambiguierungsphasen zur Fortsetzung ist eine Verdichtung sowohl intrapersoneller als auch interpersoneller Simultanaktivitäten beobachtbar, mittels derer die Interaktanten die Art der Fortführung aushandeln und koordinieren.

In Bezug darauf, liefert eine solche Ausdifferenzierung auch einen Beitrag zum klassisch-konversationsanalytischen Thema der „Projektion“ nächster Handlungen (z.B. Sacks, Schegloff & Jefferson 1974; Couper-Kuhlen 2004; Selting 2000; Goodwin 2002, 2006; Streeck 2005; Streeck & Jordan 2009): Antizipatorische Handlungen gründen einerseits auf ostentativen Äußerungsgestalten multimodaler Komplexität, die situativ desambiguierend, definite Rahmenbedingungen der Orientierung schaffen und den Umfang potentieller Gesprächsfortführungen eingrenzen (vgl. Goodwin 2000a; Hindmarsh & Pilnick 2007;

¹⁶ Schmitt (2005) bezieht sich hier auf die terminologische Verwendung des „account“-Begriffs von Garfinkel (1967): „Ethnomethodological studies analyze everyday activities as members’ methods for making those same activities visibly rational and reportable for all practical purposes, i.e. accountable.“

Dausendschön-Gay & Krafft 2009). Andererseits, wie unsere Beispiele zeigen, sind sie auf eine spezielle Form des monitoring angewiesen, in der partikuläre Ressourcen der holistischen Äußerungsgestalt als Mediator und nicht als Attraktor genutzt werden.

Unsere empirischen Ergebnisse unterscheiden sich von den analytisch rekonstruierten Verfahren zur Projektion und Antizipation nächster Handlungen in Arbeitskontexten professionellen Wissens (Operationssäle, Verkaufsstände, etc.). Während die Antizipation nächster Handlungen in professionalisierten Interaktionssituationen an routinisierte Arbeitsabfolgen gekoppelt ist (vgl. 2.1 “intercorporeal knowing”, Hindmarsh & Pilnick 2007; s.a. Dausendschön-Gay & Krafft 2009), deuten unsere Analysen auf Projektions- und Antizipationsmuster, die auf einer reinen Sprach- und Handlungskompetenz in der Laienkommunikation gründen. Ein weiterer Unterschied besteht dabei in der phänomenologischen Fokussierung: Dem turn-orientierten, syntaktisch-prosodisch motivierten Projektionsbegriff, stellen Dausendschön-Gay & Krafft (2009) das Konzept der “preparation” (“preparation work”) gegenüber und verweisen damit auf multimodale Projektionsverfahren, mit denen sich Interaktanten intra- und interpersonell auf nächste Aktivitäten orientieren. In Abgrenzung zur Projektion abgeschlossener Turnkonstruktionseinheiten (TCUs) rücken sie somit die multimodale Projektion von “larger projects” in den Forschungsfokus. Unsere Analysen hingegen verweisen primär nicht auf die Projektion nächster Handlungsaktivitäten¹⁷, sondern fokussieren multimodale Projektionsgestalten und antizipatorische Prozesse bei der Herstellung von Ko-Orientierung als Koordinierungsanforderung. Nach Deppermann & Schmitt (2007: 40f.) zielen Koordinierungsaktivitäten – im Gegensatz zur „Kooperation“ – „nicht auf die Herstellung eines gemeinsamen Produkts und haben in diesem Sinne auch keinen produktspezifischen Handlungscharakter“ (zur Kritik an der Generalisierung des Prinzips vgl. 2.3 – Beziehungskonstellation von Sequentialität und Simultaneität). Innerhalb dieser räumlichen Dimensionierung des Phänomenbereichs, der eine Radikalisierung des „Detailierungsogs“ darstellt (Bergmann 1985: 316; vgl. Deppermann & Schmitt 2007: 26), lassen sich zeitlich verdichtete Projektionsformate aufzeigen, die innerhalb weniger Sekunden antizipatorische Inferenzprozesse bei der Herstellung von Ko-Orientierung ermöglichen (vgl. Fragment 2+3).

Die phänomenologische Differenzierung zwischen attrahierender und mediatisierender Aufmerksamkeitsbindung in Transitionsphasen interaktionaler Aufgaben (vgl. Fragment 1+2) liefert im Zuge dessen den Erkenntnisgewinn, dass antizipatorische Handlungen – neben Verfahren des “intercorporeal knowing”¹⁸ (ebd.) – an die ostentative Ausführungsqualität einer Orientierungsgestalt und ihrer kontextuellen Einbettung gekoppelt sind.

Die Ausführungsqualität von Orientierungsgestalten ist auch bei Pitsch et al. (2014) mit einem Fokus auf antizipatorische Prozesse thematisiert worden. Am Beispiel einer “tutoring” -Aufgabe, in der Eltern ihren acht bis elf Monate alten Kindern zeigten, wie unterschiedlich große Becher ineinander gestapelt werden, konnten Pitsch et al. empirische Hinweise auf eine reziproke und sequentielle Anpassung zwischen den Stapelhandlungen der Eltern und den Blicken ihrer Kinder feststellen. Ihre Ergebnisse zeigen auch, dass einige Kinder die nachfolgende Stapelhandlung durch Blickfokussierung auf den nächst größeren Becher antizipieren. Je nach sequentieller Einbettung (“action-final anticipation” vs. “anticipation during sub-action” vs. “anticipation during a nesting pause”) deuten die Eltern das antizipatorische Blickverhalten des Kindes als Display des Verstehens oder mangelnde Aufmerksamkeit, die durch ostentative Trajektorien der Becher greifenden Hand korrigiert wird. In den empirischen Beobachtungen ihrer Studie lässt sich eine phänomenologische Unterscheidung zwischen “anticipating” and following¹⁹ erkennen.

Trotz der phänomenologischen Differenzierungsähnlichkeit (“anticipating” vs. “following” und “mediatisierende” vs. “attrahierende” Aufmerksamkeitsbindung) beschreiben unsere empirischen Ergebnisse die analytische Präzisierung eines anders gelagerten Desiderats. Während das “tutoring” als “interaccional

¹⁷ Als Ressource zur analytischen Rekonstruktion, inwiefern die Eröffnung einer neuen Interaktionsaufgabe als solche wahrgenommen und bearbeitet wurde, ist der “next turn proof” somit nicht Teil des fokussierten Phänomenbereichs.

¹⁸ Auch wenn Hindmarsh & Pilnick (2007) den Begriff des “intercorporeal knowing” für projektive und antizipatorische Aktivitäten in professionalisierten Handlungskontexten (Bsp.: im Operationssaal) reservieren, beschreibt ein interpersonelles Wissen über körperliche Handlungen, in denen nicht mehr die physische Aktivität als solche, sondern zusätzliche, inkorporierte Informationen wahrgenommen werden, auch ein grundlegendes Handlungsprinzip alltäglicher Interaktion. Wenn mir beispielsweise ein rohes Ei zugeworfen wird, ohne dass ich den Gegenstand in der Hand der werfenden Person erkenne, werde ich trotzdem aus der Wurftechnik/-bewegung Informationen auf die materielle Beschaffenheit des Gegenstandes erschließen können, da das Ei vermutlich nicht mit einem Schlagwurf oder einer Kugelstoßbewegung geworfen wird.

¹⁹ “the tutor’s hand motions [...] function as orienting devices to attract and guide the infant’s attention” (Pitsch et al. 2014: 35)

learning system” (ebd.: 2, bezugnehmend auf Bruner 1985) ein reziprok adaptives Verfahren²⁰ der Ko-Konstruktion darstellt, lassen sich bei uns keine empirischen Hinweise auf adaptive Verfahren zwischen ostentativen Orientierungsgestalten und antizipatorischen Inferenzprozessen bei der Herstellung von Ko-Orientierung finden. Einerseits spielt hier wiederum die zeitliche Dimensionierung²¹ der Aktivitäten eine Rolle (die Projektion von Ko-Orientierung vs. die Projektion von “larger projects”, s.o.). Andererseits begründet sich der Unterschied auch aus den spezifischen Rezipientenstatus und damit einhergehenden Kompetenzzuschreibungen: Während im “motionese” die Möglichkeit des „Nicht-Verstehens“ unterstellt wird (folglich adaptive Korrekturen des Tutors in der Interaktionsaufgabe zu einem gewissen Grad prädisponiert sind), gehen die Interaktanten im Oberseesetting davon aus, dass die Konstitution einer gemeinsamen Wahrnehmungswahrnehmung eine ausreichende Pre-Bedingung zur Herstellung von Ko-Orientierung darstellt. Das Verstehen einer multimodal orientierenden Äußerungsgestalt wird somit unterstellt und vorausgesetzt, da “mutual monitoring” eine konsistente Grundbedingung der spezifischen face-to-face Situation zwischen kompetenten Sprachmitgliedern bildet. Vor dem Hintergrund der Abhängigkeit zur kognitiven Kindesentwicklung im “tutoring”, unterscheiden sich die empirischen Vorgehen: Während Pitsch et al. die Frage adressieren, wie sich antizipatorisches Handeln (hier als Lerneffekt einer routinisierten Handlungsabfolge) auf die ostentative Ausführungsqualität der Orientierungshandlung des Tutors auswirkt, verfolgen wir den umgekehrten Weg und liefern empirische Hinweise, wie eine Orientierungshandlung kontextualisiert und multimodal gestaltet sein muss, um antizipatorisches Handeln zu ermöglichen. Die analytisch rekonstruierten Konstitutionsformen bedingen dabei die sprecherseitige Terminologie der „attrahierenden“ und „mediatisierenden“ Aufmerksamkeitsbindung.

Inwieweit eine projektive Gestalt antizipatorisches Handeln ermöglicht, bemisst sich – wie unsere Beispiele verdeutlichen – im Wesentlichen daran, ob die Orientierungsgestalt das Potential aufweist, situative Ambiguitäten instantan und ostentativ zu desambiguieren – also für den jeweiligen Ko-Partizipanten in situ interpretierbar zu machen (vgl. Fragment 1-3). Die situative Desambiguierung bildet damit eine notwendige Prämisse, die es dem jeweiligen Ko-Partizipanten erlaubt, kontextuell fundierte Hypothesen an die interaktionale Fortsetzung zu stellen. Dabei konnte gezeigt werden, dass Desambiguierungsverfahren zum Teil erst einen projektiven Charakter durch reflexive Sinnbezüge auf präkontextuelle Ereignisse entfalten (vgl. Fragment 3).

2.3 Mutual Monitoring-basierte Verfahren zur situativen Desambiguierung

“Occurences of more than one speaker at a time are common, but brief.”
(Sacks, Schegloff & Jefferson 1974: 700)

Wenn wir mutual monitoring-Aktivitäten in Bezug zu Lösungsstrategien von Fortsetzungsproblemen in den analytischen Fokus rücken, wird deutlich, z.B. in Transitionsphasen interaktionaler Aufgaben (vgl. Abschnitt 2.2), dass die verbalsprachliche Analyseeinheit der Sequentialität kein ausreichendes Beschreibungsinstrument für kinesische Handlungsaktivitäten bietet. Auch nonverbale Simultanaktivitäten der Interaktanten sind häufig, aber im Vergleich zu der verbalsprachlichen “overlap”-Konzeption von Sacks, Schegloff und Jefferson (1974) nicht unbedingt kurzzeitig (“brief”). Simultaneität stellt vielmehr eine konsistente Grundbedingung multimodaler Kommunikation dar (vgl. Deppermann & Schmitt 2007).

2.3.1 Definition von Simultaneität

Da die Verwendung der Begriffe zu temporalen Aspekten der Kommunikation in der Forschungsliteratur sehr unterschiedlich gehandhabt wird, verweist Stukenbrock (2012: 28) auf die „erforderliche theoretische Unterscheidung zwischen zufällig zustande kommenden, kontingenten zeitlichen Verhältnissen und zeitlich konstituierten Relationen, die nicht zufällig sind und denen eine Funktion zugeschrieben werden kann.“ Daraus folgernd spezifiziert sie zwei Zeitlichkeitsbeziehungen, die für unsere Analysen zentral sind:

²⁰ “Action modification and the recipient’s gaze can be seen to have a reciprocal sequential relationship and to constitute a constant loop of mutual adjustments.” (Pitsch et al. 2013: 1)

²¹ Deppermann und Schmitt (2007: 38) sprechen von „unterschiedlichen Auflösungs-niveaus“ koordinativer Prozesse.

„Gegenüber dem weiteren Begriff der *Ko-Temporalität* soll der Simultaneitätsbegriff der Gleichzeitigkeit im engen, mikroperspektivischen Sinn vorbehalten bleiben. Er wird ausschließlich zur Bezeichnung multimodaler Koordinierungsprozesse verwendet, die sich intra- und interpersonell im selben Augenblick ereignen. Auf mikroanalytischer Betrachtungsebene werden mit dem Begriff der *Teilsimultaneität* dementsprechend Momente erfasst, in denen sich bestimmte Ausdrucksformen für einen gewissen Zeitraum in ihrer Emergenz zeitlich überschneiden und in bedeutungsvoller, funktional relevanter Weise miteinander einher gehen, sich dabei jedoch hinsichtlich Anfangs- und/oder Endpunkt und Dauer voneinander unterscheiden.“ (Stukenbrock 2012: 31)

Dieser phänomenologischen Differenzierung schließen wir uns weitestgehend an und verstehen das Konzept der „Simultaneität“ als zeitlich konstituierte Relation. Eine Differenzierung hinsichtlich der Funktionalität zeitlich konvergierender Aktivitäten ist jedoch problematisch. Zum einen ist nicht klar, worauf sich die Funktionalität (im Sinne der etymologischen Herkunft des Wortes ‚Funktion‘, als Aufgabe die innerhalb eines Ganzen in bestimmter Weise wirksam ist (vgl. DWDS²²)) im engeren Sinne bezieht: Ist zeitliche Konvergenz für intrapersonelle, interpersonelle oder beide Koordinierungsprozesse funktional? Oder bezieht sich die Funktionalität auf Potentiale interaktionaler Aufgabebearbeitung /-fortsetzung?

Zum anderen legen unsere Analysen nahe, dass die Funktionalität zeitlich konstituierter Relationen einerseits von den kontextspezifischen Bedingungen der Situation abhängt, andererseits nicht per se in der Simultaneität der Aktivitäten liegt, sondern von den Interaktanten selbst hergestellt und wechselseitig angezeigt wird. Diese Argumentation stützend, verweisen Fragmente unserer face-to-face Bedingung auf einen Typus „konfligierender Simultaneität“ (vgl. 2.3.7), der sowohl zufällig als auch nicht zufällig emergieren kann. Für die kausale Emergenz lassen sich Bedingungen des Präkontextes spezifizieren, die simultane Handlungsaktivitäten der Teilnehmer begünstigen. Unabhängig von der Art der Emergenz (zufällig vs. kausal) kann dieser Simultaneitätstypus sowohl funktional wie auch dysfunktional sein. Während dysfunktionale Simultaneität sich in einer Störung der kontemporären Aufgabebearbeitung abzeichnet, ist der konfligierende Simultaneitätstypus funktional, wenn die sich störend konvergierenden Teilnehmeraktivitäten ein Ablaufformat nächster Handlungsschritte generieren. In diesen Fällen suspendiert einer der Teilnehmer seine gegenwärtig ausgeführte Simultanaktivität, wartet den Handlungsvollzug des Ko-Partizipanten ab und kehrt dann zur zeitweilig suspendierten Aktivität zurück. Die suspendierte Handlungsaktivität fungiert dabei als Display einer „in-petto-Konfiguration“.

Die Zuschreibung von Funktionalität auf Ebene intrapersoneller Simultaneität ist hingegen weniger strittig, da sich verschiedene Ausdruckressourcen in multimodalen Gestalten wechselseitig in ihrer Signalisierungsfunktion beeinflussen und im Zusammenwirken Bedeutungsdivergenzen zu ihren Teilkomponenten aufweisen (vgl. Goodwin 2000a, 2003, 2011). Daraus folgernd kann intrapersoneller Simultaneität eine Funktion für semiotische Signalisierungsprozesse zugewiesen werden.

Aber auch hier ist der Verweis auf ein funktionales Zusammenwirken einzelner Ressourcen nicht ganz unproblematisch. In unserer face-to-face Bedingung lassen sich Episoden identifizieren, in denen multimodalen Äußerungsgestalt von den Interaktanten unterschiedlich gedeutet werden, da der Ko-Partizipant die „funktionale Relation“ der Einzelkomponenten nicht holistisch erfasst und interpretiert (vgl. 2.3.7.1). Für dyspräferierte Anschlussreaktionen, die mit der gegenwärtig interaktionalen Bearbeitung konfligieren, stellt sich dann die Frage, ob die einzelnen Ausdrucksressourcen wirklich funktional konnektiert waren? Analytisch ist das nicht entscheidbar. Ergebnisse aus den AR-Studien zeigen in diesem Zusammenhang, dass die funktionale Relation der Komponenten nicht immer eine obligate Prämisse zum Verständnis des Handlungsvollzugs bildet. Interaktanten sind unter bestimmten kontextuellen Bedingungen in der Lage, teilrezipierte Äußerungsgestalten zu rekonstruieren, ihren semantischen Gehalt zu inferieren und die erzeugte Präferenzstruktur sinngemäß zu bearbeiten (vgl. 3.2).

Aus den geschilderten Gründen reservieren wir den Begriff der „Simultaneität“ für multimodal konstituierte, intra- und interpersonelle Koordinierungsprozesse, die zeitlich begrenzt und gleichzeitig stattfinden (s.a. Stukenbrock 2012: 31). Ihre Funktionalität oder Dysfunktionalität rekonstruieren wir jedoch in situ, d.h. anhand der Teilnehmerinterpretationen der Interaktanten. Aus dieser Verfahrensart resultiert die empirisch begründbare Differenzierung der phänomenologisch unterschiedlichen Simultaneitätstypen „konfligierende und funktionale Simultaneität“ (vgl. 2.3.6).

²² Digitales Wörterbuch der deutschen Sprache (Brandenburgerische Akademie der Wissenschaften): <http://www.dwds.de/>

2.3.2 Multimodale Verdichtungsräume

Das Konzept „multimodaler Verdichtungsräume“ (Stukenbrock 2008, 2012) stellt in unseren Daten ein empirisch beobachtbares Phänomen dar und liefert darüber hinaus als Beschreibungskategorie einen analytisch gewinnbringenden Ansatz zur Rekonstruktion intrapersoneller Simultaneität.

„Multimodale Verdichtungsräume entstehen dadurch, dass unterschiedliche, für den kommunikativen Akt konstitutive Ausdrucksressourcen in einer Phase der Simultaneität zusammenlaufen, in verdichteter Weise gleichzeitig „präsent“ sind, bevor sie wieder in unterschiedlichen, weniger eng aufeinander bezogenen funktionalen und temporalen Strukturen auseinander laufen.“ (Stukenbrock 2012: 31)

In Anlehnung an die in 2.3.1 geschilderte Ausführung zur Funktionalität verweisen wir darauf, dass die für den kommunikativen Akt konstitutiven Ausdrucksressourcen nicht per se in aufeinander bezogenen funktionalen Strukturen relationiert sein müssen. Entscheidend ist, ob der jeweilige Ko-Partizipant die Komposition der Ausdrucksressourcen in einer Form deuten kann, die es ihm ermöglicht, erzeugte Relevanzstrukturen zu bearbeiten oder ob der gegenwärtige Sprecher eine dyspräferierte Reaktion akzeptieren und sie als funktionale Komponente in den emergenten Handlungsvollzug integrieren kann.

Abbildung 8 zeigt die Konstitution eines multimodalen Verdichtungsraums am Beispiel einer deiktischen Zeigehandlung nach Stukenbrock (vgl. 2012: 31f.). In der abgebildeten Situation schildert ein Patient mit chronischen Schmerzen den anwesenden Ärzten seine Probleme. Vor Beginn der Schmerzschilderung blickt er zu einem der Ärzte (→@doc), reorientiert sich unmittelbar vor der Verbaläußerung nach unten auf den Beinbereich (---) und startet nachfolgend eine gestische Aktivität. Die Preparationsphase der deiktischen Verweishandlung setzt kurz vor Beginn der verbalsprachlichen Schmerzschilderung ein (~~~). Im Augenblick der Artikulation des prosodisch markierten Demonstrativums „DES“ (süddt. für „das/dies“), führt der Patient einen doppelten Stroke auf sein linkes Knie aus (**). Stukenbrock beschreibt dieses Format als zeitliche Struktur, in der „sich die in der Emergenz immer stärker konvergierenden, auf dem Gipfelpunkt zusammentreffenden Ressourcen“ verdichten.

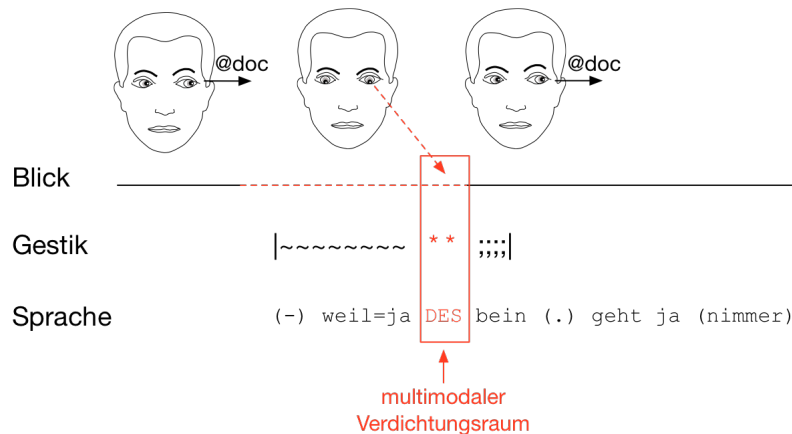


Abbildung 8: Multimodaler Verdichtungsraum nach Stukenbrock (2012: 32)

Neben der reduktionistischen Perspektive auf multimodale Verdichtungsräume als Äußerungsgestalten intrapersoneller Simultaneität stellen wir sowohl für das vorliegende Verdichtungsformat als auch für unsere analytischen Rekonstruktionen des Phänomens (2.3.6, 2.3.7, 2.3.8) eine interaktiv relevante Funktion fest: Multimodale Verdichtungsräume stehen oftmals im Kontext situativer Desambiguierung.

Im vorliegenden Beispiel desambiguiert der doppelte Stroke, um welches Knie es sich handelt. Intuitiv erscheint daher die funktionale Relation der einzelnen Ausdrucksressourcen innerhalb des geschlossenen Äußerungsformats inhärent logisch. Stellen wir uns jedoch eine präkontextuelle Konfiguration vor, in der einer der Ärzte bereits das schadhafte Knie kurz zuvor diagnostiziert hat, ändert sich das gesamte funktionale Zusammenwirken der einzelnen Ausdrucksressourcen in einen Akt intrapersoneller Orientierung oder präventiver Verständnissicherung (weitere Handlungsvollzüge sind denkbar). Das heißt konkret, dass sich die funktionale Relation der Äußerungsgestalt primär nicht nach ihrer inhärenten Logik bemisst, son-

dern anhand der kontextuellen Einbettung und interaktiven Wirksamkeit. Besonders relevant wird diese Praxis, wenn die Äußerungsgestaltung der multimodalen Verdichtung kein sprachlich fossilisiertes Produktions- und Interpretationsformat des referentiellen Wissens darstellt, das sich durch ein exaktes Zusammenfallen der einzelnen Ausdrucksressourcen konstituiert, sondern eine Gestaltform annimmt, in der die einzelnen Ausdrucksressourcen zeitlich auseinanderfallen:

„Im Hinblick auf Temporalitätsstrukturen und multimodale Verdichtungsräume bei Zeigehandlungen ist allerdings festzuhalten, dass unterschiedliche Ausdrucksressourcen wie z.B. ein deiktischer Ausdruck und eine dazugehörige Zeigegeste in einer emergierenden Zeigehandlung nicht notwendigerweise zeitlich exakt zusammenfallen müssen.“ (Stukenbrock 2012: 32)

Fälle multimodaler Verdichtung, in denen die Ausdrucksressourcen zeitlich dispersiv, aber trotzdem im Sinne der Einheitenbildung für die Äußerungsgestalt bedeutungskonstitutiv sind, stellen erhöhte Aufmerksamkeits- und Interpretationsanforderung an den Ko-Partizipanten (s. 2.3.7.1). Der multimodale Verdichtungsraum ist dann weniger gut erschließbar als in zeitlichen Verdichtungsformen komprimierter Ressourcenkonvergenz und infolgedessen auf ein präzises Sprecherdisplay angewiesen, das den relationalen Zusammenhang der Einzelkomponenten „accountable“ macht.

2.3.3 ‚Koordination‘ und die Beziehungskonstellation von Sequentialität und Simultaneität

Unter einer multimodalen Perspektive auf die „faktische Komplexität interaktiver Realität“ (Schmitt 2005: 53), in der simultane Aktivitäten eine persistente Konstitutionsbedingung des Interaktionsgeschehens bilden, stellt sich die Frage nach der gesprächsorganisatorischen Relation der Strukturbildungsmechanismen „Sequentialität“ und „Simultaneität“. Mit der – insbesondere für die deutsche Gesprächsforschung²³ – theoretisch wie methodisch und methodologisch relevanten Konzeptualisierung des Forschungsgegenstandes „Koordination“ (Schmitt (Hrsg.) 2007; erste Ansätze in Schmitt 2005), formulieren Deppermann und Schmitt ein Programm, das die beiden historischen Wurzeln der multimodalen Interaktionsforschung „interaction order“ (Goffman 1983a) und die „strenge empirische Methode“ der von Sacks und Schegloff entwickelten Konversationsanalyse zusammenführt²⁴ (ebd.: 16f.) und eröffneten zugleich eine analytische Perspektive auf die Beziehungskonstellation von Sequentialität und Simultaneität.

2.3.3.1 ‚Koordination‘ als Forschungsgegenstand

Die Begründung des „neuen Forschungsgegenstandes“ (s. Fußnotenankündigung (Holly 2008)) ‚Koordination‘ wird von Deppermann und Schmitt (2007: 16) mit der Feststellung motiviert, dass Videoaufzeichnungen immer häufiger Tonaufnahmen als empirische Basis der Untersuchung interaktionaler Phänomene ersetzen. Die „visuelle Revolution“ (Schmitt 2004: 56) mache es erforderlich, auf die durch audiovisuelle Medien sichtbar gemachte Komplexität, in der „eine Vielzahl zumeist gleichzeitiger Aktivitäten ablaufen“ (Deppermann & Schmitt 2007: 22), mit der Entwicklung neuer Analysemethoden zu reagieren. Zur Bestimmung des Gegenstandes skizzieren die Autoren ‚Koordination‘ mit den folgenden Charakteristika:

- 1) ‚Koordination‘ stellt eine permanente interaktionskonstitutive Aufgabe dar („Permanenzcharakter koordinativer Anforderungen“), in der alle in der Situation verfügbaren Ausdrucksmodi theore-

²³ Die Relevanz des Forschungsprogramms, insbesondere für die deutsche Gesprächsforschung, resultiert aus einer vergleichsweise langen Reduktion auf verbalsprachliche Phänomene innerhalb der Mainstreamforschung. Die „Medialität und Intermedialität“ von Sprache ist dabei als „freiwillige Beschränkung“ auf das Auditive lange Zeit ausgeblendet worden (vgl. Holly 2008: 277, s.a. Dausendschön-Gay & Krafft 2002: 32). In amerikanischen Arbeiten (z.B. Goodwin) und britischen Arbeiten der „work place studies“ (z.B. Heath & Luff) lässt sich hingegen eine wissenschaftliche Strömung zur systematischen Beschäftigung mit dem Visuellen seit den 70/80er Jahren beobachten.

²⁴ Die Zusammenführung wird bei Deppermann & Schmitt (2007: 17) wie folgt begründet: „Goffman hat mit seinen Untersuchungen nicht nur den Nachweis der Komplexität der interaktiven Ordnung geführt. Er hat auch die analytische Relevanz der sie konstituierenden zahlreichen multimodalen Aspekte überzeugend herausgearbeitet. Er selbst hat die generativen Mechanismen von Interaktion aber nie systematisch in ihrem tatsächlichen interaktiven Vollzug analysiert. Dies hat für den Bereich der Verbalität die Konversationsanalyse geleistet. Die hierfür notwendige strenge empirische Methode der Konversationsanalyse kann also als Orientierung für die mikroanalytische Auswertung des in Videodokumenten eingefangenen Prozesses der multimodalen Herstellung interaktiver Ordnung dienen.“

tisch gleichwertig nebeneinander stehen („Egalität der Ausdrucksmodi“) und aus einer vollzugsrekonstruktiven Perspektive konversationsanalytisch begriffen werden.

- 2) Konstitutive Aspekte koordinativer Aktivitäten sind Räumlichkeit, Zeitlichkeit, Multimodalität und Mehrpersonenorientierung.
- 3) Es lassen sich zwei Grundformen von ‚Koordination‘ unterscheiden:
 - c) Intrapersonelle Koordination: Aktivitäten, mit denen Interaktanten die unterschiedlichen Ausdrucksmodalitäten des eigenen Verhaltens abstimmen („Selbstorganisation“).
 - d) Interpersonelle Koordination: Anforderungen, die sich auf die „zeitliche, räumliche und multimodale Abstimmung eigener Handlungen und Verhaltensweisen mit denen der anderen Beteiligten beziehen“ (Merkmal der „Adaptivität“) (S. 34).

In ihrer Differenzierung gehen die Autoren von einer „Gleichzeitigkeit beider Grundformen“ im Interaktionsgeschehen aus, in der es keine „exklusiven Aktivitäts- und Ausdrucksformen für intra- und interpersonelle Koordinierungsanforderungen gibt“. Dabei räumen sie ein, dass die Rekonstruktion der einen oder anderen Anforderung nicht immer klar voneinander abgrenzbar ist. Vielmehr postulieren sie für die intrapersonelle Koordination eine partnerorientierte, interaktiv relevante Gestaltung als Regelfall, die folglich interpersonelle Koordinationsprozesse involviert (vgl. S. 37). Insofern handelt es sich um „analytisch zu trennende Konstitutionsaufgaben“ (S. 32).

- 4) Monitoring-Aktivitäten stellen eine zentrale Grundbedingung für die intra- und interpersonelle Koordination dar (s.a. Clark & Krych 2004: 62: „self-monitoring“ vs. „other-monitoring“).
- 5) ‚Koordination‘ wird von ‚Kooperation‘ unterschieden: „Koordination fungiert als Voraussetzung zielorientierter Kooperation, besitzt bei dieser funktionalen Adaption jedoch selbst keinen vollwertigen Handlungscharakter.“ „[...] Die einzelnen Koordinierungsaktivitäten [...] sind aber meist keine Gegenstände interaktiver Verhandlung und reziproker, erkennbar gemachter Deutung, sondern sie haben einen primär organisatorischen Charakter“ (S. 40f.). ‚Koordination‘ beschreibt somit „solche Verhaltensweisen und -aspekte, die im Zusammenhang und zeitgleich mit verbalen Kooperationsbeiträgen und als deren Voraussetzung in den unterschiedlichen Modi körperlichen Ausdrucks realisiert werden [...]“ (S. 22f.).

Auch wenn der Forschungsgegenstand ‚Koordination‘ für die Beschreibung und analytischen Rekonstruktion der Phänomene zur „konfligierenden und funktionalen Simultaneität“ (vgl. 2.3.6, 2.3.7 und 2.3.8) eine fruchtbare Grundlage bietet, verweisen unsere Analysen auf eine handlungstheoretische Diskrepanz, die sich auf die „Nicht-Handlungshaftigkeit“ (Deppermann & Schmitt 2007: 40) koordinativer Aspekte bezieht (Punkt 5). Deppermann und Schmitt konzipieren ‚Koordination‘ als ein Aktivitätenkonglomerat, das sowohl im Verhältnis der Gleichzeitigkeit als auch in relationalem Zusammenhang mit verbalsprachlichen Kooperationsbeiträgen steht. Verhaltensaktivitäten und Handlungen beschreiben jedoch nicht immer trennscharfe Kategorien, sondern manifestieren sich zum Teil als untrennbare Verschränkung innerhalb eines lokalen Handlungsziels. Revisionsbedürftig erscheint dabei im Speziellen die Trennung zwischen der verbalsprachlich semantischen Handlungsebene und den Ausdrucksmodi, die als Ressourcenbündel in koordinativen Prozessen eingesetzt werden. Dieser Differenzierung folgend zählen die Autoren die „Stimme“ und „Lautstruktur“ als Modulatoren der Verbalsprache zu den koordinativen Ressourcen, nicht jedoch die Verbalsprache selbst (vgl. S. 25). Unserer Ansicht nach stellt dies einen Widerspruch zum eingangs formulierten Prinzip der „Egalität der Ausdrucksmodi“ dar (vgl. Punkt 1). Diesbezüglich merkt Werner Holly in einer Rezension des Bandes an: „Ab wann wird ein Blick, eine Geste, eine Körperdrehung, eine Betonung von einer Verhaltensaktivität zu einer Handlung? Da ist eine handlungstheoretische Schwachstelle, die zu wenig plausiblen Grenzziehungen zu führen scheint (verbal/nonverbal, verteilt auf Handlung/Verhalten) und der Klärung bedarf“ (Holly 2008: 279).

Vor diesem Hintergrund verweisen wir exemplarisch auf Fragment 3 (Abbildung 7), in welchem wir eine Polyfunktionalität des adverbialen Konzessivkonnektors „!TROTZ! ALLE!DEM!“ innerhalb des multimodalen Äußerungsformats „!TROTZ! !ALLE!DEM (-) BIN ich FÜR (.)“ rekonstruiert haben. Auf der Handlungsebene wird durch die Adverbialform konzessiv ein anaphorischer Bezug auf etwas „Vorhergehendes [...], das mit dem Nachfolgenden nicht kompatibel ist“ hergestellt (vgl. Fahrländer 2013). Die

semantische Kontrastierung hat dabei jedoch gleichermaßen eine koordinative Funktion, da die Aufmerksamkeit projektiv auf eine neue Handlungsaktivität gelenkt wird, in der eine Entität maßgeblich ist, die in der phrasalen Ausführung des spezifischen Falls (...BIN ich FÜR (.)) bekannt sein muss. Innerhalb der kontextuellen und multimodalen Einbettung der Adverbialform handelt es sich somit um eine koordinative Projektionsform, die prospektiv eine Handlung mit reflexivem Sinnbezug ankündigt. Deppermann und Schmitt sprechen „prototypischen koordinativen Aktivitäten“ jedoch eine Handlungsprojektion ab²⁵ (vgl. S. 40).

Nun ließe sich einerseits anmerken, dass sich die koordinative Funktion des Konzessivkonnektors nicht eindeutig belegen lässt, da die Adverbialform Teil einer multimodalen Orientierungsgestalt mit anderen beteiligten Ausdrucksressourcen ist (z.B. eine als Beat geschlagene Verweisform (vgl. Abbildung 7)). Andererseits könnte man davon ausgehen, dass es sich nicht um eine „prototypische koordinative Aktivität“ handelt. In diesem Zusammenhang können die Analysen der AR-Bedingung weitere empirische Befunde liefern, die zur Präzisierung des Gegenstandes beitragen.

Insbesondere in Transitionsphasen von einer zur nächsten Handlungsaktivität lässt sich für die AR-Bedingung eine koordinative Relevanz diskursstrukturierender Partikel und verbaldeiktischer Ausdrücke für die Herstellung von Ko-Orientierung feststellen (vgl. 3.4.1). Aufgrund des eingeschränkten Sichtfeldes sind die Interaktanten in Transitionsphasen nicht immer ko-orientiert, da die Koordination zum gemeinsamen Handlungsabschluss einer interaktionalen Aufgabe fehlschlägt. Während einer der Akteure noch in der Bearbeitung einer kontemporären Interaktionsaufgabe verharrt (z.B. der Platzierung eines Objekts, Lokalisationsratifikation, „post-proposal accounts“ etc.), verlagert der andere bereits seine Aufmerksamkeit auf ein neues Objekt, dessen Platzierung pre-operativ als nächste interaktionale Aufgabe vorbereitet wird (z.B. Greifen des Objekts, Inspizieren des Objekts, Isolieren aus der Objektgruppe am Kartenrand etc.). In solchen Phasen divergenter Orientierung stehen die verwendeten Ressourcen einer multimodalen Äußerungsgestalt dem jeweiligen Ko-Partizipanten nur in fragmentarisch-sezierter Form zur Verfügung. Unsere Analysen liefern diesbezüglich einerseits empirische Hinweise auf eine interpersonelle Koordinierung durch diskursstrukturierende Partikel mit prospektiver Funktion (z.B. „SO:“ oder „NUN“), die eine gemeinsame Koordinierung auf eine neue interaktionale Aufgabe ermöglichen. Hierzu zählen auch Ratifikationsmarker („so;“ oder „okay;“), die in der Abfolge von Initialratifikation und Gegenratifikation eine äquivalente Koordinierungsfunktion ausüben (vgl. 3.4.1.1). Andererseits lassen sich auch koordinative Anforderung durch den Gebrauch von Diskursmarkern mit retrospektiver Funktion (z.B. „ja wobei:“, „zumal“ oder „na=ja:“) analytisch aufzeigen, die eine Re-Orientierung auf die kontemporäre Aufgabebearbeitung des aufmerksamkeitsverlagernden Ko-Partizipanten bewirken²⁶ (vgl. 3.4.1.2).

Darüber hinaus ist eine hohe koordinative Relevanz verbaldeiktischer Ausdrücke feststellbar (z.B. „HIER“). Die Verwendung von Verbaldeiktika deuten innerhalb unseres Settings, in dem die Akteure ko-präsent und in physischer Nähe zum Verweisraum sitzen, nicht nur auf potentiell relevante Lokalisationsangaben, sondern auch auf weitere körperliche Ressourcen zur Lokalisationsspezifikation (z.B. eine deiktische Geste), die in der holistischen Orientierungsgestalt verwendet werden. Ein Re-Orientierung erfolgt somit aus der praxeologischen Erkenntnis, dass der Referent der Orientierungsgestalt bei einer rechtzeitigen Aufmerksamkeitsverlagerung ermittelt werden kann. Insofern stellen verbaldeiktische Ausdrücke ein Koordinierungsangebot hoher Relevanz zur interpersonellen Herstellung gemeinsamer Orientierung dar (vgl. 3.2).

Vor dem Hintergrund der geschilderten Einwände erscheint uns eine generalisierende Differenzierung zwischen koordinativen Verhaltensaktivitäten und Handlungen nicht plausibel, da sie kein Abbild, sondern einen idealisierten Zugang zur faktischen Komplexität interaktiver Ereignisse schafft²⁷. Daher adap-

²⁵ Ein ähnlich gelagertes Diskrepanzverhältnis – das zwischen Koordinierungsangebot und -ratifizierung – wird bei Wolf (2007: 166) thematisiert.

²⁶ Eine (re-)orientierende Funktion liegt nicht allein in der ‚Semantik‘ des verwendeten Diskursmarkers als prospektiv- hinweisendes oder retrospektiv- rückweisendes Element begründet, sondern hängt entscheidend von seiner sequentiellen Einbettung im Interaktionsgeschehen ab. Ein „wobei:“ oder „na=ja:“ kann sowohl in Ko-Orientierung auf einen Sachverhalt/eine Entität eine Einschränkung ankündigen als auch im Sinne eines Expansionsmarkers zur gegenargumentativen Handlungserweiterung einer vermeintlich abgeschlossenen und ko-ratifizierten Interaktionsaufgabe fungieren.

²⁷ In diesem Zusammenhang erscheint die später im Sammelband formulierte Präzisierung epistemisch relevant, ‚Koordination‘ als Kategorie zweiter Ordnung zu verstehen, „die der Konversationsanalyse die Möglichkeit gibt, Beobachtungen zu systematisieren und in einen neuen Interpretationszusammenhang zu stellen“ (Dausendschön-Gay & Krafft 2007: 191).

tieren wir das Konzept für unsere Analysen in einer konversationsanalytisch rekonstruktiven Form, in der a priori keine Differenzierung zwischen koordinativen Aktivitäten und Handlungen getroffen wird.

2.3.3.2 Beziehungskonstellation von Sequentialität und Simultaneität

Der Fokus auf koordinative Aspekte sozialer Interaktion ist auf die Annahme zurückzuführen, dass die interpersonelle Abstimmung simultan ablaufender Handlungsaktivitäten für die Teilnehmer eine permanente interaktionskonstitutive Aufgabe bildet (vgl. 2.3.3.1, Punkt 1). Bezüglich des bei Deppermann und Schmitt formulierten Erkenntnisinteresses der empirischen Analyse des Vollzugscharakters ‚interaktiver Ordnung‘ (2007: 17) stellt sich die Frage nach der organisationalen Relation der Strukturbildungsmechanismen ‚Sequentialität‘ und ‚Simultaneität‘. Dabei liefern Deppermann und Schmitt (2007: 29f.) folgende Hinweise zur Beziehungskonstellation der Strukturprinzipien:

- 1) Das Prinzip der Sequentialität wird als „nur *ein* konstruktiver Mechanismus der Konstitution interaktiver Ordnung“ betrachtet. „Die geordnete Herstellung der komplexen, multimodalen Realität von Interaktion ist letztlich jedoch nur durch einen weiteren Mechanismus möglich, nämlich Simultaneität.“
- 2) Während sich auf der verbalsprachlichen Äußerungsebene die *Sprecher* in einem „Bezugssystem der Zeitlichkeit“ befinden (Sequentialität als Normalfall), sind *Interaktionsbeteiligte* in einem „raumzeitlichen Bezugssystem“ verortet (Simultaneität als Normalfall).
- 3) Da Räumlichkeit und Zeitlichkeit in einem multimodalen Bezugssystem untrennbar miteinander verbunden sind, existiert unter einer multimodalen Perspektive „das Interaktionsgeschehen grundsätzlich nur als permanente Gleichzeitigkeit koordinierter Verhaltensweisen.“
- 4) „Reine Nachzeitigkeit gibt es [...] immer nur bezogen auf eine isolierte Modalität (etwa Verbalität) zusammen mit einer raumgebundenen Gleichzeitigkeit von Aktivitäten auf anderen Ausdrucksebenen (Blick, Körperbewegung, Gestikulation etc.)“
- 5) „Im Rahmen des übergeordneten bzw. grundlegenden Prinzips der Raum-Zeitlichkeit von Interaktion treten Sequentialität (Nachzeitigkeit), bisher restriktiv verstanden als Nacheinander verbaler Interaktionsbeiträge, und Simultaneität (Gleichzeitigkeit) gleichberechtigt nebeneinander.“²⁸

Vor dem Hintergrund der kontextuellen Einbettung in die „Begründung eines neuen Forschungsgegenstandes“ (ebd.: 15) ist die skizzierte Beziehungskonstellation der Strukturprinzipien als deskriptive Annäherung zu verstehen, die Impulse zur analytischen Präzisierung in der Folgeforschung setzt. Im Vergleich zum empirisch fundierten und methodologisch etablierten Leitprinzip der emischen Rekonstruktion sequentieller Ordnung „from within“ (Garfinkel 1967) erscheinen die Hinweise zum Verständnis von Simultaneität und der damit verbundenen organisationalen Relation zum Strukturmechanismus der Sequentialität eher eine etische Beschreibungsdimension einzunehmen. Dementsprechend sind sie grundsätzlicher Natur und referieren metaperspektivisch auf das, was in der Videodokumentation als Überkomplexität des visuell und auditiv Wahrnehmbaren feststellbar wird. Systemtheoretisch begriffen sind die Postulate der Beziehungskonstellation als Beobachtungen erster Ordnung zu typisieren, die auf einem Unterscheidungsparadigma der Strukturprinzipien beruhen. Dies ist insofern nicht ganz kongruent, als dass Deppermann und Schmitt aus der etischen Beschreibungsdimension interaktiv relevante Handlungsimplicationen emischer Qualität, wie den „Permanenzcharakter koordinativer Anforderungen“ (vgl. 2.3.3.1, Punkt 1) ableiten. Hier stellt sich die Frage, ob die beobachtbare Persistenz simultaner Aktivitäten auch aus einer emischen Perspektive eine teilnehmerrelevante Kategorie bildet? Und ferner, ob die Unterscheidung der Strukturbildungsmechanismen ‚Sequentialität‘ und ‚Simultaneität‘ emischer oder etischer Natur ist?

²⁸ Die Voraussetzung „gleichberechtigt nebeneinander tretender Strukturprinzipien“ ist bereits von Schmitt (2005: 53) formuliert worden, der von einer „Gleichzeitigkeit von Simultaneität und Sequentialität als zentralen Ordnungsmechanismen des interaktiven Handlungsvollzugs“ ausgeht.

2.3.3.3 Zentrale Ergebnisse zur organisationalen Relation der Strukturbildungsmechanismen

Die Analysen zur „konfligierenden und funktionalen Simultaneität“ (vgl. 2.3.6, 2.3.7 und 2.3.8) liefern grundlegende empirische Erkenntnisse, die sich dem dringlich zu schließenden Desiderat der organisationalen Relation der Strukturbildungsmechanismen ‚Sequentialität‘ und ‚Simultaneität‘ aus einer emischen Perspektive widmen. Aus didaktischen Gründen stellen wir einige dieser Erkenntnisse dem analytischen Teil als Orientierung voran, da die empirische Erforschung „struktureller Eigenschaften kinesischer Realisierung“ – im kontrastiven Bezug zu bestehenden konversationsanalytischen Konzepten – umfangreiche systematische Analysen erfordert (vgl. Schmitt 2005: 53). Für unseren Datentyp der aufgabenfokussierten face-to-face Interaktion stellen wir fest:

i. Simultaneität ist permanent da, aber nicht permanent relevant.

Die Einschränkung auf den spezifischen Aufgabentyp erscheint uns insofern wichtig, als dass in anderen Typen fokussierter Interaktion eine weitaus konsistentere Relevanz zur interpersonellen Koordinierung persistent gleichzeitig ablaufender Aktivitäten vorliegen dürfte. Speziell Formen vorwiegend körperlicher Interaktion, wie Goffmans Beispiel des Boxkampfes oder Tanzpaares (1964: 135), verdeutlichen anschaulich die Folgen impermanenter Koordinierung in Aufgabentypen hoher Koordinierungsrelevanz – im ungünstigsten Fall den “Knockout” des Boxers und blaue Füße der Tanzpartnerin. Das Ausmaß der Koordinierungsanforderungen steht somit in Abhängigkeit zum Aufgabentyp und ist dadurch an die Bearbeitung spezifischer lokal und global auftretender konversationeller Probleme²⁹ gekoppelt. Neben dem Aufgabentyp stellt die kontextuelle Bedingung der Interaktionssituation einen entscheidenden Einflussfaktor auf den Grad der Koordinierungsrelevanz dar. So stellen wir für die mutual monitoring limitierende AR Bedingung fest, dass die Bearbeitung erzeugter Relevanzstrukturen nicht normativ bindend ist. Die Nicht-Reaktion auf explizite Koordinierungsanforderung wird unter den spezifischen Bedingungen toleriert (vgl. 3.3, “wait & see” Strategie).

Unter einer Perspektive, in der die Ausprägung koordinativer Anforderungen als Kontinuum zwischen Permanenz und Impermanenz begriffen wird, stellt sich die Frage, unter welchen spezifischen Bedingungen Handlungsaktivitäten koordinative Relevanzen erzeugen und wann keine Koordinierungsrelevanz besteht. Stukenbrocks Hinweis auf die „erforderliche theoretische Unterscheidung zwischen zufällig zustande kommenden, kontingenten zeitlichen Verhältnissen und zeitlich konstituierten Relationen, die nicht zufällig sind [...]“ (2012: 28; vgl. 2.3 – „Definition von Simultaneität“) bildet in diesem Zusammenhang für uns ein grundlegendes Unterscheidungsparadigma.

Die zufällige Kontingenz simultaner Aktivitäten resultiert primär aus der genuin körperlichen Anwesenheitsbedingung als Konstitutionsform sozialer face-to-face Interaktion, die als persistente Gleichzeitigkeit visuell wahrnehmbar wird. Mit Fokus auf die verschiedenen Ausdrucksdimensionen existiert diese Gleichzeitigkeit sowohl für nonverbale wie verbale Aktivitäten. So tritt die Verbalsprache genau genommen auch nicht in ihrer orthographischen Sequentialisierung auf, sondern kann durch verschiedene Formen suprasegmentaler Merkmale konturiert und – selbst bei gleicher sequentieller Lexemabfolge – bedeutungsdistinktiv markiert werden. In diesem Sinne bildet die Verbalebene a priori einen multimodalen, bzw. multidimensionalen Verdichtungsraum, in dem die verschiedenen Ausdrucksdimensionen in einem auditiven Strang konvergieren und somit dem Prinzip der Sukzession unterworfen sind. Im Vergleich zur Lautsprache, die nach Frey in 9 Dimensionen beschrieben werden kann, weist das körperliche Ausdrucksarsenal mit 104 Dimensionen jedoch eine hochdifferenzierte Komplexität auf³⁰ (Frey 1984: 30f.), die in gesteigerten intra- und interpersonellen Koordinierungsanforderungen empirisch sichtbar wird. Auch wenn es sich um eine rein formale Beschreibung morphologischer Bewegungsvarianten handelt, die einerseits aufgrund der schweren Segmentierbarkeit des körperlichen Ausdrucks kein faktisches Abbild darstellen kann (wir verstehen es hier als Annäherung) und andererseits stets nur in Ansätzen an der Bedeutungs-

²⁹ Die Bearbeitung konversationeller Probleme und Aufgaben stellt ein konstitutives Merkmal sozialer Interaktion dar, welches bereits bei Goffman (1974) in der “frame analysis” theoretisch expliziert (vgl. “guided doings”, 22ff) und mit Beginn der “conversation analysis” in verschiedensten Studien tradiert wurde (z.B. Schegloff & Sacks 1973) . In der Gesprächsanalyse liegt mit dem Konzept des Handlungsschemas (Kallmeyer und Schütze 1976) ein Ansatz zur Kategorisierung unterschiedlicher Gesprächstypen und ihrer spezifischen Aufgabenstrukturen vor (Fiehler et al. 1999; Spiegel & Spranz-Fogasy 2001).

³⁰ Frey unterscheidet folgende Ausdrucksträger und Bewegungsdimensionen des Körpers (vgl. Frey 1984: 30f.): Kopf (3), Gesicht (49), Schulter (4), Rumpf (3), Oberarm (6), Hände (18), Becken (2), Oberschenkel (5), Füße (14). Die Beschreibungsdimension der Lautsprache setzt sich aus Sprachlaute (1), Lautstärke (1), Stimmhöhe (1) und Klangfarbe (6) zusammen.

konstitution in multimodalen Äußerungsgestalten beteiligt ist, veranschaulicht Freys Differenzierung die „Überkomplexität“ visuell wahrnehmbarer kinesischer Informationen im interaktionalen Geschehen. Dennoch erleiden wir trotz der massiven und permanenten Gleichzeitigkeit an visuellen Informationen keinen „sensory overload“. Wie ist das möglich? Wie wird mittels der Strukturbildungsmechanismen interaktive Ordnung in einer für die Interaktanten nachvollziehbaren Weise hergestellt?

Unsere Analysen deuten in diesem Zusammenhang auf die außerordentliche Relevanz des Strukturbildungsprinzips der „Teilsimultaneität“ (vgl. Stukenbrock 2012: 31). Für „intrapersonelle Simultaneitätstypen“³¹ stellen wir am Beispiel von Orientierungshandlungen fest:

ii. Intrapersonelle Simultaneität erfolgt nur in Teilsimultaneität.

Dieses Prinzip gründet auf der empirischen Beobachtung, dass multimodale Äußerungsgestalten in einer für den jeweiligen Ko-Partizipanten gerichteten Form gestaltet und dem Sukzessionsprinzip unterworfen werden. Das heißt, intrapersonelle Simultanaktivitäten sind stets in einer genuin sequentiellen Lesart formiert, die für ein sinngemäßes Verständnis der sprecherseitigen Äußerungsgestalt vom Ko-Partizipanten erschlossen werden muss. Neben unseren Fallanalysen in 2.3.6, 2.3.7 und 2.3.8 kann ebenfalls das Konzept des „multimodalen Verdichtungsraums“ als Beispiel teilsimultaner Äußerungsgestalten aufgeführt werden (vgl. Abbildung 8). Auch hier setzen die beteiligten Ausdrucksformen der Schmerzschilderung sukzessiv nacheinander ein (teilsimultan), so dass sie sich bedeutungskonstitutiv für einen gewissen Zeitraum überschneiden. Intrapersonelle Simultaneität ist dabei an eine wahrnehmungsspezifische Anforderung gebunden, in der Handlungen nicht in „simultanen Abbildern“ konstituiert, sondern nur gerichtet, in vektorialen Ablaufformaten vollzogen werden können. Terminologisch kann diese Art der Gleichzeitigkeit als „sequentielle Simultaneität“ begriffen werden – eine Simultaneitätsform, die in sequentielle Teilsimultaneitäten aufgelöst wird. Unsere Analysen liefern zusätzlich empirische Hinweise auf eine interaktionsrelevante Funktion bei der Bearbeitung von Fortsetzungsproblemen: Situative Ambiguitäten können durch Formen „sequentieller Simultaneität“ instantan linearisiert und so in ein für den Ko-Partizipanten geeignetes Interpretationsformat überführt werden (z.B. Fragment 4, „Instantane situative Desambiguierung“).

Wenn intrapersonelle Formen „sequentieller Simultaneität“ aufgrund der Anwesenheitsbedingung kopräsender Akteure aufeinandertreffen, entsteht interpersonelle Simultaneität als Permanenzbedingung sozialer Interaktion („zufällige Kontingenz simultaner Aktivitäten“ = körperliches Verhalten kopräsender Akteure). Diese Form der Simultaneität ist axiomatisch. Sie bildet eine Normalformervartung der Interaktanten, die durch ‘monitoring’-Aktivitäten begleitet wird, per se jedoch keine Koordinierungsanforderungen stellt.

iii. Interpersonelle Simultaneität bildet eine Normalformervartung.

Explizite Koordinierungsanforderungen hingegen sind an das Konzept der Orientierungsangebote gekoppelt, die nicht nur Hinweise darauf liefern, wie eine Aktivität zu verstehen und zu behandeln ist, sondern auch, wann eine Aktivität koordinierungsrelevant wird (vgl. „On-Line-Hilfen“: Dausendschön-Gay & Krafft 2002; „Rendering activities visible“: Heath & Luff 1992a; Heath et al. 2002; „Kontextualisierungshinweise“: Auer 1986; 1992; Gumperz 1992;).

Unter einer Perspektive, in der die Ausprägung koordinativer Anforderungen variiert, stellt sich abschließend die Frage, unter welchen interaktionalen Bedingungen eine erhöhte Koordinierungsrelevanz zu beobachten ist. Unsere Analysen zeigen, dass sofortiger Handlungsbedarf in Fällen besteht, in denen das Prinzip der Teilsimultaneität verletzt wird. Dies ist insbesondere dann der Fall, wenn intrapersonelle Simultanaktivitäten der Interaktanten zeitgleich einsetzen („echte Simultaneität“), auf der gleichen Modalitätsebene konkurrieren und so die sequentiellen Abläufe der Aufgabenbearbeitung stören. In diesen Fällen sprechen wir von „konfligierender Simultaneität“ (vgl. 2.3.6, 2.3.7), da keine eindeutige Desambiguierung der Situation vorliegt, sich die sequentiellen Lesarten der aufeinandertreffenden Teilnehmeraktivitäten gegenseitig annullieren und die Interaktanten das Problem der interaktionalen Fortsetzung koordinieren müssen. Für die Bearbeitung des Fortsetzungsproblems lassen sich unterschiedliche Verfahren rekonstru-

³¹ In Anlehnung an das Konzept der „intra- und interpersonellen Koordination“ (vgl. i, Punkt 3a+b) unterscheiden wir zwischen „intra- und interpersoneller Simultaneität“ und markieren somit die verschiedenen Konstitutionsformen von Simultaneität als Ausgangsbedingungen zur Koordinierung.

ieren („Formen gestischer Sequentialisierung“, „Sequentialisierung durch Direktiva“ und „Sequentialisierung durch Blickkoordination“ (vgl. 2.3.6, 2.3.7)), mit denen Handlungsparallelität in ein sequentielles Ablaufformat überführt wird. Das empirisch begründete Prinzip lautet daher:

- iv. Wenn interpersonelle Simultaneität sequentielle Abläufe der Aufgabenbearbeitung stört, wird Handlungsparallelität dem Prinzip der Sukzession unterworfen.

Die Beziehungskonstellation der Strukturprinzipien ‚Sequentialität‘ und ‚Simultaneität‘ kann abschließend präzisiert werden: Als Unterscheidungskriterium der Beobachtung erscheinen sie eine primär etische Beschreibungsdimension einzunehmen. Insofern bestätigen unsere analytischen Ergebnisse, dass unter einer multimodalen Perspektive „das Interaktionsgeschehen grundsätzlich nur als permanente Gleichzeitigkeit (koordinativer Verhaltensweisen)“ (vgl. 2.3.3.2, Punkt 3) existiert. Mit Verweis auf Stukenbrocks Unterscheidungsparadigma zwischen „zufällig zustande kommenden, kontingenten zeitlichen Verhältnissen und zeitlich konstituierten Relationen [...]“ (2012: 28) klammern wir den „Permanenzcharakter koordinativer Anforderungen“ (vgl. 2.3.3.1, Punkt 1) jedoch aus. Handlungsleitend erscheint hier vielmehr eine Maxime der Offenheit und Flexibilität, in der monitoring-Aktivitäten das Erkennen markierter Koordinierungsrelevanzen sicherstellen.

Aus einer emischen, vollzugsrekonstruktiven Perspektive ist es zusätzlich relevant, das Postulat „reine Nachzeitigkeit gibt es [...] nur bezogen auf eine isolierte Modalität (etwa Verbalität) zusammen mit einer raumgebundenen Gleichzeitigkeit von Aktivitäten auf anderen Ausdrucksebenen [...]“ (vgl. 2.3.3.2, Punkt 4) analytisch zu präzisieren. Die beiden Strukturbildungsmechanismen erscheinen hier weniger distinktiv als angenommen, da die „Gleichzeitigkeit der Aktivitäten auf den anderen Ausdrucksebenen (Blick, Körperbewegung, Gestikulation etc.)“ (ebd.) überwiegend nicht in „echter Simultaneität“ erfolgt, sondern durch sukzessive Nominierung in „sequentielle Lesarten“ strukturiert wird. Insofern markiert „sequentielle Simultaneität“ eine Grundordnung des interaktiven Geschehens. Verstehen ist in diesem Sinne an die Konstitution einer Richtung gebunden. Im Vergleich zur Verbalität, die zumindest in ihrer Konvergenz der Ausdrucksdimensionen als „reine Nachzeitigkeit“ begriffen werden kann, handelt es sich jedoch primär um Mikroformen der Nachzeitigkeit, die auf dem Prinzip der „Teilsimultaneität“ beruhen.

2.3.4 Verhältnis der overlap-Konzeption zur kinesischen Simultaneität

Unter Verweis auf die in 2.3.3 vorangestellte Präzisierung der Beziehungskonstellation der Strukturbildungsprinzipien ‚Sequentialität‘ und ‚Simultaneität‘ stellt sich die Frage, wie sich die verbalsprachliche Konzeption der verbalen Gleichzeitigkeit im Rahmen der Rederechtsorganisation zur Simultanform kinesischer Handlungsweisen verhält. Das Interesse ist dabei nicht nur epistemischer Natur. Die Validität der organisationalen Prinzipien (2.3.3, i.-iv.) bemisst sich letztendlich auch an ihrem analytischen Potential, eine für kinesische Simultanaktivitäten äquivalente Topographie zur verbalsprachlich explizierten „overlap“-Konzeption zu erstellen, in der die strukturellen Gemeinsamkeiten und Unterschiede ableitbar und empirisch replizierbar sind.

Ausgehend von einer phänomenologischen Fokussierung auf die Verbalsprache haben Sacks, Schegloff und Jefferson (1974) in ihrer „simplest systematics“ ein grundlegendes Differenzierungsparadigma eingeführt, das zwischen „silences“ und „overlaps“ unterscheidet, wobei „silences“ eine weitere Subkategorisierung in „gaps“, „pauses“ and „lapses“³² erfahren, die im Modell als „ways of conceiving the appearance of silence in conversation“ (ebd.: 715) beschrieben werden. Die drei Kategorien sind syntaktisch motiviert und werden zum Teil als ineinander transformierbare Einheiten betrachtet:

“[...] parties’ treatment of silence in conversation is contingent to its placement. To put it roughly: intra-turn silence (not at a transition-relevance place) is a ‘pause’, and initially not to be talked in by others; silence after a possible completion point is, initially, a gap, and to be minimized; extended silences at transition-relevance places may become lapses. But some silences are transformable. Thus, if a developing silence occurs at a transition-place, and is thus a (potential) gap, it may be ended by talk of the same party who was talking before it;

³² „Lapses“ werden bei Sacks et al. auch als Diskontinuitätsform beschrieben: “Discontinuities occur when, at some transition-relevance place, a current speaker has stopped, no speaker starts (or continues), and the ensuing space of non-talk constitutes itself as more than a gap – not a gap, but a lapse” (SSJ 1974: 714).

so the 'gap' is transformed into a 'pause' (being now intra-turn). This is one way that 'gap' is minimized.“
(Sacks et al. 1974: 715)

Die intrapersonelle Konstitutionsform “pause” und die zeitlich expandierte Form “laps” bilden gewissermaßen Sonderformen, so dass dem Modell eine übergeordnete Relevanz der Kontrastierung zwischen “gap” und “overlap” zu entnehmen ist, die sich auch in der Formulierung der als Regularitäten (“grossly apparent facts”, SSJ 1974: 700) verstandenen Prinzipien widerspiegelt. Auf den Gegenstand bezogen sind vor allem vier der vierzehn Prinzipien des turn-taking Modells relevant (ebd.: 700f.):

- 1) “Overwhelmingly, one party talks at a time.”
- 2) “Occurrences of more than one speaker at a time are common, but brief.”
- 3) “Transitions [...] with no gap and no overlap are common. Together with transitions characterized by slight gap or slight overlap, they make up the vast majority of transitions.”
- 4) “Repair mechanisms exist for dealing with turn-taking errors and violations; e.g. if two parties find themselves talking at the same time, one of them will stop prematurely, thus repairing the trouble.”

Phänomenologisch werden “gap” und “overlap” als komplementäre Erscheinungsformen verstanden, deren Auftreten an “transition relevance places” (TRPs) gebunden ist und unter den Regularitäten des turn-taking Systems zu minimieren versucht wird³³. Während sich der Minimierungsanspruch des “overlaps” aus der physikalischen Ko-Okkurrenz akustischer Signale als klassischer Störfall manifestiert, erscheint das Prinzip der Vermeidung von Gesprächslücken intuitiv weniger verständlich. In diesem Zusammenhang stellen Kallmeyer & Schmitt (1996: 46; s.a. Schmitt 2005: 51) fest: „Die Konversationsanalyse hat dem Modell der Redevertelung die Vorstellung zugrunde gelegt, daß Redegelegenheiten in der Interaktion prinzipiell ein knappes Gut sind, d.h. im Normalfall mehr als einer sprechen will.“ Insofern ist nicht entscheidbar, ob diese Auffassung auch im turn-taking Modell unterstellt wurde oder ob es sich um eine post-inskribierte Lesart der Community handelt.

Der Minimierungsanspruch von “gap” und “overlap” ist in seiner konzeptionellen Generalisierung kritisch ausdifferenzieren. Einerseits lässt sich aus einer anthropologisch-ethnographischen Perspektive anmerken, dass die „soziostilistischen Normalformen auch in Bezug auf die Bedeutung des Sprechens für die Herstellung und Aufrechterhaltung sozialer Situationen“ mit Verweis auf „Rede- und Schweigekulturen“ sehr unterschiedlich sind (vgl. Kallmeyer & Schmitt 1996: 46; s.a. Sidnell 2001). Andererseits ist darauf hinzuweisen, dass keine definitorische Übereinkunft über eine konstitutive Normalform der Interaktion vorliegt, bzw. vorliegen kann. Ob längere Gesprächspausen und “overlaps” letztendlich von der Normalform erwartung abweichen, hängt vom Interaktionstyp und den lokal wie global zu bearbeiteten interaktionalen Aufgaben ab (z.B. „aufgabenfokussierte Interaktion“ vs. spontane “face engagements”). Schmitt (2005) merkt diesbezüglich an:

„Da es bei der turn-Organisation im Prinzip nicht um Rederecht als solches, sondern um qualifizierte Gelegenheiten für bestimmte Aktivitäten geht [...] ist der Status des legitimen Sprechers kontextsensitiv und auch davon abhängig, wie der Sprecher die qualifizierte Gelegenheit faktisch nutzt.“ (Schmitt 2005: 51)

Diese Kontextsensitivität manifestiert sich neben exogenen Faktoren auch in dynamischen Veränderungen und Neukonfigurationen, die innerhalb größerer Interaktionsepisoden emergieren (vgl. Kapitel 4). Die Minimierung von “gap” und “overlap” ist somit nicht als generalisierendes Prinzip zu verstehen, sondern als eine Leistung des turn-taking Systems auch unter „den Bedingungen der Knappheit von Redegelegenheiten zu funktionieren“ (vgl. Kallmeyer und Schmitt 1996: 46).

Als Potential des Systems begriffen, stellen Studienergebnisse zur statistischen Distribution von Gesprächspausen und verbalsprachlicher Simultaneität keine Falsifikation der initialen Prinzipien dar (SSJ 1974, vgl. insbesondere die Punkte 1-4), sondern empirische Hinweise auf ihre Gewichtung in sozialen Interaktionsituationen. Auch wenn diese Studien unter Verweis auf die a) Hinweise zur Kontextsensitivität, b) Abhängigkeit vom Interaktionstyp und c) der zum Teil unzureichenden Berücksichtigung der in der

³³ “The rule-set, along with the constraints imposed mutually by the opinions in it, eliminates gap and overlap from most of conversation by eliminating gap and overlap from most single turns” (SSJ: 1974: 705).

Folgeforschung geleisteten Präzisierung der “overlap” -Konzeption (siehe unten) nicht als generalisierendes Handlungsprinzip verstanden werden sollten, manifestiert sich eine Tendenz, nach der vor allem das Prinzip “slight gap/slight overlap” (Punkt 3) eine zentrale Relevanz im Rahmen der Rederechtsorganisation einzunehmen scheint (vgl. Heldner & Edlund 2010; Weilhammer & Rabold 2003; Kurtic et al. 2013).

Mit Fokus auf die verbalsprachliche Gleichzeitigkeit lässt sich in der Folgeforschung zur “turn-taking machinery” eine Ausdifferenzierung der “overlap”-Konzeption feststellen, in der die Minimierungsanforderungen anhand des Störpotentials bemessen werden (vgl. Abbildung 9). Dabei wird zwischen zwei Grundformen unterschieden:

“There are [...] two distinct forms of overlap competition [...]; one announces a trouble and attempts to remedy it [...] and one neither recognizes nor attempts to remedy an observable trouble. We might call the former procedure *Marked Competition* and the latter *Unmarked Competition*.” (Jefferson 2004: 49f.)

Dieser Unterscheidung folgend lassen sich phänomenologisch unterschiedliche Erscheinungsformen der verbalen Gleichzeitigkeit kategorisieren (vgl. Schegloff 2000: 5f.; s.a. Mondada & Oloff 2011: 322). Als unproblematisch behandeln die Interaktanten Simultanformen, in denen der Rezipient als aktiver „Mitwirker“ (Dausendschön-Gay 2000) fungiert und somit potentiell zur Bedeutungskonstitution der emergierenden Äußerungsgestalt im Sinne der Ko-Konstruktion beiträgt. Hierzu zählen Hörersignale der Verständnis- und Orientierungssicherung, deren terminologische Variationen nicht immer phänomenologisch begründet sind, sondern zum Teil aus einer Abgrenzung unterschiedlicher Perspektiven auf das Phänomen resultieren (“backchannels” (Yngve 1970); “continuers” (Schegloff 1982); “reactive tokens” (Clancy et al. 1996); “response tokens” (Gardner 2001)). Ebenfalls unproblematisch werden Instanzen behandelt, in denen eine Beendigung des Redebeitrags auf Basis syntaktisch- prosodischer Projektionen antizipiert wird (“terminal overlaps” (Jefferson 1983)). “Terminal overlaps” werden dabei als zeitlich kurze Formen charakterisiert (vgl. “slight overlap” (SSJ 1974: 700)): “Here the very auspices under which the overlap comes into being project its almost immediate self-liquidation, as the incipiently finishing speaker brings the prior turn into completion [...]” (Schegloff 2000: 5). Weitere Instanzen unproblematischer Gleichzeitigkeit bilden Hilfestellungen bei Formulierungsschwierigkeiten³⁴ (“conditional access to the turn” (Lerner 1996)) und chorale Sprachformen (“laughter”, “collective greetings”, “leave-takings” etc.).

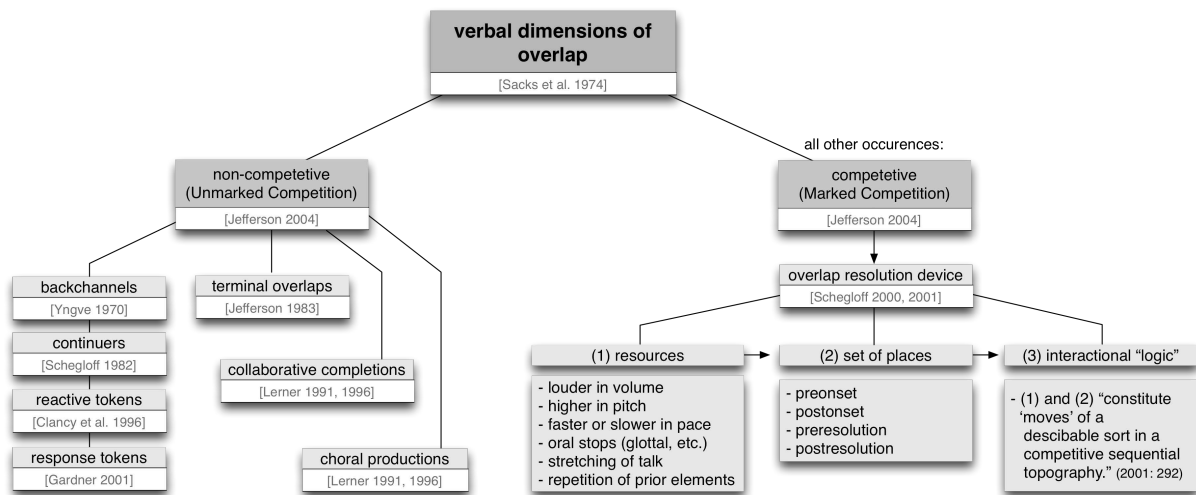


Abbildung 9: “Overlap”-Differenzierung anhand des Störpotentials, basierend auf Schegloff 2000

Alle anderen Vorkommen verbalsprachlicher Gleichzeitigkeit manifestieren sich in einem Störpotential, das von den Interaktanten bearbeitet wird. Phänomenologisch sind diese Instanzen bei Schegloff als Kontinuum zwischen “hitches” und “perturbations” konzeptualisiert worden:

³⁴ Nach Schegloff (2000) manifestieren sich diese Hilfestellungen vornehmlich bei Wortfindungsproblemen und kollaborativen Satzkonstruktionen.

“[...] stretches of overlapping talk are characterized by hitches and perturbations in the talk. By that I mean deflections in the production of the talk from the trajectory which it had been projected to follow. These can then be momentary arrests in the continuity or “progressivity” (Schegloff 1979) of the talk’s production – what I call ‘hitches’ – or marked departures from the prosodic character of the talk’s articulation to that point – what I call ‘perturbations.’” (Schegloff 2000: 11)

Die Verfahren zur Bearbeitung konkurrierender Verbalaktivitäten sind als “overlap resolution device” (Schegloff 2000, 2001) systematisiert worden, der aus drei Komponenten besteht:

- 1) Einem Set an Ressourcen, aus denen sich Okkurrenzen verbaler Gleichzeitigkeit im Spektrum zwischen “hitches” und “perturbations” konstituieren. Dies sind vor allem Lautstärke, Tonhöhe, Sprachgeschwindigkeit, Abbrüche (glottal, labial, dental etc. stops), Längung und die Wiederholung voriger Elemente, die in kombinierten Variationen auftreten können.
- 2) Orte, an denen diese Ressourcen eingesetzt werden (“pre-/postonset” und “pre-/postresolution”, s.a. Jefferson 1983).
- 3) Eine „interaktive Logik“, nach der die Komponenten (1) und (2) in wahrnehmbaren Handlungsvollzügen (“moves”) konstituiert werden.

In Bezug auf die Entwicklung einer äquivalenten Topographie für kinesische Simultanaktivitäten erscheint Schegloffs Hinweis relevant, dass die Grundformen “non-competitive” und “competitive” keine trennscharfen Kategorien bilden: “It is worth underscoring that the ‘exemption’ for these cases of overlapping talk pertains only in the first instance, i.e. for developing a first-order-account” (ebd. 2000:6). Eine übergeordnete Bedeutung liegt hier in der Frage, ob die Interaktanten selbst die verbalsprachliche Simultaneität als legitime Handlungsweise oder Störung einer spezifischen Situation betrachten. Darüber hinaus stellen wir für konfligierende Simultanaktivitäten auf kinesischer Handlungsebene fest, dass ein anfängliches Störpotential in der weiteren Bearbeitung einer interaktionalen Aufgabe durchaus funktional sein kann. Dies ist insbesondere dann der Fall, wenn aus einer Störungsbearbeitung konkrete Handlungsabfolgen resultieren. Eine suspendierte Simultanaktivität fungiert in diesen Fällen als Display einer “in-petto-Konfiguration”, die nach Beendigung der kontemporären Aktivität als nächster “move” relevant wird (vgl. 2.3.7.1).

Unter einer multimodalen Perspektive auf Interaktion fällt auf, dass zwar die Beteiligung nonverbaler Ausdrucksressourcen bei der Rederechtsorganisation untersucht wurde (z.B. Kendon 1967, 1978; Argyle & Cook 1976; Goodwin 1980, 1981; Egbert 1997; Bavelas et al. 2002; Mondada 2007b), jedoch kaum systematische Analysen zu den “embodied practices of overlap resolution” vorliegen. Erste Ansätze finden sich bei Oloff (2012, 2013; s.a. Mondada & Oloff 2011), die das Phänomen des Turnverzichts in Sequenzen verbaler Gleichzeitigkeit (“withdrawal from a turn during simultaneous talk”) mit einem Fokus auf visuell- gestische Ressourcen untersucht und konzeptuell an die Formen der Handlungsdemonstration (“participation”, Goodwin & Goodwin 2004) geknüpft hat:

“The problem of continuing or stopping a turn-in-progress in overlapping talk can be closely linked to the participation framework (Goodwin and Goodwin 2004), as speakers do visibly take into account their recipient’s availability and coordinate their turn construction with the dynamic changes of the participation framework and the interactional space.” (Oloff 2012: 207)

In analytischer Präzisierung der Schegloff’schen Topographie zur Bearbeitung verbalsprachlicher “overlaps” wird auf die Beteiligung unterschiedlicher Ausdrucksressourcen beim “overlap resolution device” verwiesen:

“[...] speakers can make use of a variety of resources – not only increasing/decreasing volume, prosody, repetition and restarts, sound stretches and rush-throughs (Schegloff 2000), but also various kinds of gestures, changes of bodily posture, movements, object manipulations, and gaze. Even if overlap is a phenomenon primarily defined by the juxtaposition of verbal and vocal resources, participants also deploy visible resources in order to manage it.” (Mondada & Oloff 2011: 337)

Es ist festzustellen, dass Oloff und Mondada zentrale Beobachtungen und Erkenntnisse in Bezug auf die multimodalen Verfahren der “overlap resolution” liefern³⁵, den “overlap” selbst jedoch nicht als multimodale Konstitutionsform thematisiert haben.

Empirische Analysen, die sowohl in der analytischen Fokussierung auf den Gegenstand als auch seiner interaktiven Bearbeitung multimodale Konstitutionsformen adressieren, finden sich bei Schmitt (2005). Basierend auf Daten eines Lehr-Lern-Settings, in dem eine Gruppe von Filmstudenten mit zwei Dozenten Ideen für eine Film besprechen, wird die kinesische Turnbeanspruchung einer Studentin analysiert. In Konkurrenz zu den gegenwärtigen Ausführungen eines Dozenten beansprucht sie über einen Zeitraum von insgesamt 26 Sekunden – mit wachsender Expressivität an Melde- und Bittgesten und unter Bezug auf die aktuelle Kommunikationsentwicklung – die Sprecherrolle. Anders als es in der verbalsprachlichen Konzeption des “competitive overlap” diskutiert wurde, ist weder eine Orientierung an turn-internen Pausen noch an transition relevance places beobachtbar (ebd.: S. 33). Ein Störpotential ihrer Meldeaktivität wird jedoch laut Schmitt auf zweierlei Weise deutlich. Einerseits reagiert der Dozent auf die Beanspruchung des Rederechts mit einer „Störungsabwehr“ („sofort“), die ihre Etablierungsanstrengungen zeitweilig suspendiert, andererseits stellt die zeitliche Dauer und Vehemenz ihres Verhaltens einen „Verstoß gegen die situative Interaktionsordnung“ dar. Dabei kann davon ausgegangen werden, dass sie die „Beeinträchtigung der Darstellungsgelegenheit“ ihrer Kommilitonen auch selbst erkennt, da sie ihr Verhalten nach dem Erhalt des Rederechts mit dem account „ich muss es jetzt loswerden ich platze“ legitimiert (ebd.: 44).

Unter der Fragestellung, inwieweit die kinesisch realisierte Turnbeanspruchung mit der konversationsanalytischen Vorstellung des “overlap” korrespondiert, stellt Schmitt fest: Alle Aspekte der verbalsprachlich explizierten “overlap”-Konzeption sind auch in der kinesischen Realisierung der Studentin konstitutiv, manifestieren sich jedoch in einem deutlich geringeren Störpotential. Insofern scheint es hier keine Normalformerwartung im Sinne der verbalsprachlichen Konstitutionsform “one speaker at a time” zu geben (ebd. 46f.). In kontrastiver Exploration der Ursachenfaktoren wird auf unterschiedliche Formen der Aufmerksamkeitskonkurrenz verwiesen:

“Im klassischen Fall des verbalen Overlaps handelt es sich um eine Konkurrenz innerhalb einer sensorischen Modalität (Akustik). Mit dem hier untersuchten Verhalten ist zwar auch eine Aufmerksamkeitskonkurrenz und damit ein Störpotential verbunden, die Störung erfolgt jedoch nicht innerhalb der gleichen Modalität. Dadurch wirkt sich das Verhalten weniger beeinträchtigend aus.“ (Schmitt 2005: 47)

Daraus folgernd markiere die Ko-Okkurrenz verbaler Aktivitäten gewissermaßen einen Sonderfall. Auf die konzeptuelle Relevanz solcher Fälle verweisend, in denen Handlungsaktivitäten innerhalb der gleichen Modalität konkurrieren, eröffnet Schmitt Anschlussoptionen für die Folgeforschung zur Präzisierung des Gegenstandes.

Anknüpfend an Schmitts Analysen, präsentiert Pitsch (2006) am Beispiel von Schülermitschriften im bilingualen Geschichtsunterricht Fälle zur Koordinierung parallelemergender Handlungen im Zusammenhang zwischen Unterrichtsdiskurs, Orientierungshinweisen und Mitschreib-Aktivitäten. Ihren Analysen liegt die, insbesondere für den spezifischen Aufgabentyp ‚Unterrichtskommunikation‘, relevante Unterscheidung zwischen a) „erwünschter“ und b) „konkurrierender Simultaneität“ zugrunde.

Mitschreib-Aktivitäten werden als erwünschtes Beteiligungsformat des Unterrichtsdiskurses adressiert, die im Gegensatz zur “overlap”-Konzeption als „Störfall“ keinen Minimierungsanforderungen unterliegen. ‚Mitschreiben‘ und die inhaltlich-thematische Entwicklung des Unterrichtsdiskurses beschreiben vielmehr „zwei eigenständige Aktivitätsebenen mit eigenen Logiken, die sequentiell aneinander gekoppelt und miteinander durch Orientierungshinweise verbunden werden“ (ebd.: 189). Für das erwünschte Beteiligungsformat stellt Pitsch fest, dass formale Abweichungen der Gestaltung und inhaltliche Fehler zwischen einer angebotenen Tafelinskription und einer Mitschrift auf Koordinierungsprobleme der parallel-emergenten Aktivitäten (‚Unterrichtsdiskurs‘ und ‚Mitschrift‘) zurückzuführen sind. Koordinierungspro-

³⁵ Siehe z.B. Oloffs “preliminary observations” (2012: 230): “First, if, in overlap, a speaker gazes at a non-gazing (and overlapping) addressee, he is likely to withdraw quickly from the turn as soon as he notices his interlocutor’s absence of gaze. Second, if a speaker is addressed in overlap by one of his co-participants, his withdrawal from the turn is often closely timed with a re-orientation of his gaze toward the overlapping speaker. Third, if a mutual gaze between two participants is already established when overlap occurs, a re-orientation to an alternative, non-overlapping participant may lead to a quick abandonment of the turn by the speaker who lost his recipient.”

bleme manifestieren sich im Zuge dessen einerseits in Anforderungen zur Re-Organisation emergierender Mitschriften als Folge einer zeitlichen Divergenz zwischen Angebotsstruktur („Tafelinskription“) und Startpunkt der Aufzeichnung sowie Erweiterungen und Bedeutungsverlagerungen materieller Strukturen der Tafelinskription im Rahmen des prozessualen Voranschreitens inhaltlicher Arbeit. Andererseits resultieren sie auch aus dem selbständigen ‚Notieren‘ (in Abgrenzung zum ‚Kopieren‘) kommunikativ angebotener Inhalte als sequentielle Abkopplung der parallelen Ebenen ‚Mitschreib-Aktivität‘ und ‚Unterrichtsdiskurs‘ (vgl. Pitsch 2006: 186ff.).

Im Rahmen einer „overlap“ äquivalenten Konzeption konkurrierender Simultanaktivitäten präsentiert Pitsch zwei Fälle, in denen (1) „simultane Ereignisse mit gleichem interaktiven Fokus“ und (2) die „Gleichzeitigkeit von noch laufender und bereits projizierter nächster Handlungen beim Übergang zwischen zwei Aktivitäten“ fokussiert werden. Ausgehend von einem klassischen Fall der verbalsprachlichen Gleichzeitigkeit in (1) zeigt sie, dass die durch den „overlap“ entstandene Situation alternativer Fortsetzungsoptionen nicht nur durch eine verbalsprachliche Sequenzierung („overlap resolution device“), sondern auch durch ein körperliches Projektionsformat bearbeitet wird. Die vorangehende Verbalaktivität wird im untersuchten Fall zwar suspendiert, aber körperlich als nächste Aktivität aufrechterhalten (S. 198f.). Das Verfahren wird als eine „Transformation der Gleichzeitigkeit“ konzeptualisiert, in der die problematische Simultaneität auf Verbalebene in eine unproblematische auf der Ebene des körperlichen Ausdrucks umgewandelt wird.

Der Bearbeitung konkurrierender Aktivitäten durch körperliche Projektion wird in (2) ein Fall gegenübergestellt, in dem eine projektive Gestalt konkurrierende Aktivitätsanforderungen erzeugt. Im untersuchten Fall handelt es sich um ein intrapersonelles Format, das durch eine beidhändige Äußerungsgestalt konkurrierende Aufmerksamkeitsangebote konstituiert. Auch wenn die Reaktionen der Gesprächspartner auf das kommunikative Angebot der simultanen Orientierung aufgrund der Kameraposition schwer erkennbar sei, weist Pitsch darauf hin, dass das Ausbleiben einer Reparatur bezüglich der konkreten Aufmerksamkeitsfokussierung des multiplen Orientierungsangebotes auf eine positive Bestätigung des Verfahrens deute (ebd.: 200). Pitsch kommt zu dem Ergebnis, dass die intrapersonelle Konstitution des Orientierungsangebotes auf die „interaktive *Dimension* der Koordinierung paralleler Aktivitäten in einem Individuum“ verweist. Das Beispiel liefert damit einen empirischen Hinweis auf die Konzeption intrapersoneller Koordinierung nach Deppermann & Schmitt, in der von einer interaktiv relevanten Gestaltung als Regelfall ausgegangen wird (vgl. Deppermann & Schmitt, 2.3.3.1, Punkt 3).

2.3.5 Ableitung des Forschungsdesiderats

Aus einem Überblick der zentralen analytischen Ergebnisse zu Simultanaktivitäten und multimodalen Verfahren der „overlap resolution“ ergeben sich folgende Beobachtungen, aus denen sich das zugrundeliegende Forschungsdesiderat der nachfolgenden Analysesektion konstituiert:

- Es liegen Analysen vor, die mit einem Fokus auf gestisch- visuelle Beteiligungsweisen verkörperte Verfahren der „overlap resolution“ am Beispiel des Turnverzichts adressiert haben (Oloff 2012, 2013; Mondada & Oloff 2011). Die gewonnenen empirischen Hinweise konzentrieren sich auf multimodale Bearbeitungsstrategien verbalsprachlicher Störfälle des „overlap“.
- Schmitt (2005) liefert in seiner Analyse aus einem Lehr-Lern-Setting empirische Hinweise auf Formen kinesischer Simultaneität als Störfaktor, verweist in der Kontrastierung zur verbalsprachlichen „overlap“-Konzeption auf unterschiedliche Störpotentiale und eröffnet der Folgeforschung differenzhypothetisch typäquivalente Störfälle zwischen verbalsprachlicher und kinesischer Simultanaktivität.
- Pitsch (2006) greift die unterschiedlichen Störpotentiale auf und eröffnet mit den in der Spezifik der Unterrichtskommunikation liegenden Kategorien „erwünschte und konkurrierende Simultaneität“ eine maximale Kontrastierung von Simultanaktivitäten, die sich phänomenologisch an die verbalsprachliche Konzeption zwischen „Unmarked“ und „Marked Competition“ anschließt. Zusätzlich wird mit dem Fall der Modalitätssubstitution (kinesische Aufrechterhaltung der durch den verbalsprachlichen „overlap“ entstandenen Fortsetzungsoptionen) eine Perspektive auf das Potential dynamisch-multimodaler Anpassungen bei Störfällen eröffnet.

Aus einem Überblick multimodal- analytischer Arbeiten zur kinesischen Simultaneität und “overlap resolution” wird deutlich, dass der Phänomenbereich bislang nur in Ansätzen durchdrungen und mit Fokus auf spezifische Einzelphänomene des Gegenstands untersucht wurde. Dabei ist auch anzumerken, dass die Untersuchung im Rahmen institutionalisierter Lehr-Lern-Settings (Schmitt 2005; Pitsch 2006) eine Form der sozialen Wirklichkeit abbildet, in der institutionelle Handlungsnormen und -erwartungen sprachliche Handlungsmuster tradieren (vgl. Brünner 2005: 171ff.). Insofern liegt kein gleichberechtigter, sondern ein intrapersonell regulierter Zugang³⁶ zum Rede- und Handlungsrecht vor. Die asymmetrische Beschränkung der Rederechtsverteilung prädestiniert dabei auch das Störpotential konkurrierender Aktivitäten und verweist infolgedessen auf die Erfordernis zur analytischen Präzisierung des Gegenstands in symmetrischen Kommunikationssituationen, in denen beide Interaktanten in ihren Mechanismen zur Regulierung parallelememergenter Aktivitäten gleichberechtigt sind. Das Oberseeszenario liefert dafür eine idealtypische Situation, in der sich die Interaktanten einen gleichberechtigten Zugang zu Handlungsoptionen unterstellen. Unter diesen kontextuellen Bedingungen sind Okkurenzen simultaner Aktivitäten beobachtbar, die zum Teil auf gleicher Modalitäts-, Fokus- und Handlungsraumbene konkurrieren. Ähnlich zur verbalsprachlichen “overlap”-Konzeption ist in den untersuchten phänomenologischen Erscheinungsformen eine massive Verdichtung multimodaler Ausdrucksressourcen sowohl im Hinblick auf die Konstitution der parallelemergenten Aktivitäten als auch ihrer interaktiven Bearbeitung beobachtbar. Daraus resultiert auch eine zeitliche Verdichtung, die analytisch ein vehement höheres Auflösungs-niveau fordert, in dem die aufeinandertreffenden Simultanformen in ihrer zumeist sukzessiven Nominierung der sich teilsimultan überlappenden Ausdrucksdimensionen rekonstruiert werden. Der Zeitraum zwischen emergierendem “overlap” und abgeschlossener “overlap resolution” umfasst dabei oftmals nur wenige Sekunden. In der kontrastiven Analyse zwischen „funktionaler“ und „konfligierender Simultaneität“ manifestiert sich im Zuge dessen eine zentrale Feststellung als Schlüsselprinzip: Das Störpotential simultaner Aktivitäten bemisst sich daran, in welcher Art und Weise Interaktanten Probleme der interaktionalen Fortsetzung bearbeiten.

Aus den geschilderten Gründen adressieren wir nachfolgend – in Fokussierung auf kinesische Simultanformen – die multimodale Verfahrensvielfalt der “overlap resolution” mit dem Ziel einer Systematisierung, die sich rückbindend an die hier geführte Diskussion der analytischen Ausführung anschließt. Die rekonstruierten Verfahren sind nicht als vollständige Beschreibung des kommunikativen Inventars dieser Praktik zu verstehen, sondern als ein erster Versuch, die Verfahrensvielfalt zu systematisieren, um einen Ansatz zur Präzisierung des Gegenstandes für die Folgeforschung zu eröffnen.

2.3.6 Phänomenologische Differenzierung: Konfligierende & funktionale Simultaneität

Das nachfolgende Eingangsbeispiel veranschaulicht eine empirisch begründbare Differenzierung in zwei kontrastierende Typen von Simultanaktivitäten:

- Simultanaktivitäten, die in der Bearbeitung interaktionaler Aufgaben funktional sind
- Simultanaktivitäten, die ein Störpotential in der Bearbeitung interaktionaler Aufgaben bilden

Die Analyse zeigt, dass in den Verfahrensgruppen Fortsetzungsprobleme unterschiedlich gelöst werden. In funktionalen Simultanaktivitäten sind Verfahren der situativen Desambiguierung beobachtbar, die instantan – teils durch multimodale Verdichtung – potentielle Fortsetzungshandlungen zugunsten einer faktischen Fortsetzungsaktivität elidieren. Konfligierende Simultanaktivitäten hingegen konstituieren selbst eine Situation ambiger Fortsetzungsmöglichkeiten. Erst die Reparaturverfahren konfligierender Simultaneität haben eine situationsdesambiguierende Funktion: Simultanaktivitäten werden entweder suspendiert oder zugunsten des gegenwärtig emergenten Handlungsvollzugs elidiert. Beide Verfahrensgruppen sind dabei eng an die interdependente Beobachtung der Handlungsaktivitäten beider Partizipanten gekoppelt.

Die nachfolgende Differenzierungsreihenfolge richtet sich nicht nach der empirischen Häufigkeit ihres Vorkommens, sondern ist der sequentiellen Emergenz im Fallbeispiel geschuldet.

³⁶ Hinweise hierzu finden sich auch bei SSJ (1974: 729f.) im Rahmen der Diskussion zu “differential turn-taking systems” und Anpassungen der Allokationstechniken in verschiedenen Kommunikationsformen (“debates”, “meetings witg chair-persons”, etc.).

Konfligierende Simultaneität:

Gestische Sequentialisierung – Handlungsrektion bei divergenten Teilnehmerinterpretationen

Im Vorfeld des nachfolgend dargestellten Transkriptausschnitts haben sich die Interaktanten auf die Platzierung einer Quadstrecke verständigt, ohne sich auf eine konkrete Route einigen zu können. Die Entscheidung über den Streckenverlauf wurde dabei vertagt. Der Beginn des Transkripts markiert nun einen Zeitpunkt, in dem die Frage der Fortführung ungeklärt ist, da keiner der beiden Interaktanten ein neues Verhandlungsobjekt zur Platzierung vorgeschlagen hat. Während dieser Transitionsphase inszeniert VP-A mit „und äh: ich würde sie doch SEHR bitten beim SIE zu bleiben;“ eine Distanzierung, die ironisierend das aufoktroierte Rollenverhältnis reproduziert (01). Entscheidend für die nachfolgend ausgeführte Analyse ist, dass A’s inszenierte Mahnung einen Zeitpunkt im Diskurs markiert, in dem beide Interaktanten durch die Zuwendung auf die noch zu verhandelnden Objekte im “stack” (02) auf eine Beendigung der Transitionsphase drängen – sie kehren sprichwörtlich zum „Ernst der Lage“ zurück.

FRAGMENT 4:

F2F-Baseline:	VP-Gruppe 08 (VP-A: Naturschützer / VP-B: Investor)
Zeit:	19:37 - 19:43

```

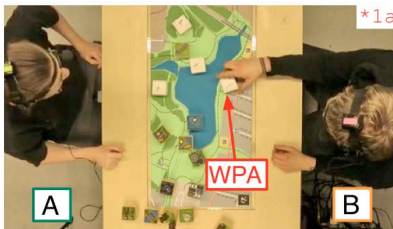
01 A-ver: und äh: ich würde sie doch SEHR bitten beim SIE zu bleiben;
    A-gaz: | >>>B(face).....>>>map...

02 B-ver: | ((lacht)).....| |JA;
    B-gaz: | >>>A(face).....>>>stack.....>>>WPA.....
    A-gaz: |.....| >>>stack.....>>>WPA...

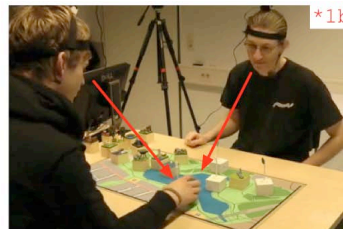
03 B-ver: |äh=ach das war hier WASSERSchutzgebiet-
    B-act: | ((greift das Objekt WPA, dreht und verschiebt es leicht))
    B-gaz: |@WPA.....
    A-gaz: |@WPA.....

```

*1a+b



Topview
19:37.64



Fokus auf A
19:37.64

Vor diesem Hintergrund bildet B’s Aufmerksamkeitsverlagerung vom “stack” auf das Objekt “Water Protection Area” (WPA) (02) eine für VP-A relevante Orientierungshandlung, die er seinerseits mit der Verlagerung der Aufmerksamkeit auf das kontextuell relevant gesetzte Objekt bearbeitet (02). Zu diesem Zeitpunkt sind beide Partizipanten ko-orientiert. Sequentiell nachfolgend dreht B das Objekt, um die Aufschrift des Schildes lesen zu können und identifiziert es dann mit „äh=ach das war hier WASSERSchutzgebiet-“ (03+*1a+b).

Aus analytischer Sicht ist nicht eindeutig entscheidbar, ob B’s Orientierung auf das Objekt WPA der alleinigen Objektidentifikation dient oder als “pre-proposal account”³⁷ (Houtkoop-Steenstra 1990) eine

³⁷ Eine Objektidentifikation als “pre-proposal account” (Houtkoop-Steenstra 1990) stellt in unseren Daten ein relativ häufiges Verfahren dar. Insbesondere in der AR-Bedingung (beispielhaft VP-Gruppe 05) werden verbalsprachlich formulierte Probleme der Objektidentifikation strategisch genutzt, indem die Ko-Orientierung auf das zu identifizierende Objekt als Ausgangspunkt für einen sequentiell nachfolgenden Platzierungsvorschlag verwendet wird.

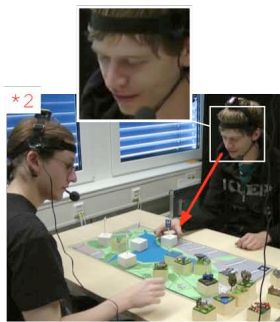
ankündigende Handlung zum Objektvorschlag darstellt, bzw. die Möglichkeit einer Anschlusshandlung offen lässt. Die Aktivitäten der Interaktanten deuten auf divergente Teilnehmerinterpretationen: A's Orientierungsverlagerung zurück zum "stack" (04+*2) lässt auf eine Teilnehmerinterpretation schließen, die B's Handlung als „rekapitulierende Objektidentifikation“ ohne weitere Handlungsimplikationen auslegt. B's fortwährende Objektfokussierung und die am Objekt verweilende Greifhand hingegen deuten auf eine potentielle Anschlusshandlung (04+*2). Diese Teilnehmerinterpretation manifestiert sich zusätzlich in der Beobachtung, dass B die Objekt-umschließende-Greifhand erst reaktiv – infolge der von A zum "stack" geführten Hand – in ihre "home position" zurückführt (04+*3+*4a). Im Hinblick auf die ostentativ ausgeführte multimodale Orientierungsgestalt A's (*3), die den faktischen Beginn einer neuen Aufgabenbearbeitung antizipierbar macht, bildet die simultane Handlungsaktivität B's einen potentiellen Störfaktor, der für die kontemporäre Fortsetzung der Interaktion dysfunktional ist. Infolgedessen bearbeitet B die emergente Störung prospektiv durch ein Rückzugsangebot und der Fokussierung der Aufgabeninitiierung A's (*4a).

Zusammenfassend verweist die strukturelle Ordnung des Verfahrens auf einen Typus von Simultanaktivität, bei dem zwei emergierende Orientierungsgestalten konfligierend aufeinander treffen, einer der beiden Akteure das Störpotential erkennt und selbstinitiiert repariert, indem die eigene Simultanaktivität in die sequentielle Ordnungsstruktur der emergenten Handlungsaktivität des Ko-Partizipanten überführt wird.

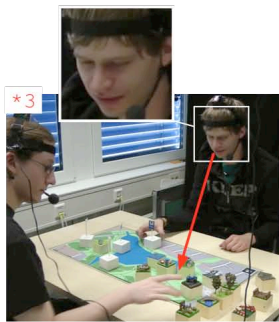
```

04 A-ver:                |<<p>na=gut äh:> hier HINTEN können wa ja
A-gaz:   >>>|stack(BR)...|>>>loc1.....
A-rH:    |grasp>>>@BR.....lift>>>|
A-lH:                                |grasp@BR.....place>>>>>>
B-gaz:   .....>>>rH(A).....>>>map.....>>>lH(A)>>>>>>>
B-rH:    @WPA.....>>>homeposition.....
                   *2  *3          *4a+b  *5          *6

05 A-ver: ne RICHTig schöne brücke noch machen;|
A-gaz:   .....|
A-lH:    >>>>>>>>>>>>>>>@loc1.....|
B-gaz:   >>>>>>>>>>>>>>>@loc1.....|
                                   *7
  
```



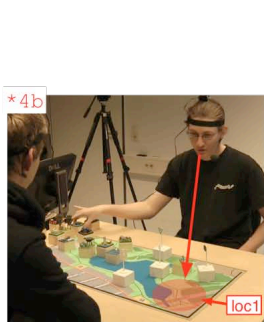
Fokus auf B
19:38.16



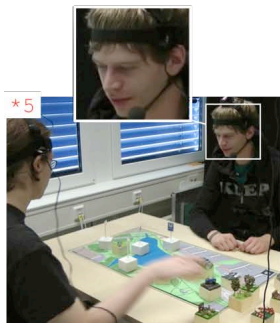
Fokus auf B
19:38.52



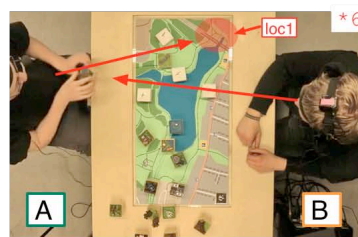
Fokus auf B
19:39.36



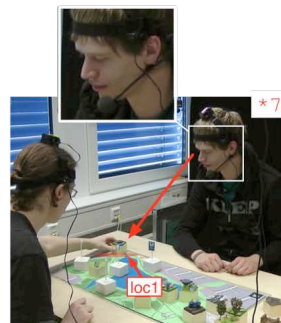
Fokus auf A
19:39.36



Fokus auf B
19:39.88



Topview
19:40.76



Fokus auf B
19:42.60

Funktionale Simultaneität:

Instantane situative Desambiguierung durch multimodale Verdichtung

Neben dem konfligierenden Simultaneitätsformat und seiner Bearbeitung durch die gestische Handlungsretraktion B's lässt sich im vorliegenden Beispiel (Fragment 4 ab Transkriptzeile 04) eine komplexe Orientierungsgestalt beobachten, in der VP-A verschiedene Ausdrucksressourcen einsetzt, die in einer Phase der Simultaneität konvergieren. Die Ressourcen erfüllen dabei funktional unterschiedliche Aufgaben: A's zum Objekt "Bridge" (BR) gewandte Körperpositur sowie die zum Objekt geführte Greifhand bedingen eine Herstellung intersubjektiver Wahrnehmung auf das Referenzobjekt (*3+*4a). Kurz nachdem A physischen Kontakt zum Objekt hat, wendet er seinen Blick über die Karte zu einem spezifischen Lokalisationsort (*4a+b, „loc1“). Diese Aufmerksamkeitsverlagerung impliziert eine Wahrnehmung der Aufmerksamkeitsorientierung des Ko-Partizipienten auf den gegenwärtigen Handlungsvollzug (A's Sichtfeld „kreuzt“ den Ko-Partizipienten). Erst nach Sicherstellung der Wahrnehmungswahrnehmung markiert A den Beginn der Handlungsaktivität verbalsprachlich mit „<p>na=gut>“ (04).

Die physische Manipulation des Objekts und die Fokussierung auf den Lokalisationsort konstituieren eine neue semiotische Einheit, deren Bedeutungserweiterung in der räumlichen und diskursiven Relation zwischen dem Referenzobjekt und seinem Lokalisationsort besteht (*4b). Damit ist zu diesem Zeitpunkt erwartbar, dass der Objektauswahl weder ein Objektidentifikationsproblem noch ein ausgiebiger "pre-proposal account" (Houtkoop-Steenstra 1990) folgen, sondern sich ein unmittelbarer Vorschlag zur Objektplatzierung anschließt. Insofern desambiguiert die Aufteilung des Orientierungsangebots die situative Ambiguität, da eine erwartbare Anschlusshandlung instantan konzeptionalisiert wird. A's Aufmerksamkeitsorientierung auf „loc1“ bildet somit ein relevantes Orientierungsangebot, das B durch eine Orientierungsverlagerung seinerseits bearbeitet, noch bevor die kommunikative Handlung verbalsprachlich als Platzierungsvorschlag erkennbar wird (*5).

Im weiteren Interaktionsverlauf wird der „multimodale Verdichtungsraum“ aufgelöst: A übergibt das „Referenzobjekt“ von der rechten in die linke Hand und stellt so eine proximale Konfiguration her, in der alle verfügbaren Ausdrucksressourcen (Positur, Blick, Referenzobjekt) vektoriell auf den „Lokalisationsort“ gerichtet sind (*6). B reagiert auf die Konkretisierung der emergenten Platzierungshandlung mit der Aufmerksamkeitsverlagerung von dem für ihn unspezifischen Lokalisationsort auf das zum Vorschlag präparierte Objekt (*6) und folgt anschließend der Platzierung auf den spezifischen Lokalisationsort (*7), während A jetzt auch verbalsprachlich den Handlungsvollzug als Vorschlag zur Objektplatzierung markiert („hier HINTEN können wa ja ne RICHTig schöne brücke machen;“ 04+05).

Zusammenfassend verweist die analytische Ausführung auf einen Simultaneitätstypus, der ein funktionales Orientierungsformat für die kontemporäre Aufgabenbearbeitung bereitstellt. Die Funktionalität besteht darin, dass erwartbare nächste Handlungsschritte des spezifischen Bearbeitungsformats „Vorschlag zur Objektplatzierung“ (Objektidentifikation, Objektidentifikationsproblem, "pre-proposal account") zugunsten einer instantan vorgeschlagenen Objektplatzierung elidiert werden. Das Ablaufformat wird verdichtet, so dass der Zweck der Objektauswahl in situ – durch die Herstellung einer Relation zwischen Referenzobjekt und Lokalisationsort – projiziert wird. Dies zeigt im Besonderen, dass McNeills (1992) Regeln zur „semantischen und pragmatischen Synchronie“ keine universellen Prinzipien darstellen. Im vorliegenden Fall treten zwar Sprache und Gestik koinzident auf, dienen jedoch weder der Übermittlung derselben Bedeutungseinheit ("idea unit") („Semantische Synchronie“) noch erfüllen sie dieselbe pragmatische Funktion („Pragmatische Synchronie“). Während A's rechtshändige Objektauswahl eine Ko-Orientierungsfunktion einnimmt, dient die Verbalsprache in diesem Moment der Objektlokalisierung („<p>na=gut äh:> hier HINTEN können wa ja“).

Im Hinblick auf argumentationstheoretische Überlegungen stellt ein verdichtetes Ablaufformat im Kontext von Vorschlagshandlungen einen resoluten Handlungsvollzug dar, der anders als die präparativen Vorschlagskonstitutionen erst nach dem Vollzug – bei uns der Objektplatzierung – handlungsimplicative Wirkungen hat. Dem Ko-Partizipienten wird somit eine Einflussnahme während der Handlungsdurchführung verwehrt.

Die zeitliche Verdichtung konstituiert zudem ein spezifisches Anforderungsprofil an die monitoring-Aktivitäten des Ko-Partizipienten, der die komplexe Orientierungsgestalt als instantanen Handlungsvollzug zur Objektplatzierung deuten muss. Dabei ist feststellbar, dass die intrapersonelle Simultanaktivität in

einer sequentiellen Lesart enkodiert ist und in dieser Lesart vom Ko-Partizipanten dekodiert werden muss, um den holistischen Handlungsvollzug dem „intentionalen Zweck“ entsprechend interpretieren zu können. Sequenzen aus der AR-Bedingung verweisen in diesem Zusammenhang auf die Entstehung von divergenten Teilnehmerinterpretationen, wenn die komplexe Orientierungsgestalt aufgrund der AR-spezifischen Sichtfeldeinschränkung nicht in ihrer sequentiellen Lesart interpretiert werden kann (vgl. 2.4.1).

2.3.7 Bearbeitung konfligierender Simultanaktivitäten

Auf Basis der in 2.3.6 vorgenommenen Differenzierung in funktionale und konfligierende Simultanaktivitäten lassen sich empirische Phänomene der face-to-face Bedingung typologisieren. Wir beginnen nachfolgend mit einer Ausdifferenzierung des konfligierenden Typus, ohne die Bearbeitungsreihenfolge einer Priorisierung unterstellen zu wollen. Simultanaktivitäten, die ein Störpotential in der Bearbeitung interaktionaler Aufgaben bilden, sind insofern konzeptuell interessant, da unser Versuchsdesign Fälle evoziert, in denen visuelle Ereignisse auf gleicher Modalitätsebene konkurrieren (vgl. Schmitt 2005: 47f.).

Die in unseren Daten empirisch rekonstruierbaren Bearbeitungsverfahren konfligierender Simultaneität lassen sich phänomenologisch in drei Gruppen einteilen, in denen wiederum zum Teil strukturell unterschiedliche Realisierungsvarianten auf Fallebene aufgezeigt werden können. Die Gliederung dieses Unterkapitels richtet sich nach der phänomenologischen Differenzierung der Verfahren in (1) Formen gestischer Sequentialisierung (2.3.7.1), (2) Sequentialisierung durch Direktiva (2.3.7.2), und (3) Sequentialisierung durch Blickkoordination (2.3.7.3).

2.3.7.1 Formen gestischer Sequentialisierung

In Bezug auf unser Eingangsbeispiel zur phänomenologischen Differenzierung von Simultanaktivitäten präsentieren wir nachfolgend strukturelle Ausgestaltungen der Verfahrensgruppe „gestische Sequentialisierung“ (vgl. Fragment 4, 2.3.6). Der verwendete Terminus dient der Typologisierung und soll nicht auf monoaktivistische Verfahren hindeuten. Die Bezeichnung ist als Charakterisierung des primären Handlungsvollzugs zu verstehen, der durch das Zusammenwirken multimodaler Ressourcen konstituiert wird.

Handlungssuspension bei divergenten Teilnehmerinterpretationen

Das nachfolgende Fallbeispiel zeigt, wie Simultanaktivitäten mit Störpotential in Phasen emergieren, in denen die Interaktanten den gegenwärtigen Handlungsvollzug unterschiedlich interpretieren. Eine „störende Simultanaktivität“ ist dabei in der Regel nicht auf ein individuelles „Fehlverhalten“ zurückzuführen, sondern emergiert im interaktiven Zusammenwirken interpersoneller Handlungsaktivitäten. Eine Besonderheit des vorliegenden Falls (Fragment 5) besteht darin, dass die konfligierenden Simultanaktivitäten (interpersonell) aus der divergenten Teilnehmerinterpretation einer funktionalen Simultanaktivität (intrapersonell) hervorgehen.

Während in Fragment 4 die störende Simultanaktivität vollständig elidiert wurde, verweist das nachfolgende Fragment 5 auf einen anderen Typus: Simultanaktivitäten, die im Moment des konfligierenden Aufeinandertreffens ein Störpotential für den Interaktionsfortgang bilden, in der holistischen Aufgabebearbeitung jedoch funktional sind, da sie ein Ablaufformat nächster Handlungsschritte generieren. Im nachfolgenden Beispiel wird die Simultaneität durch eine Form der gestischen Sequentialisierung aufgelöst: VP-A unterbricht den gestischen Handlungsvollzug der Fortsetzungsaktivität, bearbeitet die Handlungsimplicationen eines alternativen Platzierungsvorschlags seines Ko-Partizipanten und kehrt dann zur suspendierten Handlungsaktivität zurück.

Noch vor dem dargestellten Transkriptausschnitt argumentiert VP-B (Naturschützer) ab Minute 10:04 für eine alternative Platzierung des Quadparks (QP) und stellt das Objekt auf eine Fläche, die er als „Freifläche/Parkplatz“ interpretiert. VP-A erwidert, dass es sich nicht um eine freie Fläche, sondern um ein Wohngebiet handle, das nicht bebaut werden dürfe. Während dieses Einwands nimmt er das Objekt von der Wohngebietsfläche, stellt es vor sich auf der Karte ab und zeigt VP-B, wo sich die Parkflächen auf der

Karte befinden. Der Beginn des Transkripts markiert den Augenblick, in dem A zu der provisorischen Platzierung des Quadparks zurückkehrt.

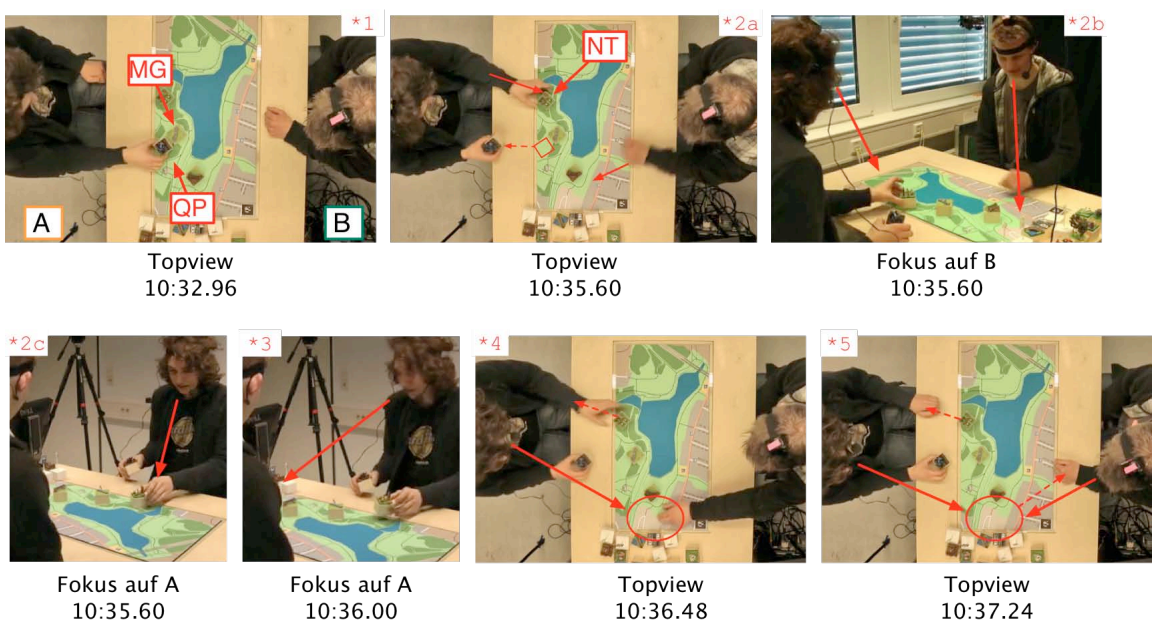
FRAGMENT 5:

F2F-Baseline:	VP-Gruppe 03 (VP-A: Investor / VP-B: Naturschützer)
Zeit:	10:32 - 10:55

```

01 A-ver: |und HIER direkt neben der minigolfanlage is ja auch kacke;|
      A-rH: |@QP...*.>>>>>>pos-1.....|
                *1

02 B-ver: |aber guck mal was ham wir denn HIER,=zum beispiel|
      B-act: |dG>>>>>>>>>>>>>>>>>>>>>>>>>@loc>>>>>>>>>>HP|
      A-rH: |>>>pos-2.....|
      A-lH: |grasp>>>>@NT..lift>>>>place>>>>>>>>>>>>>>>HP....|
      A-gaz: |@NT.....>>>dG(B)..@loc.....|
      B-gaz: |@map(loc).....@loc.....|
                *2a+b+c      *3      *4      *5
  
```



```

03 A-ver: das=sieht=das sieht NICHT so aus als könnte man DIEse fläche
04 A-ver: nutzen=das is wahrscheinlich i=irgendwie nen gewerbegebiet;
  
```

Die provisorische Platzierung wird sowohl verbalsprachlich, durch den Hinweis der Unvereinbarkeit der Objekte Quadpark (QP) und Minigolfanlage (MG), als auch physisch, durch die Wiederaufnahme und Rückführung des Objekts zum Kartenrand, aufgelöst (01, *1, *2a). Für sich genommen bildet diese multimodale Gestalt syntaktisch, prosodisch und performativ (im Sinne des multimodalen Handlungsvollzugs) eine geschlossene Äußerungseinheit, auf die B mit der turn-Übernahme in 02 reagiert.

Betrachten wir hingegen A's holistische (beidhändige) Äußerungsgestalt, wird deutlich, dass A eine turn-Fortsetzung durch die linkshändige Präparation des Objektes Naturlehrpfad (NT) konstituiert (*2a-c). Folglich handelt es sich hier um eine zweigliedrige multimodale Äußerungsgestalt, in der aufgabenteilig einerseits die provisorische Objektplatzierung aufgelöst wird (rechte Hand), andererseits – in partieller Gleichzeitigkeit anschließend – eine Fortsetzungsmarkierung erfolgt (linke Hand), die den Objektrückzug als „Mittel zum Zweck“ ausweist.

Auch wenn es sich im vorliegenden Fall um eine temporale Abfolge aus „Objektretraktion“ (rechte Hand) und „Objektmanipulation“ (linke Hand) handelt, überlagern sich beide Ausdrucksressourcen in

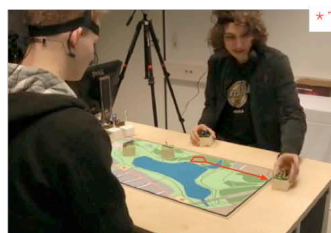
Eine methodologische Relevanz der arbeitsteiligen Gestaltform ist aus B's reaktiver Handlungsaktivität ableitbar. Er reagiert nicht auf die ganzheitliche, aufgabenteilige Äußerungsgestalt im multimodalen Verdichtungsraum, sondern bereits nach der Rückführung des Objekts zum Kartenrand, in partieller Simultaneität zu A's linkshändiger Objektmanipulation (*2a-c, 02). Die holistische Äußerungsgestalt, in der die Objektmanipulation die semiotische Signalisierungsfunktion der Objektretaktion (*2a-c) verändert (s.o.), wird somit nur teilrezipiert und nicht ihrer sequentiellen Lesart entsprechend interpretiert. Die konversationsanalytische Annahme einer holistischen Produktion und Rezeption multimodaler Äußerungsgestalten (z.B. Goodwin 2000a, 2003) ist infolgedessen – für den hier geschilderten Fall – zu vage und verweist auf die Notwendigkeit, interaktionale Effekte teilrezipierter Äußerungsgestalten systematisch zu erforschen, um die Axiomatisierung der konversationsanalytischen Multimodalitätstheorie zu präzisieren (s. Kapitel 3). Dabei stellt sich die Frage, ob die holistische Rezeption multimodaler Äußerungsgestalten eine Grundvoraussetzung zur Interaktionsfähigkeit der Interaktanten darstellt oder ob die Emergenz von Anschluss-handlungen, die Möglichkeiten der interaktionalen Fortsetzung bereitstellen, das prädominante Kooperations- und Ordnungsprinzip bildet.

Im vorliegenden Fall führen die divergenten Teilnehmerinterpretationen zu einer konfligierenden Simultaneität, die jedoch in ein für die gesprächsorganisatorische Fortsetzung funktionales Ablaufformat gewendet wird. In diesem Sinne ist die Divergenz der Teilnehmerinterpretationen unproblematisch, da die Handlungsaktivität B's von A sinnhaft als parenthetischer Eingriff ins Handlungsgeschehen gedeutet und in den eigenen Handlungsablauf integriert wird. An dieser Stelle ist beobachtbar, dass A die linkshändige Fortsetzungsaktivität (*2a-c) während der konfligierenden Simultanaktivität B's abbricht, indem er zuerst den Aufmerksamkeitsfokus auf den Handlungsvollzug seines Ko-Partizipanten verlagert (*3) und sequentiell nachfolgend die Hand vom Manipulationsobjekt zurückführt (*4+*5). Erst nachdem B seinen alternativen Platzierungsvorschlag „aber guck mal was ham wir denn HIER,=zum beispiel“ (02) verbalisiert und die Verweishand in ihre "home position" zurückgeführt hat (*5), reagiert A verbalsprachlich mit einer Vorschlagsablehnung, die argumentativ den markierten Lokalisationsort als nicht bebaubar ausweist (03+04).

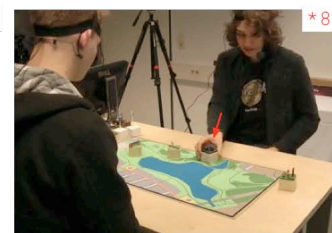
- 05 A-ver: |°hh |und AUßERdem- eine=sehen sie doch mal-|das is
 A-act: |grasp>>>|NT.....|>>>HP
- 06 A-ver: nen WANDERweg für äh: RENTner; wir wollen HIER aber eine
- 07 A-ver: !SPORT!Anlage raus machen |da wollen wir nich irgendwelche
 A-act: |((nimmt NT von der Karte und
 *6
- 08 A-ver: rentner in ihrem wald spazieren gehen=sondern wir machen
 A-act: platziert es am Kartenrand + platziert QP auf dem einstigen
 *7 *8



Topview
10:51.32



Fokus auf A
10:52.36



Fokus auf A
10:53.32

- 09 A-ver: den wald weg und bauen da ne QUADstrecke-
 A-act: Platzierungsort von NT))

Nach der Vorschlagsablehnung kehrt A zum Objekt „Naturlehrpfad“ (NT) zurück und finalisiert die in 01 und 02 präparierte Fortsetzungshandlung. Die Rückkehr zur suspendierten Fortsetzungshandlung wird dabei sowohl durch die kinesische Wiederaufnahme des Objekts NT als auch durch die verbalsprachliche

Konnektiveinheit „und AUßERdem-“ markiert (05). Zusammen mit dem nachfolgenden Textelement „sehen sie doch mal-“ wird eine Äußerungseinheit konstituiert, die den parenthetischen Handlungseingriff B's funktional in die argumentative Struktur des eigenen Handlungsvollzugs einbindet: B's Platzierungsvorschlag wird hier nicht nur im ersten Schritt als impraktikabel ausgewiesen („nicht bebaubare Fläche“ (03+04)), sondern darüber hinaus – durch die nachdrückliche Relevanzsteigerung des in 05 angekündigten Sachverhalts – im zweiten Schritt als vermindert relevant markiert. Die Folgehandlung A's ab Zeile 06 lässt dabei erkennen, dass es ihm nicht nur um eine alternative Platzierung des Quadparks geht, sondern darum, den Naturlehrpfad (NT) aus der Neukonzeption des Naherholungsgebiets zu streichen. Infolgedessen nimmt er das Objekt NT von der Karte (*6+*7) und platziert den Quadpark genau dort, wo sich zuvor der Naturlehrpfad befand (*8).

Handlungssuspension bei aufeinandertreffenden Gestenvektoren

Eine strukturell ähnliche Form der gestischen Sequentialisierung von konfligierenden Simultanaktivitäten zeigt Fragment 6. Im Unterschied zu Fragment 5 ist die Emergenz der aufeinandertreffenden Handlungsaktivitäten jedoch nicht auf divergente Teilnehmerinterpretationen einer funktionalen Simultanaktivität zurückzuführen, sondern auf eine sequentiell nachfolgende turn-Expansion, die zeitlich mit dem Versuch der Rederechtsübernahme des Ko-Partizipienten in der “self selection technique ‘starting first’” (Sacks, Schegloff & Jefferson 1974: 718) zusammenfällt. Das für diesen Mechanismus beschriebene Prinzip “first starter gets the turn” (ebd.) scheint jedoch auf der Ebene multidimensionaler Handlungsaktivitäten als organisatorische Strukturform weniger bindend als es für rein verbalsprachliche Koordinierungsprozesse angenommen wurde. Während in den beiden vorigen Beispielen (Fragment 4 & Fragment 5) der verbalsprachliche “first starter” tatsächlich den turn übernimmt und der jeweilige Ko-Partizipient die in der Situation simultan ausgeführte gestische Fortsetzungsaktivität elidiert oder suspendiert, ist es im nachfolgenden Beispiel umgekehrt: Der verbalsprachliche “first starter” bricht das Äußerungsformat ab und lässt seinen Ko-Partizipienten gewähren.

Unsere Analysen deuten darauf hin, dass die verbalsprachlich explizierte Form “first starter gets the turn” in multimodalen Handlungssituationen deutlich komplexer ist und nicht allein durch eine situative Prädominanz verbalsprachlicher Einheiten konstituiert wird. So verweisen die beiden vorigen Beispiele (Fragment 4 & Fragment 5) auf weitere Determinanten, wie z.B. die situative Beurteilung, inwieweit die emergente Handlungsaktivität des Ko-Partizipienten eine ostentative Gestaltform konstituiert, die eine faktische Fortsetzung der Aufgabenbearbeitung erkennbar macht (v.a. Fragment 4). Umgekehrt scheint jedoch auch das self-monitoring und die Einschätzung der interaktiven Wirksamkeit eigener Handlungsaktivitäten maßgeblich zu sein. In den beiden vorigen Fragmenten waren die Fortsetzungsaktivitäten in ihrer figurativen Gestaltform ambig und von verminderter Zielorientiertheit, was zu divergenten Teilnehmerinterpretationen und konfligierenden Simultanaktivitäten führte. Eine Beurteilung des eigenen Äußerungsformats in Bezug auf die ostentative Figurierung und Interpretierbarkeit der Gestaltform mag folglich auch dazu beitragen, ob dem Ko-Partizipienten in der jeweiligen Situation Handlungsfreiheit gewährt und die eigene Aktivität zurückgenommen wird.

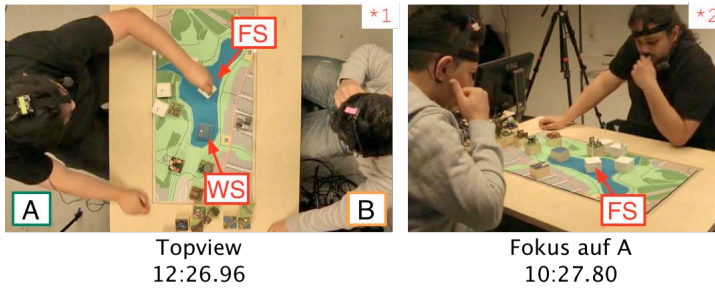
Im nachfolgenden Fragment schließt sich ein weiteres Ordnungsprinzip an: Die gestische Sequentialisierung bei aufeinandertreffenden Gestenvektoren. Gestische Aktivitäten, die sich in einem gemeinsamen Handlungsraum aufeinander zubewegen, bilden dabei keine universale Konstitutionsform konfligierender Simultanaktivitäten. Vielmehr sind sie stark kontextsensitiv und von der situativen Konfiguration des Interaktionssystems abhängig. In Phasen, in denen die Interaktanten dynamisch einzelne Objekte manipulieren, sich uneins über einen geeigneten Platzierungsort sind oder spielerisch inszenierend die Wechselseitigkeit der konträren Rollenverhältnisse reproduzieren, bildet eine trajektorische Konvergenz oder ein Überlappen gestischer Aktivitäten eine handlungskonforme Beteiligungsweise. In Phasen hingegen, in denen strikt sequentielle Handlungsabläufe ausgehandelt und praktiziert werden, konstituieren aufeinandertreffende Gestenvektoren ein unerwartetes Störpotential, das von den Interaktanten in ein sequentielles Ablaufformat transformiert wird. Das vorliegende Beispiel ist kontextuell der letzteren Bedingung zuzuordnen.

FRAGMENT 6:

F2F-Baseline:	VP-Gruppe 05 (VP-A: Naturschützer / VP-B: Investor)
Zeit:	12:22 - 12:37

01 A-ver: |DAS: (--) könnte man NEHMEN wenn DAFÜR auf JEDENfall
 A-act: |((nimmt das Objekt FS aus dem stack und platziert es

02 A-ver: der bereich hier geschützt ist;|
 A-act: mittig auf dem See)).....>>>HP|
*1 *2



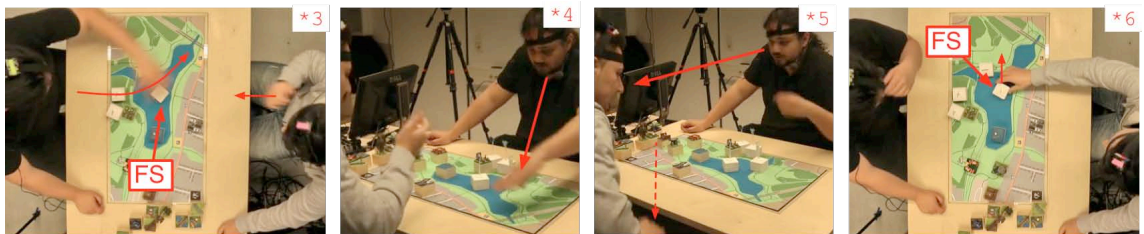
Topview
12:26.96

Fokus auf A
10:27.80

Der Ausgangspunkt für die konfligierenden Teilnehmeraktivitäten liegt in einer Situation, in der VP-A einen Kompromiss einleitet, der sich auf eine im Vorfeld situierte Vorschlagshandlung seines Ko-Partizipanten zur Platzierung eines Wasserskigebietes (WS) bezieht. Die Platzierung des Wasserskigebietes wird dabei unter dem Vorbehalt ratifiziert, einen anderen Teil des Sees zu schützen (01+02), der verbalsprachlich als „der bereich hier“ ausgewiesen und symbolisch durch die Platzierung des Objekts “Fish-Sign” (FS) markiert wird (*1). A’s konzessive Vorschlagshandlung wird durch die gestische Retraction der ausführenden Verweishand beendet (*2).

Wie in allen Versuchspersonengruppen der face-to-face Bedingung werden die Objekte als symbolische Repräsentanten eines Planungskonzepts verwendet, dessen räumliche Dimension durch deiktische Handlungen am Phantasma ausgehandelt wird. Insofern bildet die deiktische Einheit aus verbalsprachlicher Markierung und physischer Objektplatzierung speziell im vorliegenden Fall, in dem sich der Kompromiss auf abzugrenzende Areale bezieht (FS vs. WS), eine unterspezifizierte Verweisform. A’s konzessive Vorschlagshandlung eröffnet damit Handlungsimplicationen, die zur Ausdifferenzierung der Areale animieren.

03 A-act: dg>>>>>>|((@Bereich))>>>HP|
 A-ver: |<<all>der ganze->|
 B-ver: |ja ich mein soweit=joa| |soweit würden die menschen|
 B-act: |RH>>>map>>>HP.....|grasp>>>|@FS...>>>dG@loc>>>>>>>>HP|
*3 *4 *5 *6



Topview
12:28.16

Fokus auf A
10:28.32

Fokus auf A
10:28.76

Topview
12:30.44

04 A-ver: |^mhm; |^MHM;
 B-ver: |auch nicht RAUSSchwimmen und (-) wasserski fahren|

05 A-ver: das wäre |soweit noch in ordnung;|
 B-act: |grasp>>>>>>>>>>>>>>>>>>@H|

In den sich teilsimultan überlappenden Fortsetzungsaktivitäten beider Interaktanten ab Zeile 03 werden die im Präkontext verankerten Handlungsimplicationen (01+02) in unterschiedlicher Weise eingelöst. So ist es einerseits VP-A selbst, der eine Spezifikation des zu schützenden Areals nachholt, indem er die dimensionale Ausdehnung gestisch skizziert (*3+*4) und verbalsprachlich mit „<<all>der ganze->“ expliziert. Andererseits, in unmittelbar sequentiell nachfolgender Teilsimultaneität (200ms nach A's Gestenonset), reagiert auch sein Ko-Partizipant VP-B. In diesem Augenblick laufen die Gestentrajektorien der emergenten Äußerungsgestalten gegenvektoriell zusammen (*3). In Reaktion darauf bricht B als verbalsprachlicher “first starter” seinen turn „ja ich mein soweit=joa“ ab und retrahiert die zur Handlungsaktivität präparierte Gestenhand (*5).

Erst nachdem A seine Hand aus dem Verweisraum zurückzieht und durch Blickadressierung seinen Redebeitrag als terminierten Handlungsakt ausweist (*5) (vgl. Kendon 1967), reagiert B mit der Wiederaufnahme seiner suspendierten Handlungsaktivität. Er greift das Objekt “FishSign” (FS), vergrößert leicht den Abstand der konfliktären Objekte (FS+WS) und kommentiert A's konzessive Vorschlagshandlung mit „soweit würden die menschen auch nich RAUSschwimmen (-) und wasserski fahren“ (03+04, *6). Sequentiell folgend interpretiert A diese Handlung als ratifizierenden Akt und projiziert sein Verständnis ebenfalls ratifizierend mit „das wäre soweit noch in ordnung;“ (05). B manipuliert währenddessen bereits ein neues Verhandlungsobjekt (Hotel (H)) und bestätigt damit die Beendigung des Handlungsvollzugs.

Anmerkung

Das vorliegende Beispiel zur Handlungssuspension aufeinandertreffender Gestenvektoren bildet epistemisch einen besonders interessanten Fall, der sich zusammenfassend mit den folgenden phänomenologischen Eigenschaften klassifizieren lässt:

- Zum Zeitpunkt der parallelen Handlungsemergenz liegt kein eindeutiger Legitimationsanspruch auf das Rede-/Handlungsrecht vor.
- Die interpersonellen Simultanaktivitäten der Interaktanten setzen mit nur minimaler Verzögerung ein (vgl. auch Fragment 5), so dass sie unter einer emischen Perspektive als Form „echter Simultaneität“ begriffen werden können, auf die erst im Handlungsvollzug der sich überlappenden Aktivitäten eine Reaktion erfolgen kann.
- Die Handlungsaktivitäten der Teilnehmer konkurrieren
 - a) auf gleicher Modalitätsebene (Gestik)
 - b) auf gleicher Fokusebene (Abgrenzung verschiedener Bereiche (WS+FS)), und
 - c) auf gleicher Handlungsraumbene (der als See definierte Bereich)

Das Beispiel rekurriert in ideal- und prototypischer Weise auf die verbalsprachlich explizierte Form des “overlap” aufgrund der Projektion einer geschlossenen Turnkonstruktionseinheit (vgl. SSJ 1974: 706f. cf. §4.3-(b)). Im Hinblick auf die zeitliche Dimensionierung des Expansionsverzichts, die von der Parallelemergenz der Aktivitäten bis zur gestischen Retraktion ungefähr 600ms (vgl. screenshots *3-*5) umfasst, zeigt sich eine phänomenologische Äquivalenz zum Prinzip: “Occurrences of more than one speaker at a time are common, but brief” (Sacks, Schegloff & Jefferson 1974: 700). Ein analytisch schwer zu lösendes Problem erscheint uns dabei in der Aufdeckung der Regelmechanismen zum Expansionsverzichts zu liegen, da die Auflösung kinesischer Handlungsemergenz ungleich komplexer geregelt zu sein scheint als der für die selbstselektive Rederechtsübernahme dargelegte Mechanismus “first starter gets the turn”⁴⁰.

Ein methodisches Problem existiert in der auf den verschiedenen Ausdrucksdimensionen der Körperlichkeit beruhenden Variationsbreite zur Bearbeitung kinesischer Parallelemergenz. Während sich im Rahmen der Verbalität der “overlap resolution device” (Schegloff 2000, 2001) in der Konvergenz des Auditiven erschöpft, können kinesische Ausdrucksressourcen einerseits in sehr unterschiedlichen Konfigurationen assembliert werden (vgl. Goodwin 2000a, 2003: “contextual configuration”) und andererseits

⁴⁰ Der Mechanismus ist nicht als bindendes, deterministische Prinzip zu verstehen, sondern als „Regularität“, die auf einer empirisch beobachtbaren Häufigkeit beruht (vgl. SSJ 1974: 700 “grossly apparent facts” / “set of empirical constraints”)

durch einen Wechsel auf andere, in der Parallelaktivität nicht oder nur peripher beteiligte Modalitäten bearbeitet werden (vgl. 2.3.7.2 und 2.3.7.3 „Sequentialisierung durch Direktiva und Blickkoordination“). Die Schwierigkeit besteht infolgedessen darin, den Phänomenbereich exakt zu definieren und strukturell gleichwertige Fälle zu finden. Aufgrund der geschilderten phänomenologischen Eigenschaften und der figurativ klaren Gestaltform markiert das vorliegende Beispiel unserer Ansicht nach einen geeigneten Ausgangspunkt für die Folgeforschung, um die Regelung des Expansionsverzichts nicht nur deskriptiv zu beschreiben, sondern die zugrundeliegenden Ursachen zu eruieren. Eine systematische Fallkollektion homomorpher Instanzen jenes klar umrissenen Phänomenbereichs verspricht essentielle Erkenntnisse zu den organisatorischen Prinzipien zu liefern, wie derartige Situationen – abseits nicht kontrollierbarer Persönlichkeitsfaktoren wie Extro- und Introvertiertheit – aufgelöst werden. Als Ausgangspunkt lassen sich folgende potentiell relevante Prinzipien bestimmen, die in der Folgeforschung auf ihre empirische Validität überprüft werden könnten, um zu einem präzisen Verständnis der Regelmechanismen des Expansionsverzichts zu gelangen:

- “First come, first served” - Prinzip: Der kinesische “first starter” (im Beispiel: VP-A) erwirbt einen legitimen Anspruch auf das Rede- und Handlungsrecht.
- Das verbalsprachlich explizierte Prinzip “first starter gets the turn” wird außer Kraft gesetzt, wenn sich der Ko-Partizipient mit einer kinesischen Handlungsmodalität bereits als “first starter” etabliert hat: VP-B ist verbalsprachlicher “first starter”, VP-A hat jedoch bereits eine kinesische Handlungsaktivität initiiert.
- Der relevante Handlungsraum ist besetzt: Die kinesische Handlungspriorität VP-A’s ist nicht relevant, ausschlaggebend ist, dass VP-B seine Handlung nicht ausführen kann.
- Skalärer Unterschied der Gesten im Sinne des definitorischen Merkmals “effort” (vgl. Kendon 1980: 212; McNeill 1992: 83): A hat bereits einen höheren „Arbeitsaufwand“ betrieben als B.
- Es gibt einen „legitimere Sprecher“: A ist der legitimere Sprecher, da der gestische Verweis semantisch an seine konzessive Vorschlagshandlung anschließt. Die präkontextuelle Unterspezifikation weist ihn als legitimen Sprecher aus.

2.3.7.2 Sequentialisierung durch Direktiva

Eine weitere Verfahrensgruppe zur Lösung von Fortsetzungsproblemen stellen Formen der „Sequentialisierung durch Direktiva“ dar. In Abgrenzung zur sprechakttheoretischen Definition, betrachten wir nicht direktive Sprechakte, sondern Handlungsakte, die ein multimodal konstituiertes Anforderungsprofil an den Ko-Partizipienten stellen. Dabei ist das Fortsetzungsproblem in den folgenden Beispielen anders definiert als in der Verfahrensgruppe der „gestischen Sequentialisierung“ (vgl. 2.3.7.1). Die Frage der interaktionalen Fortsetzung ist nicht dadurch bestimmt, dass die emergente Handlungsparallelität systemisch nicht händelbar ist, sondern dadurch, dass einer der beiden Interaktanten ein individuelles Handlungsziel als bearbeitungsrelevante Interaktionsaufgabe statuiert. Es handelt sich im Folgenden also nicht um reaktive Verfahren im Sinne der Tilgung oder Suspension eigener Simultanaktivitäten, sondern um proaktive Verfahren, in denen die simultane Aktivität des Ko-Partizipienten strategisch-lenkend elidiert wird.

Die nachfolgenden drei Fragmente der Verfahrensgruppe entstammen derselben Versuchspersonengruppe (VP-03), liegen zeitlich eng beieinander und verweisen auf eine spezielle Interaktionsdynamik, in der die Sequentialisierungsverfahren durch Direktiva in ihrer Explizitheit zunehmen. Äquivalent zu den bisherigen Beispielen des Kapitels, erfolgt die analytische Rekonstruktion der Verfahren an dieser Stelle aus einer „mikrophänomenologischen Perspektive“ (vgl. Deppermann & Schmitt 2007: 38), in der die koordinativen Vorgänge des Sequentialisierungsverfahrens im Interessenfokus stehen. Zu einem späteren Zeitpunkt (vgl. Kapitel 4) stellen wir anhand der nachfolgenden Trias an Fällen zur „Sequentialisierung durch Direktiva“ einen makrostrukturellen Ansatz vor, der sich der analytischen Rekonstruktion der Phänomene in ihrem „handlungsschematischen Zusammenhang“ (ebd.) widmet. Unter einer makroskopischen Perspektive lässt sich zeigen, dass die empirisch beobachtbare interaktionale Dynamik nicht zufällig emergiert, sondern in der Interaktionsgeschichte der Beteiligten präkontextuell verankert ist. Ko-Orientierung ist dabei als bidirektionales Phänomen zu verstehen, in dem lokale Koordinierungsanforderungen der Orientierung auf handlungsschematische, globale Zusammenhänge bezogen werden, die selbst

wiederum kontextuelle Relevanzen für lokale Ko-Orientierungsphänomene bilden. Interaktanten stehen somit vor der Aufgabe, gemeinsame Orientierung auf beiden Ebenen zu koordinieren, um die lokale Ko-Orientierung auf spezifische Entitäten in ihrem globalen Sinnzusammenhang zu begreifen.

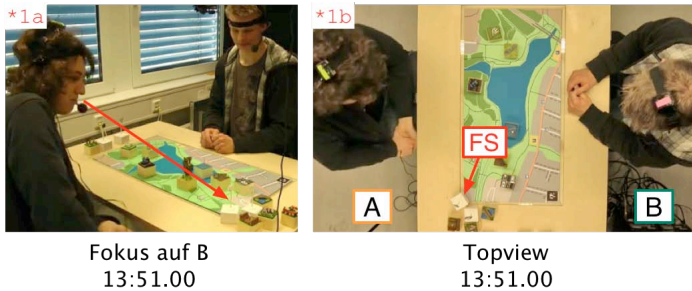
Konziliante Form: Frageadressierung zur Fortführung der gegenwärtigen Interaktionsaufgabe

Im Vorfeld (12:53-13:47) der im Transkript dargestellten Interaktionssituation platziert VP-A Objekte auf der Karte, die seinen Interessen als Investor entsprechen. Die Entwurfsplanung wird mit einer abschließenden Erläuterung rekapituliert, die mit der Bewertung der aufgestellten Konstellation schließt (01). Der Einstieg in die Interaktion markiert sowohl im Sinne semantischer wie auch syntaktisch-prosodischer Einheitenbildung eine Transitionsphase interaktionaler Aufgaben, in der unterschiedliche Fortsetzungsaktivitäten der Partizipanten beobachtbar sind.

FRAGMENT 7:

F2F-Baseline:	VP-Gruppe 03 (VP-A: Investor / VP-B: Naturschützer)
Zeit:	13:47 - 14:00

01 A-ver: |so; das wär so meine IDEALvorstellung glaub ich;
 A-gaz: |>>>>stack.....
*1a+b



Den Abschluss der Entwurfsplanung markiert VP-A durch eine multimodale Einheit aus verbalsprachlichem Diskursmarker „so,“ mit finalisierender Bewertung („das wär so meine IDEALvorstellung glaub ich;“) und nonverbaler Positur- sowie Aufmerksamkeitsverlagerung auf die verbleibenden Verhandlungsobjekte im “stack” (*1a+b).

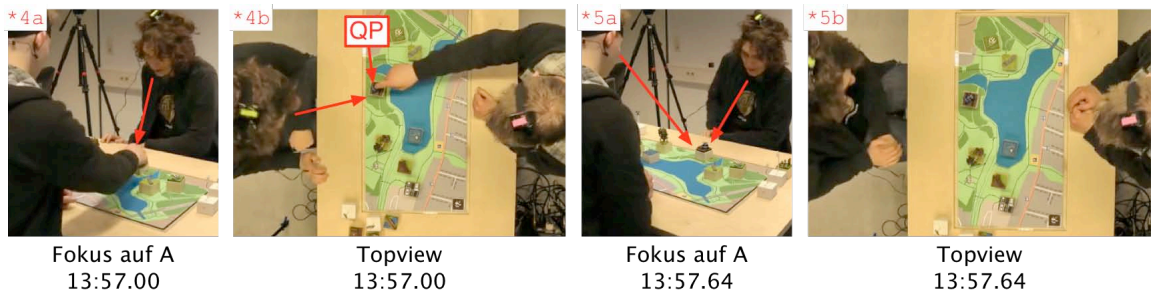
02 B-ver: ähm: | |meinen=das=was glauben sie was sie für
 | (1.2) |
 B-act: |dG~~~~map~~~~>>>>H.....
 A-act: |grasp>>>FS..lift..place>>>>>>>>>>>>>>>HP
 A-gaz:@FS.....>>>dG(B).....
*2a+b *3a+b



```

03 B-ver: ne perSONENgruppe hiermit ANsprechen, |
B-act: >>>>>>>QP.....>>>>>>>>>>>>>HP |
A-gaz: .....@QP..... |
          *4a+b                               *5a+b

```



Auf die turnübergaberelevante Stelle der terminierenden Aufgabenbearbeitung A's reagiert B, indem er sich verbalsprachlich durch die Hesitationsvokalisation "ähm:" (02) als "first starter" positioniert. In der Literatur sind Verzögerungspartikel unter anderem als "turn keeping device" diskutiert worden, die funktional der Sicherung der Sprecherrolle an turnübergaberelevanten Stellen dienen (vgl. Olbertz-Siitonen 2003: 7f.). Für Verzögerungspartikel wie „äh“ oder „ähm“ postuliert Olbertz-Siitonen ein Funktionscluster, in dem Planen, Hervorheben und Aufmerksamkeitssicherung zusammenwirken. Wie der vorliegende Fall zeigt, lassen sich auf pragmatischer Handlungsebene kontextuelle Bedingungen spezifizieren, in denen die funktionale Rolle der Hesitationsvokalisation auf der Ebene interaktiver Wirksamkeit außer Kraft gesetzt ist. A befindet sich bereits selbst in einer Planungsphase, die sich in seiner körperlich- ostentativen Orientierungsgestalt auf die verbleibende Objektgruppe am Kartenrand äußert (*1a+b). B's markierte Planungsaktivität ist hier infolgedessen hinsichtlich ihrer interaktiven Wirksamkeit zur Aufmerksamkeitssicherung des Ko-Partizipanten annulliert. Aus dieser Situation heraus emergieren sequentiell nachfolgend die simultanen Handlungsaktivitäten der Teilnehmer.

Nach einer Pause von 1.2 Sekunden beginnt B seine rechte Hand in deiktischer Verweisformation (mit ausgestrecktem Zeigefinger) trajektorisch diskontinuierlich und teils regressiv (mit Unterbrechungen in der gestischen Ausführungskontinuität) über die Karte zum Objekt "Hotel" (H) zu führen (*2a+b). A reagiert in Teilsimultaneität auf die jetzt auch körperlich projizierte Zögerlichkeit mit der physischen Auswahl des Objekts "FishSign" (FS) aus der bereits zuvor fokussierten Objektgruppe (*2a+b).

Sequentiell nachfolgend bricht A – in Reaktion auf B's multimodales Äußerungsformat aus emergenter Fragadressierung „meinen=das=was“ (02) und deiktischem Verweis – seine Handlungsaktivität ab (*3a+b). Interessanterweise ist B's deiktische Zeigegeste auf das Objekt "Hotel" für den emergenten Handlungsakt semantisch irrelevant, da das Objekt nicht den Referenten der verbalsprachlichen Repräsentation „was glauben sie was sie für ne perSONENgruppe hiermit ANsprechen,“ darstellt (vgl. 02+03). Gesprächsorganisatorisch jedoch übt der deiktische Verweis eine Attraktorfunktion aus (2.2, Abbildung 5), die die Aufmerksamkeit des Ko-Partizipanten bindet und zum faktischen Verweisort „Quadpark“ (QP) lenkt (*3a+b, 03, *4a+b). In Ko-Orientierung auf das Objekt „Quadpark“ verändert A abschließend seine Beteiligungsrolle im "participation framework" und konstituiert sich körperlich als Beobachter der Fortsetzungsaktivität seines Ko-Partizipanten (*5a+b).

Konsequente Form: Imperativ zum Erhalt des Handlungsrechts

Im weiteren Handlungsverlauf (14:00-15:36) manifestiert VP-B den Anspruch eines apodiktischen Rollenverhältnisses im "participation framework", indem er sich durch fortlaufende Handlungsaktivitäten als gegenwärtiger Sprecher etabliert. Zum einen drängt er seinen Ko-Partizipanten durch weitere Fragen in eine prototypische Rezipientenrolle defensiver Rechtfertigung („was GLAUBen sie WAS für leute kommen in so=nen Hotel hier,“). Zum anderen unterbreitet er insistierend Vorschläge zur Platzierung und Deplatzierung von Objekten, ohne seine Aktivitäten im Sinne eines kooperativen Handlungsvollzugs mit den Interessen des Ko-Partizipanten abzugleichen oder ratifizieren zu lassen. Im Zuge dessen stellt er das Objekt „Naturlehrpfad“ (NT) auf ein Waldstück, welches bereits zuvor von VP-A mit dem Objekt „Quadpark“ (QP) besetzt wurde (vgl. Fragment 7, *4b). Dazu entfernt er das Objekt von der Karte

(“QAUDstrecke !GAR! nich=GAR nich;“) und platziert es in den „passiven Handlungsraum“ der verbleibenden Planungskonzepte am Kartenrand. Das nachfolgende Fragment zeigt nun A's zweiten Versuch (1:36 min nach seinem ersten Versuch (Fragment 7)) seinen Rezipientenstatus in ein handlungsaktives Beteiligungsformat zu wandeln.

FRAGMENT 8:

F2F-Baseline:	VP-Gruppe 03 (VP-A: Investor / VP-B: Naturschützer)
Zeit:	15:36 - 15:43

01 B-ver: so: DANN (-) |°hh ähm:: |°h hier ham wir HIER
 B-lH: |lH~~~~stack (QP) |
 B-rH: |grasp>>>>>>>>>>>>>>>MG
 A-act: |grasp>>>>>>>>>>>>>>>H
*1 *2 *3



Im Übergang zur nächsten Aktivität etabliert sich B wiederum – wie schon zuvor (vgl. Fragment 7) – „prophylaktisch“ als handlungswilliger Sprecher, indem er ohne konkrete Objektauswahl seinen Anspruch auf die interaktionale Fortsetzung verbalsprachlich projiziert („so: DANN (-) °hh ähm::“ (01)). Auf kinesischer Ebene manifestiert sich sein verbalsprachlich explizierter Anspruch als Planungsaktivität und Aufmerksamkeitssicherung: Die trajektorielle Diskontinuität, in der er zuerst seine linke Hand zum Objekt „Quadpark“ (QP) führt (*2), sie dann retrahiert, um mit der rechten in erkennbarer Zögerlichkeit das Objekt “MiniatureGolf” (MG) auszuwählen (*3), verweist auf eine räumliche Beanspruchung des Handlungsrechts, in der die Fortsetzungsaktivität nicht a priori geplant ist, sondern im Vollzug selbst entwickelt wird. A reagiert auf die projizierte Zögerlichkeit mit dem Versuch, selbst ein Objekt als nächste interaktionale Aufgabe zu etablieren. Im Vergleich der Screenshots *2 und *3 wird sichtbar, wie er seine rechte Hand aus der Ruheposition (“rest position”, Kendon 1980: 212) in eine palmar dem Objekt “Hotel” (H) zugewendete Greifposition bringt.

02 B-ver: WARTe ich- (.)|will nur
 B-rH:@MG|((führt gegriffenes Objekt MG zur Kartenmitte))
 A-act:|((führt Hand ohne H zurück in ihre HP))
*4 *5 *6a+b

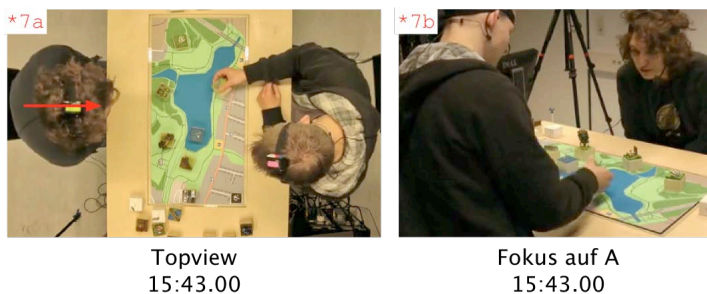


Die emergierende Handlungsparallelität und damit einhergehende Bedrohung des Handlungsrechts wird von B umgehend mit dem imperativen Modus „WARTe ich- (.) will nur“ (02) bearbeitet. Dabei zeigt sich,

dass B im Moment der Intervention sowohl in der körperlichen wie gestischen Orientierung vollständig auf das Objekt “MG” ausgerichtet ist (*4). Aus den divergenten Aufmerksamkeitsfoci der Teilnehmer und B’s unmittelbarer Reaktion auf die emergente Handlungsparallelität A’s, lässt sich ein hochfunktionaler Koordinierungsmechanismus ableiten, der auf Aktivitäten im “peripheral monitoring” (vgl. Heath et al. 1995) zurückzuführen ist. In diesem Zusammenhang stellen Gullberg und Holmqvist (2002: 212) fest: “As long as gestures are moving, peripheral vision is sufficient for detecting (and processing) the broader gestural information (location, direction, or size) even when gestures are performed in the periphery.”

Im weiteren interaktionalen Verlauf bearbeitet A die Handlungsaufforderung, indem er seine Greifhand retrahiert (*5) und sich vom Ort der mahndend angezeigten Handlungsaktivität körperlich entfernt (*6a+b).

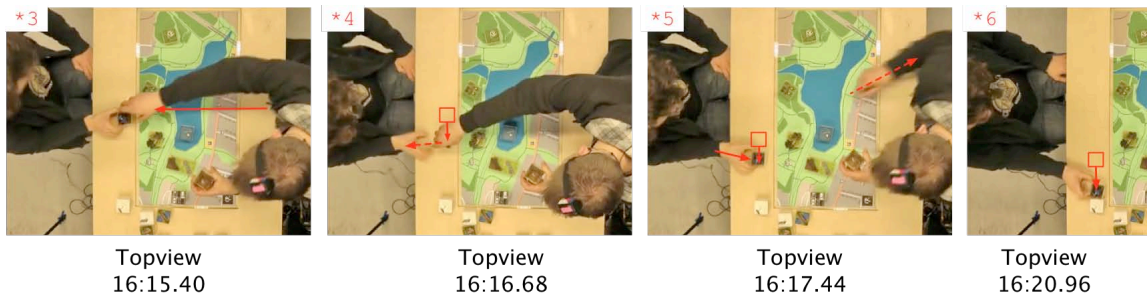
03 B-ver: | das hier auch;
 A-act: | ((lehnt sich nach Vorne zum Objekt MG))
*7a+b



Zum Abschluss werden die ausgehandelten Rollen im “participation framework” konsolidiert: Aus der körperlichen Abwahl der eigenen Aufmerksamkeitspräferenz (vgl. *6a+b) ratifiziert A den Handlungsanspruch seines Ko-Partizipanten, indem er sich als direkter Beobachter der emergenten Fortsetzungsaktivität B’s konstituiert (*7a+b).

Insistente Form: Vorschlagsablehnung durch physischen Handlungsabbruch

Das dritte Fallbeispiel der Verfahrengruppe setzt unmittelbar nach der erfolgreichen Aushandlung und Platzierung des in Fragment 8 vorgeschlagenen Objekts “MiniatureGolf” (MG) ein. Während die vorigen Fragmente Verfahren zeigten, in denen konfligierende Parallelaktivitäten durch verbalsprachliche Handlungen im Frage- oder Imperativmodus elidiert wurden, lässt sich nachfolgend eine abermalige Intensivierung des Handlungsanspruchs feststellen: A’s ostentativ ausgeführte Vorschlagskonstitution wird von B umgehend durch einen physischen Handlungsabbruch und insistenden Hinweis auf die Unvereinbarkeit mit Belangen des Naturschutzes bearbeitet. Aus einer emisch-vollzugsrekonstruktiven Perspektive stellt sich nicht nur der Handlungsabbruch der parallelemergenten Aktivität als eine im Nachdruck intensivierte Form dar, sondern auch A’s Versuch, sich als handlungslegitimer Sprecher zu etablieren: Während in den vorigen Beispielen A’s Handlungsanspruch in einer vom kontextuellen Geschehen separierten Gestaltform erfolgte, bezieht er ihn dieses Mal ostentativ auf die emergente Äußerungsgestalt seines Ko-Partizipanten.



Äquivalent zu den vorigen Beispielen, wird auch im dritten Anlauf der Versuch der Handlungsübernahme abgewehrt. Interessant ist dabei, dass die Auflösung der konkurrierenden Handlung in einer sehr vehementen Form erfolgt: B dringt in den peripersonalen Handlungsraum seines Ko-Partizipanten ein (*3), setzt sich über den exklusiven Manipulationsanspruch, der sich in der am Objekt ruhenden Greifhand manifestiert, hinweg und schiebt das Objekt mit den Worten „quadfahren (.) lassen wir ganz weg;“ symbolisch zur Seite (03, *4). Die anschließende Ratifikation A’s erfolgt auf rein nonverbaler Handlungsebene in zwei Schritten: 1) Während B seinen Arm retrahiert, greift A das Objekt “QP” erneut und schiebt es ebenfalls leicht zur Seite (*5). 2) Erst nachdem B sein Anliegen der Nichtvereinbarkeit verbalsprachlich spezifiziert (04+05), erfolgt eine weitere Ratifikation, in der A das Objekt nun vollständig aus dem relevanten Handlungsraum bewegt (*6).

2.3.7.3 Sequentialisierung durch Blickkoordination

Abschließend präsentieren wir ein Fallbeispiel, in dem die analytisch rekonstruierten Bearbeitungsformate „gestische Sequentialisierung“ (2.3.7.1), „Sequentialisierung durch Direktiva“ (2.3.7.2) und einer nachfolgend fokussierten Bearbeitungsform – der „Sequentialisierung durch Blickkoordination“ – in einem Funktionscluster zusammenwirken. Der Auslöser zur Bearbeitung der konfligierenden Teilnehmeraktivitäten liegt – anders als in den vorigen Beispielen – in einer Situation der verbalsprachlichen Gleichzeitigkeit (“overlap”, Sacks, Schegloff & Jefferson 1974).

Die Beteiligungsstruktur und die Funktion der Blickkoordination beim organisierten Sprecherwechsel sind in zahlreichen Studien empirisch ausdifferenziert worden (Kendon 1967, 1978; Argyle & Cook 1976; Goodwin 1980, 1981; Novick et al. 1996; Bavelas et al. 2002; Schmitt 2005). Dabei stellen wir fest, dass Blickkoordination vorwiegend in seiner “floor-apportionment” Funktion (Kendon 1967: 56) analysiert wurde. Sprecher beenden ihren turn meistens mit einem Blick zum Rezipienten, der dann zu Beginn der Turnübernahme seinen Blick vom Ko-Partizipanten abwendet (Kendon 1967). Dieses Verfahren konnte später von Novick et al. 1996 als “mutual break’ pattern” bestätigt werden. Bei der Turnübergabe entstehen dabei kurze Perioden, in denen die Partizipanten ihre Blicke aufeinander richten (ebd.); bei Bavelas et al. 2002 als “gaze window” bezeichnet).

“The speaker typically seeks a response from the listener by looking at him or her, which begins a brief period of mutual gaze. When the listener responds within this gaze window, the speaker quickly looks away, terminating the window and continuing to hold the turn.” (Bavelas et al. 2002: 570)

Die in der Literatur umfassend beschriebene Funktion der Blickkoordination als “floor apportionment” (Kendon 1967: 56) fokussiert ein Ordnungsprinzip, in dem zum Zeitpunkt des Sprecherwechsels die Interaktionsrollen „Sprecher“ und „Rezipient“ klar voneinander abgrenzbar sind. Jedoch sind insbesondere in aufgabenfokussierten Interaktionen, in denen die Interaktanten permanent ihre Aufmerksamkeit zwischen der Planungs-, Bearbeitungs- und Monitoringaktivität koordinieren müssen, simultane Aktivitätstypen konstituierender Bestandteil der Interaktionssituation. Im nachfolgenden Fragment fokussieren wir daher die Funktion der Blickkoordination als “overlap resolution device” (Schegloff 2000), da in der bei uns fokussierten Interaktionsbedingung das Prinzip der Monitoringmaximierung gemeinsamer Wahrnehmung, wie es Goffman für Situationen des “eye-to-eye ecological huddle” beschrieben hat (Goffman 1963: 95), nicht uneingeschränkt gültig ist. Während Schegloffs Konzeption des “overlap resolution device” verbalsprachliche Verfahren fokussiert, wird nachfolgend gezeigt, dass Blickkoordination als initi-

ierende Komponente des Verfahrens fungieren kann und über die Bearbeitungsfunktion des verbalsprachlichen overlaps hinaus, im vorliegenden Fall bei der Herstellung von Ko-Orientierung funktional ist.

FRAGMENT 10:

F2F-Baseline:	VP-Gruppe 02 (VP-A: Investor / VP-B: Naturschützer)
Zeit:	15:56 - 16:23

```

01 A-ver: |man muss den leuten ja auch irgendwas ANbieten damit sie
    A-act: |((offerierende Geste: "palm-up-open-hand".....

02 A-ver: überhaupt ins hotel auch kommen; (.) |und ähm: gut wir ham
    A-act: .....|grasp@WS.....

03 A-ver: WANDERwege un:d spielplatz=aber jetzt nich groß die
    A-act: .....

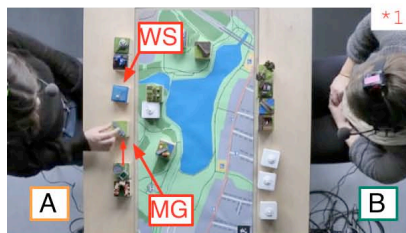
04 A-ver: attraktionen=deswegen würd ich das schon WICHTig finden;
    A-act: .....
  
```

Wir steigen in die Interaktion ein, kurz nachdem VP-A einen Vorschlag zur Platzierung des Objekts „Wasserski“ (WS) verbalisiert hat und VP-B die Objektrealisierung „in erster Instanz“ argumentativ ablehnt, woraufhin A ihren „post-proposal account“ (Houtkoop-Steenstra 1990) in den Zeilen 01-04 erweitert. Die Erweiterung des accounts deckt sich mit der empirischen Feststellung: „when [...] recipients do-not-yet-accept the proposal, the proposers tend to talk more to the proposal“ (ebd.).

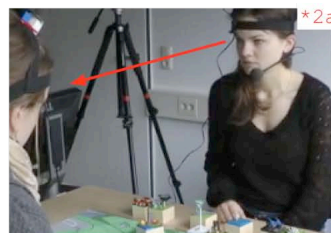
A’s Argumentationsstruktur ist dabei nicht nur verbalsprachlich aufgebaut, sondern in ihrer multimodalen Äußerungsgestalt physisch mit der Integration des Planungskonzepts „Wasserski“ verbunden: Zeitlich und syntaktisch präzise koordiniert sie ihre „palm-up-open-hand“ Geste (vgl. Müller 2004) – als ikonische Realisierung des Arguments „man muss den leuten ja auch irgendwas ANbieten“ (01) – mit der Verbalisierung des abschließenden Hauptarguments zurück zum vorgeschlagenen Objekt. Metaphorisch betrachtet „hält sie an ihrer Vorschlagsrealisierung fest“ und signalisiert dieses nach gestaltpsychologischen Prinzipien mit einem „Festhalten“ des Objekts.

```

05          | (4.0) ..... |
A-ver:          |un(.)t(-) |also..... |
B-ver:          |ich=ver- |                               |ja,
A-gaz: | >>>B... >>>MG..... | >>>B.....
B-gaz: @map..... | >>>A
A-act: ..... | ((rückt MG leicht zu WS)) |
                *1                               *2a+b
  
```



Topview
16:14.84



Fokus auf A
16:16.52



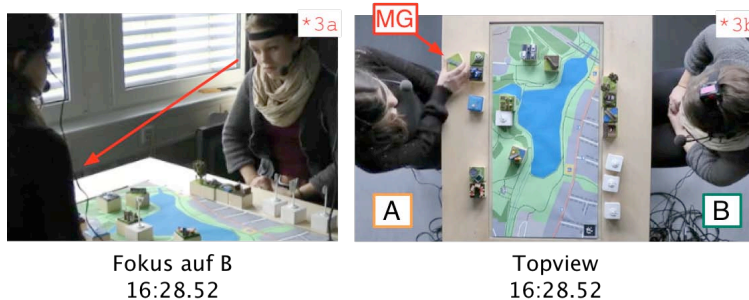
Fokus auf B
16:16.52

Auch wenn die oben geschilderte Nachdrücklichkeit, mit der A ihr Planungskonzept zu realisieren versucht, nicht im Sinne interaktiver Wirkung in B’s Folgehandlung als Teilnehmerinterpretation analytisch klar zugänglich ist, verweist ihre Nicht-Reaktion (05) darauf, gegenwärtig der Argumentation A’s nichts entgegenzusetzen zu können. Dabei ist die Turnübergabe deutlich von A durch eine abfallende Intonation zum Ende des Hauptarguments (04) und abschließender Blickadressierung (05) (vgl. Kendon 1967) markiert. B’s Aussetzen der konditionell relevanten Folgehandlung deutet folglich auf eine Teilnehmerinter-

pretation, in der die Emphase und inhärente Logik der multimodalen Äußerungsgestalt in der Argumentationsstruktur A's eine Gegenargumentation schwer anschließbar macht und einstweilen aussichtslos erscheinen lässt. Aus dieser Konfiguration heraus entsteht eine lange Gesprächspause von 4.0 Sekunden, in der die Rederechtsverteilung durch B's Nicht-Reaktion zeitweilig suspendiert ist und ein verbalsprachlicher "overlap" folgt.

Die konkurrierenden Verbalaktivitäten werden hier als problematischer overlap behandelt und von den Partizipanten bearbeitet. Deutlich nach den akustisch wahrnehmbaren Abbrüchen der Sprecheraktivitäten (markiert durch gutturale Abschlusslaute (-)), projiziert A mit dem Gliederungspartikel „also“ ihren Anspruch auf die turn-Fortsetzung und rückt währenddessen das Objekt „Minigolf“ (MG) in die Nähe des initialen Objektvorschlags „Wasserski“ (WS). Die kinesische turn-Fortführung stellt eine proaktive Form der gestischen Sequentialisierung dar, mit der A eine Fortsetzungsaktivität markiert. Der Fortsetzungsanspruch wird von B jedoch nicht fokussiert. Erst auf A's Blickrichtungswechsel zu B (*2a) reagiert B ihrerseits mit einer Blickrichtungsänderung auf A (*2b). Innerhalb dieser Konfiguration ("gaze-window", Bavelas et al. 2002) übergibt B das Rederecht zur turn-Fortführung durch die verbale Aufforderung „ja,“. Die Fortsetzungsaufforderung markiert dabei gleichzeitig eine Tilgung oder Suspension ihres eigenen Rederechtsanspruches.

```
06 A-ver: |golfen will ich nich;..|      |da is der PLATZ viel zu klein
A-gaz:  |>>>MG.....|>>>B|
A-act:  |((stellt MG zur Seite))|
                *3a+b
```



```
07 A-ver:  =wo soll ich denn hier nen GOLFplatz hinmachen,
```

A setzt daraufhin ihren turn fort, stellt das Objekt „MG“ zur Seite und weist darauf hin, dass das hier als „Golfplatz“ definierte Objekt aufgrund seiner Größe nicht realisierbar sei (*3a+b+06+07). Diese Handlungsaktivität stellt eine Fortsetzung des initialen "accounts" dar und weist eine Argumentation aus, in der andere Freizeitaktivitäten nicht realisierbar sind, so dass dem initialen Vorschlagsobjekt „Wasserski“ eine zentrale Bedeutung zukommt. Dabei ist festzustellen, dass erst die Blickkoordination des "overlap resolution device" zur Ko-Orientierung auf das Objekt „MG“ führt – also die gemeinsame Wahrnehmungswahrnehmung im „gaze window“ funktional von A zur Aufmerksamkeitssteuerung auf das Objekt „MG“ genutzt wird.

2.3.8 Funktionale Simultaneität

Im Unterschied zur konfligierenden Simultaneität (vgl. 2.3.7) bildet das funktionale Simultaneitätsformat keine Beeinträchtigung der kontemporären Aufgabenbearbeitung. In die Kategorie der funktionalen Simultanaktivität lassen sich einerseits Verfahren der situativen Desambiguierung subsumieren, die instantan, teils durch multimodale Verdichtung, potentielle Fortsetzungshandlungen zugunsten einer faktischen Fortsetzungsaktivität elidieren (vgl. Eingangsbeispiel, Kapitel 2.3.6). Andererseits lassen sich auch Sequenzen identifizieren, in denen die simultanen Aktivitäten der Partizipanten im Zusammenhang intrapersoneller Vorbereitungs- und Überprüfungsphasen zur interaktionalen Fortsetzung stehen. Die interdependente Beobachtung der jeweiligen Handlungsaktivitäten ermöglicht dabei den Teilnehmern eine Beurteilung der interaktional relevanten Fragen: Was macht gerade der Ko-Partizipant, erzeugt die Aktivität unmittelbare Handlungsimplicationen und ist die Handlungsaktivität des Ko-Partizipanten (sofern sie keine konditionellen Relevanzen erzeugt) strategisch nutzbar, um eigene Handlungsaktivitäten vorzubereiten?

2.3.8.1 Formen situativer Desambiguierung

Verfahren der situativen Desambiguierung stellen die größte Gruppe funktionaler Simultaneität dar. Sie sind kontinuierlich in Handlungsabläufe von Fortsetzungsaktivitäten integriert und lassen sich in unterschiedlichen Ausdifferenzierungen der Face-to-Face-Bedingung beobachten. Im Eingangsbeispiel haben wir bereits das Handlungsformat der „instantanen situativen Desambiguierung durch multimodale Verdichtung“ (vgl. 2.3.6) analytisch rekonstruiert. Dieses Verfahren ist durch eine sprecherseitige Aktivität gekennzeichnet, in der phänomenologisch unterschiedliche Handlungsaktivitäten (im Beispiel eine symbiotische Einheit aus Objektidentifikation und projizierter -lokalisierung) in sukzessiver Nominierung zu einer instantanen Fortsetzungsaktivität verdichtet werden. Ein strukturell ähnliches Verfahren bildet die situative Desambiguierung in Teilsimultaneität, in der die phänomenologisch divergente Zweitaktivität die Erstaktivität in einem „bifurkativ- multimodalen Verdichtungsraum“ spezifiziert (vgl. 2.3.7.1, Fragment 5, Abbildung 10).

Die Desambiguierungsformate unterscheiden sich dabei hinsichtlich ihrer sequentiellen Lesart: Während das erste Fallbeispiel eine prospektive Lesart vorgibt, in der dem Ko-Partizipanten das Ziel der Objektauswahl projiziert wird, ist der Sinngabungsprozess im zweiten Fallbeispiel an eine retrospektive Lesart gekoppelt, in der die Zweitaktivität die Erstaktivität (im Beispiel eine Objektretraktion) in ihrer Signalisierungsfunktion verändert: Die Objektretraktion wird in ihrer Lesart als terminierter Handlungsakt in die Lesart eines zeitweilig suspendierten Handlungsakts transformiert. Während beide Formate sprecherseitige Desambiguierungsverfahren beschreiben, verweist das nachfolgende Fallbeispiel auf ein rezipientenseitiges Verfahren der situativen Desambiguierung.

Situative Desambiguierung durch visuelle Überprüfung des Bearbeitungsstatus

Während im Eingangsbeispiel zur Differenzierung von Simultanaktivitäten (vgl. 2.3.6) ein Handlungsformat analysiert wurde, in dem eine sprecherseitige Aktivität in einer multimodalen Verdichtung aus Objektidentifikation und -lokalisierung eine situativ-desambiguierende Funktion ausübte, ist es im nachfolgenden Fallbeispiel umgekehrt: Die Handlungsaktivität des Sprechers ist in Bezug auf ihre Fortsetzungsqualität vage und wird vom Ko-Partizipanten durch eine visuelle Überprüfung der sprecherseitigen Handlungsaktivitäten adressiert, um den Bearbeitungsstand der gegenwärtigen Interaktionsaufgabe in Erfahrung zu bringen.

Nachdem die Versuchsperson in der Rolle der Naturschützerin (VP-A) eine spezifische Region zur Platzierung des Objekts „Petting Zoo“ (PZ) vorgeschlagen hat, reagiert die Investorin (VP-B) mit einer unspezifischen Verbalaussage, die weder den Platzierungsvorschlag ratifiziert noch auf weiteren Aushandlungsbedarf hindeutet. A's Blickrichtungswechsel von der Region auf der Karte, in der ihr Planungsvorhaben realisiert werden soll, auf ihre Interaktionspartnerin B, adressiert dabei die Unklarheit über den gegenwärtigen Status der Aufgabenbearbeitung.

FRAGMENT 11:

F2F-Baseline:	VP-Gruppe 11 (VP-A: Naturschützer / VP-B: Investor)
Zeit:	22:53 - 23:30

01 A-ver: |vielleicht wäre es DANN SINNVoll ähm die TIE:re auch (-)
A-act: |((greift Objekt "PZ" und verweist auf eine spezifische Region,

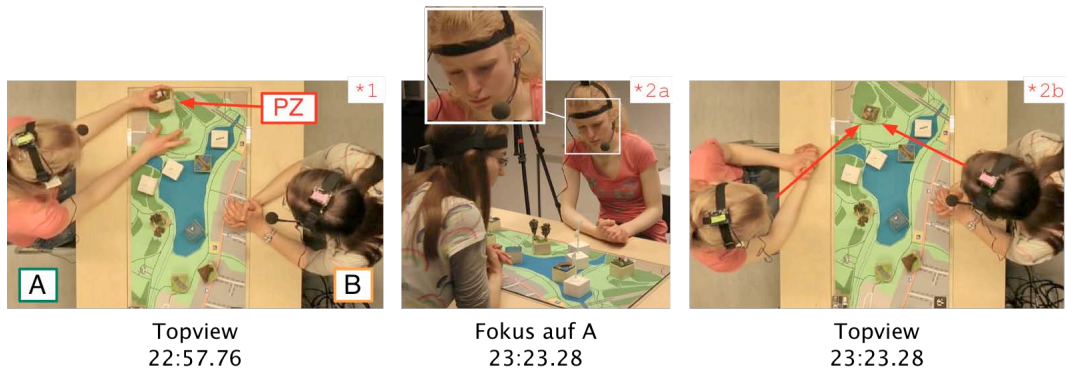
02 A-ver: ein wenig weiter HIERüber anzusiedeln-
A-act: in der das Planungsvorhaben realisiert werden soll))

*1

(22:58-23:19: A argumentiert für die Realisierung des Objekts "PZ")

03 B-ver: |ja:- (-) die fläche könnten wa sonst noch äh:: bele(gen)
B-gaz: |@Fläche.....
A-gaz: |@Fläche.....

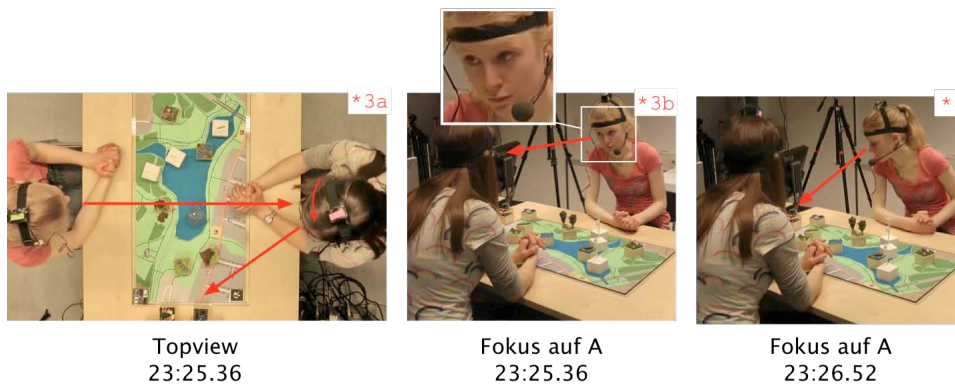
*2a+b



Zu Beginn des Fragments greift VP-A das Objekt “Petting Zoo” (PZ), verweist deiktisch auf eine spezifische Fläche der Karte und verbalisiert einen Vorschlag zur Platzierung des Planungskonzepts (01+02, *1). Im Anschluss an die Vorschlagsverbalisierung argumentiert VP-A 21 Sekunden lang für die Realisierung des Streichelzoos, um die präferierte Reaktion – eine Akzeptanz ihres Vorschlags – zu erreichen (“post-proposal account”, Houtkoop-Steenstra 1990). Ihre argumentative Einheit ist dabei durch Blicke zur Ko-Partizipientin B mit einem Adressatenzuschnitt markiert, der eine Reaktion auf die Vorschlagseinheit konditionell relevant setzt. Infolgedessen reagiert B mit einer unspezifischen Verbalaussage, die weder A’s Platzierungsvorschlag ratifiziert noch auf weiteren Aushandlungsbedarf hindeutet (03+04). Vielmehr wird die Bearbeitungsrelevanz im leise verbalisierten turn-Nachschub „<p>müsste (man) sehen</p>“ (04) auf einen späteren Zeitpunkt verschoben. Während B’s Reaktion sind beide Interaktionspartnerinnen auf das Objekt “PZ” in seiner vorgeschlagenen Platzierung orientiert (*2a+b).

```

04 B-ver: =<p>müsste (man) sehen</p>|
    B-gaz: .....|>>>>stack.....
    A-gaz: .....|>>>B(face)>>>stack
                *3a+b      *4
  
```



```

05 A-ver: °hh GUT; jetzt nochmal um auf das hotel zurückzukommen-
  
```

B’s Teilnehmerinterpretation des Platzierungsangebots zeigt sich deutlich in der darauffolgenden Handlung: Die Positurveränderung (der Oberkörper wird in eine aufrechte, sich vom Objekt entfernende Position gebracht (vgl. *2b vs. *3a)) und die Kopfdrehung zum “stack” zeigen, dass der gegenwärtige Platzierungsvorschlag für sie keine akute Bearbeitungsrelevanz mehr besitzt (04+*3a). Dabei wird B’s “next turn proof” erst in seiner multimodal komplexen Gestalt aus verbalem Redezug und sequentiell nachfolgender Positur- und Blickrichtungsänderung für A als ein (vorläufiger) Abschluss der gegenwärtigen Interaktionsaufgabe interpretierbar. Simultan zur Positurveränderung B’s blickt sie auf und folgt der Blickrichtung ihrer Ko-Partizipientin zu den übrigen Objekten am Kartenrand (*3b+*4). Die Blickrichtungsverlagerung fungiert hier folglich als Monitoringinstrument des “next turn proof” und ermöglicht VP-A eine online-Interpretation des gegenwärtigen Bearbeitungsstandes ihres initialen Platzierungsvorschlags. Das Resultat

der online-Interpretation äußert sich nachfolgend in der Eröffnung einer neuen interaktionalen Aufgabe – der Aushandlung des Objekts “Hotel” (05).

Zusammenfassend veranschaulicht das Fallbeispiel die unterschiedliche funktionale Verwendung der Ressource “Blick”. Während auf Sprecherseite die Blickorientierung auf die verbleibenden Objekte am Kartenrand eine indexikalische Funktion einnimmt, wird der Blick auf Rezipientenseite als Monitoringinstrument des “next turn proof” eingesetzt und liefert im Zuge dessen Informationen, wie die Sprecherin die gegenwärtige Interaktionsaufgabe verstanden und bearbeitet hat.

2.3.8.2 Simultaneität in Überprüfungs-, Planungs- und Kooperationsphasen

Simultanaktivität in der Suchraumalignierung

Die deiktische Orientierungsfunktion des Blickes ist oftmals nicht auf einen klar begrenzten Verweisraum bezogen (vgl. Fragment 11, in dem der “stack” mit den noch zu verhandelnden Objekten einen physisch abgrenzbaren Verweisraum konstituiert), in den der Rezipient aufgrund des online-Monitorings der Sprecheraktivitäten reaktiv orientiert wird. Im folgenden Fallbeispiel wird deutlich, dass der Ko-Partizipant auch proaktiv in Sequenzen agiert, in denen eine Verweisraumkonstitution durch den gegenwärtigen Sprecher antizipierbar ist. Im Gegensatz zu klassischen Deixistheorien, in denen eine starke Fokussierung auf die aktiven Verfahren der Orientierung gelegt wird, lassen sich in natürlichsprachlichen face-to-face Interaktionen auch Ko-Orientierungssequenzen erkennen, die im Vorfeld der Verweisraumkonstitution und faktischen Referenzierung stattfinden.

FRAGMENT 12:

F2F-Baseline:	VP-Gruppe 05 (VP-A: Naturschützer / VP-B: Investor)
Zeit:	13:55 - 12:37

```

01 B-ver: so=n STREI:CHELzoo (--) is ja AUCH nich schlecht|(und)
    A-ver:                                     | ^mhm;

02 B-ver: |hier in der Nähe (-) (XXX)=da SPRICHT doch nichts GEGEN;|
    B-act: |((platziert Objekt PZ auf der Karte))
    B-gaz: |@PZ.....
    A-gaz: |@PZ.....

```

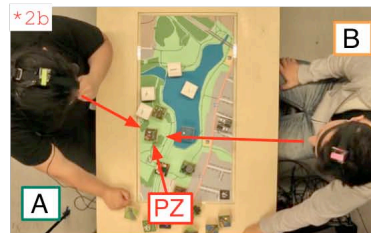
*1 *2a+b



Fokus auf B
14:00.72



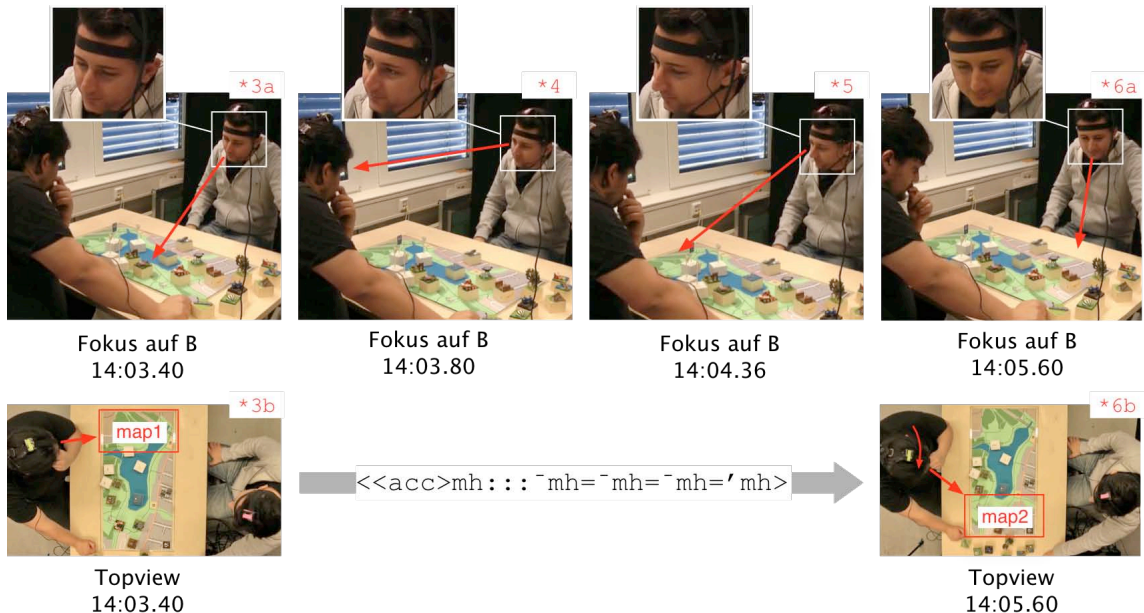
Fokus auf B
14:02.80



Topview
14:02.80

Der Transkriptausschnitt beginnt mit einer Vorschlagskonstitution, in der VP-B in der Rolle des Investors das Objekt “Petting Zoo” (PZ) im Verweisraum lokalisiert (01+02+*1). Der “post-proposal account” (Houtkoop-Steenstra 1990) nach der verbalsprachlichen Objektidentifikation „so=n STREI:CHELzoo“ (01) wird dabei weniger stark elaboriert als im vorigen Fragment und ist eher als suggestiv ausgelegte Persuasionsstrategie zu kennzeichnen (“hier in der Nähe (-) (XXX)=da SPRICHT doch nichts GEGEN;“ (02)). In dieser Zeit sind beide Interaktionspartner auf das Objekt “PZ” ko-orientiert (*2a+b).

03 A-ver: |<<acc>mh:::~mh=~mh=~mh='mh>| |is es SICHER das
 A-gaz: |>>map1.....>>map2.....
 B-gaz:>A(face)>map1...>>map2....>>RC.....
 | (1.2) |
 B-act: |grasp>>>|@RC..turn.....
 *3a+b *4 *5 *6a+b



A reagiert auf B's Vorschlagskonstitution mit einer rhythmisch-phrasalen Hesitationsvokalisation (03), die dem Ko-Partizipanten Hinweise auf eine planerische Aktivität gibt. Das auditive Display eines an sich intrinsisch-kognitiven Prozesses hat gesprächsorganisatorisch die Funktion, den Ko-Partizipanten in die Planungsaktivität mit einzubeziehen. In diesem Zusammenhang weisen Müllerova & Nekvapil (1990) auf die Funktion der Aufmerksamkeitsicherung von Verzögerungspartikeln hin. Da gefüllten Pausen in den meisten Fällen diskursrelevante Informationen oder Handlungsaktivitäten folgen (Keseling 1989), stellt Olbertz-Sütonen (2007: 8) heraus: „Relevantem (Hervorzuhebendem) geht Planung voraus, die für den Gesprächspartner in Form von Zögern erkennbar wird, was wiederum dessen Aufmerksamkeit lenkt.“

Das hier beschriebene Phänomen ist auch im vorliegenden Fragment analytisch rekonstruierbar. Während der Artikulation des ersten Verzögerungspartikels „mh:::“ der Hesitationsphrase orientiert sich A auf den von ihm aus linken Kartenabschnitt, hier als „map1“ bezeichnet (03+*3a+b). 400 Millisekunden später reagiert B, indem er seinen Blick von der Karte löst, A's Aufmerksamkeitsverlagerung wahrnimmt (*4) und seinem Blick auf den Kartenabschnitt „map1“ folgt (*5). Zum Ende der phrasalen Hesitationsvokalisation ändert A seine Orientierung auf den hier als „map2“ gekennzeichneten Kartenabschnitt zu seiner rechten Seite. Die Aufmerksamkeitsverlagerung A's auf den zweiten Suchraum wird dabei fast zeitgleich vom Ko-Partizipanten B nachvollzogen (*6a+b). Das multimodale Äußerungsformat aus prosodisch markierter Hesitationsphrase und sich verlagernder Körper-, Kopf- und Blickorientierung übt folglich eine aufmerksamkeitssteuernde Funktion aus und führt sequentiell zur simultanen Orientierung auf die durch A konstituierten Suchräume. Dabei ist es unerheblich, dass keiner der im vorliegenden Fallbeispiel eröffneten Suchräume „map1“ und „map2“ zum Verweisraum wird. Ausschlaggebendes Element der Aufmerksamkeitsicherung während der Hesitationsvokalisation ist die projizierte und für B antizipierbare potentielle Handlungsimplication, die ein Verweis auf einen alternativen Platzierungsraum für ihn hätte.

Analytisch interessant ist die zeitliche Ausdehnung der Aufmerksamkeitsicherung. Nach einer Pause von 1.2 Sekunden, in denen die Akteure auf den zweiten Suchraum „map2“ ko-orientiert sind, widmet B seine Aufmerksamkeit einem neuen Objekt („RopesCourse“ (RC)(vgl. 03)). Hier wird deutlich, dass die Aufmerksamkeitsicherung in engem Zusammenhang zur Hesitationsvokalisation steht. Die nicht gefüllte

Pause von 1.2 Sekunden dient in der Teilnehmerinterpretation B's als Signal, sich anderen Aktivitäten zuwenden zu können, da A's vokalisierte Planungs- oder Überprüfungsaktivität nicht unmittelbar in einer diskursrelevanten Information oder Handlungsaktivität mündet (vgl. Keseling 1989). Tatsächlich werden die von A konstituierten Suchräume nicht genutzt und A reorientiert mit „is es SICHER das keine tiere AUSbrechen,“ (03+04) auf den initialen Platzierungsvorschlag. Die retrospektive Relevanzsetzung der initialen Platzierung wird umgehend von B durch die Retraktion der Greifhand vom Objekt “RC” und einer verbalsprachlichen Stellungnahme bearbeitet. Nach kurzer Aushandlung (04-07) ratifiziert A den initialen Platzierungsvorschlag seines Ko-Partizipanten (08).

```

04 A-ver:   keine tiere AUSbrechen, |
      B-act:   .....>>>>>>>>>>>>>>>>HP|
              | (1.1) |
      B-ver:   |!JA! wir werden schon dafür

05 A-ver:   |und es wird auch (.) FACHkräfte geben die sich
      B-ver:   sorgen=das |da keine TIERE-

06 A-ver:   rund um die uhr um die TIERE |kümmern;
      B-ver:   |natürlich=<<p>natürlich;>

07 A-ver:   DANN::: |vielleicht noch ein bisschen weiter !DA! HIN
      A-act:   |((schiebt PZ leicht zur Seite))

08 A-ver:   aber es is GRUNDlegend in ordnung;

```

Die analytisch rekonstruierte „sequentielle Suchraumsalignierung“ weist phänomenologische Übereinstimmungen zu experimentellen Befunden von Richardson und Dale (2005) auf, die eine signifikante Ähnlichkeit der Augenbewegungsmuster von Zuhörern und Sprechern einer visuellen Szenerie fanden. Ihre Ergebnisse deuten zusätzlich auf eine Abhängigkeit zwischen dem Grad der Ähnlichkeit und der Verstehensleistung der Zuhörer: So stieg die Verstehensleistung je ähnlicher die Augenbewegungsmuster zwischen Sprecher und jeweiligem Zuhörer waren. Aus den Ergebnissen explizieren sie die These: “Perhaps it is not that moving the eyes closely in step with a speaker supplies visual data, but that it allows the listener to use spatial structure to organize information in the same way as the speaker“ (Richardson und Dale 2005: 1058).

Aus einer praxeologischen Perspektive erscheint diese These in ihrer Generalisierung nicht haltbar, da sie sich auf eine idealisierte Vorstellung der Sprecher- und Zuhörerrolle bezieht. Ein Blick auf die methodische Realisierung des Experiments zeigt, dass Sprecher und Zuhörer keine kopräsenten Akteure waren, sondern zeitlich unabhängig voneinander über die visuelle Szenerie sprachen bzw. der monologischen Einspielung beim Betrachten der Szenerie zuhörten. In der faktischen Komplexität natürlichsprachlicher Interaktionssituationen ist die Organisation visueller Informationen an die situativen Anforderungen des jeweiligen Aufgabenformats gebunden. In der vorliegenden Analyse manifestiert sich im Zuge dessen der empirische Hinweis, dass Ähnlichkeiten im Blickbewegungsmuster der Interaktanten den situativen Bearbeitungsanforderungen der gegenwärtigen Interaktionsaufgabe geschuldet sind.

Handlungsalignierte Simultanaktivität

Handlungsparallelitäten sind konstituierender Bestandteil aufgabenfokussierter Interaktionssituationen und als sinnhafte Optimierung der Aufgabenbearbeitung funktional. Im Vergleich zum Phänomen der „sequentiellen Suchraumsalignierung“ (Fragment 12) verweist das nachfolgende Fragment nicht auf einen prüfenden, sondern auf einen planerischen Typus simultaner Handlungsaktivitäten.

Der Beginn des Transkriptausschnitts markiert eine Transitionsphase interaktionaler Aufgaben. Im Vorfeld des Transkriptausschnitts beansprucht VP-A, in der Rolle der Investorin, das anfänglich VP-B zugewiesene Objekt “RopesCourse” (RC) („aber dann gehört der EHER MIR oder,“). Daraufhin händigt ihr VP-B das Objekt aus. Der Beginn des Transkriptausschnitts markiert einen Zeitpunkt, in dem B die Aushändigung des Objekts mit „aber man DARFS auch nich überTREIBEN;“ abschließend kommentiert (01).

FRAGMENT 13:

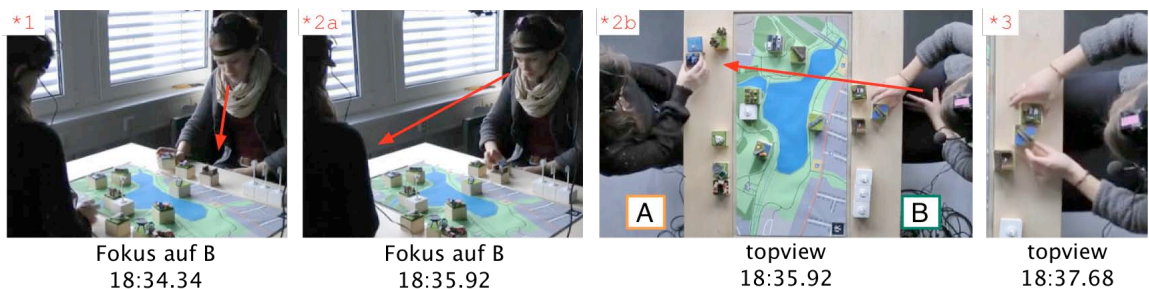
F2F-Baseline:	VP-Gruppe 02 (VP-A: Investor / VP-B: Naturschützer)
Zeit:	18:25 - 18:47

```

01 A-ver:                                     |ja das STIMMT wir
    B-ver: |aber man DARFS auch nicht überTREIBEN;|
    B-act: |grasp@PZ... >>>>grasp@BR...lift>>>hold.....

02 A-ver:  dürfens ECHT=nich übertreiben; (-) okay den SPIELplatz ham wa-
    B-act:  .....

03 A-ver: |hier das is weg=das is AUCH weg-|
          |(--)|
    B-ver:                                     |<<p>ja->
    A-act: |((stellt WS+QP zur Seite)).....|
    B-act: .....|((stellt BR zur Seite))((ordnet PZ+BR+SP...
    B-gaz: |@SP+PZ.....>>>QP.....|>>>PZ+BR+SP.....
          *1                               *2a+b           *3
  
```



Während B in einer indefiniten Pronominalform ihre Ko-Partizipantin dazu auffordert, nicht zu viele naturgefährdende Maßnahmen zu realisieren, greift sie zuerst zum “PettingZoo” (PZ), ändert dann ihre Auswahl auf das Objekt “Bridge” (BR) und hält dieses leicht angehoben in salienter Position (01). A’s verbalsprachlich ratifizierende Anschlussreaktion „ja das STIMMT wir dürfens ECHT=nich übertreiben;“ (01+02) zwingt B – im Sinne der Vermeidung konfligierender Simultanaktivitäten – zur Suspension der emergenten Handlungsaktivität.

Während B weiterhin das Objekt in suspendierter Handlungsbereitschaft hält, schließt A mit der Eröffnung einer thematisch neuen Interaktionsaufgabe an, in der sie eine Objektplatzierung rekapituliert („okay den SPIELplatz ham wa-“, 02). Wie ihr nachfolgender turn zeigt, handelt es sich dabei um eine Sortierungsaktivität, in der abgewägt wird, ob bestimmte Planungskonzepte geeignete Maßnahmen zur Realisierung darstellten. Die Aussortierung ist aufgrund des fehlenden Adressatenzuschnitts (A blickt nur auf die Objekte) deutlich als monologischer (intrapersoneller) Handlungsakt markiert und erzeugt somit keine Relevanzstrukturen. Infolgedessen inspiziert B zu diesem Zeitpunkt die eigenen vor ihr befindlichen Objekte (*1). Sie reagiert erst auf A’s physische Aussortierung der Planungskonzepte “WaterSki” (WS) und “QuadPark” (QP), indem sie ihren Blick auf die zum Kartenrand geführten Objekte richtet (*2a+b). Nachdem sie A’s Handlungsaktivität als Aneinanderreihung nicht verwendbarer Objekte am Kartenrand wahrgenommen hat, verlagert sie ihre Aufmerksamkeit zurück auf die eigenen Objekte und beginnt diese ebenfalls zu ordnen (*3).

```

04 A-ver: |weiß ich nich was DAS is |(-) RUNDfahrt is auch übertrieben|
    A-act: |((stellt KT zu WS+QP...))|grasp@BBQ.....
    B-act: .....|
          *4
  
```

```

05 A-ver: =man kann ja wandern (1.9) also !GRILLEN! ähm (2.6)
A-act: .....frwd>B.....
B-act: |grasp@sign.....
A-gaz: @BBQ.....
B-gaz: |>>>BBQ.....
          *5a+b

```



topview
18:39.28



topview
18:46.16



Fokus auf B
18:46.16

06 A-ver: n=je also das man am see grillen darf würd ich auch auf

07 A-ver: jedenfalls vorschlagen;

Sequentiell nachfolgend sortiert B ein weiteres Objekt aus (“KidsTrain” (KT)) und stellt es zu der Gruppe ausrangierter Planungskonzepte (04, *4). Erst mit der Vorschlagskonstitution des Objekts „Barbecue“ (BBQ) (04+05, *5a) beendet A ihre intrapersonelle Sortierungsepisode und erzeugt durch die multimodale Konfiguration aus i) verbalsprachlicher Handlungsankündigung „also !GRILLEN! ähm“ und ii) salienter Konstitution des Vorschlagsobjekts (*5a) eine Relevanzstruktur, die B mit einer Aufmerksamkeitsverlagerung auf das Objekt BBQ bearbeitet (*5b). Die durch A’s Vorschlagskonstitution erzeugte Handlungsimplication forciert hier folglich auch die Beendigung der intrapersonellen Planungsaktivität B’s.

Über die analytisch rekonstruierten kontextuellen Bedingungen zur simultanen Gleichzeitigkeit hinaus verweist das Fragment auf die interaktive Konstitution intrapersoneller Handlungsaktivitäten. Teilnehmer orientieren sich am Handlungsgeschehen der jeweiligen Ko-Partizipanten und versuchen in aufgabenfokussierter Interaktion emergierende Zeitfenster, in denen keine Relevanzstrukturen erwartbar sind, sinnhaft für eigene Aktivitäten zu nutzen (z.B. Objektsichtung, Planung, Vorbereitung, Verschaffen eines Überblicks etc.).

Kooperative Simultaneität: Distaler Objektvorschlag

Im nachfolgenden Fragment wird deutlich, dass die simultan ausgeführten Planaktivitäten B’s (vgl. Fragment 13) reflexiv sinnstiftend für eine nachfolgende Vorschlagskonstitution ihrer Ko-Partizipantin werden. Zur Erinnerung: B hatte im vorigen Interaktionsausschnitt ihre emergente Handlungsaktivität – die Auswahl und potentielle Vorschlagskonstitution des Objekts “Bridge” (BR) – aufgrund der Fortsetzungsaktivitäten A’s suspendiert. Während A intrapersonell nicht geeignet erscheinende Objekte aussortierte, elidierte B ihre Handlungsaktivität, replatzierte das Objekt am Kartenrand und ordnete es anschließend – zusammen mit zwei weiteren Objekten – als triadische Aneinanderreihung vor ihr befindlicher Objekte an (vgl. Fragment 13, *2b, *3, *4). Die intrapersonellen Aktivitäten beider Interaktanten wurden im Zuge dessen durch A’s Vorschlag beendet, einen Grillplatz (BBQ) zu bauen.

Fragment 14 behandelt die Transitionsphase zwischen der Aushandlung des Grillplatzes und einer sich anschließenden Vorschlagshandlung, die reflexiv sinnstiftend aus B’s erneuter Manipulation der triadisch angeordneten Objekte hervorgeht.

FRAGMENT 14:

F2F-Baseline:	VP-Gruppe 02 (VP-A: Investor / VP-B: Naturschützer)
Zeit:	20:01 - 20:13

18:54-20:01: Aushandlung der Objektplatzierung "Grill" (BBQ)


```

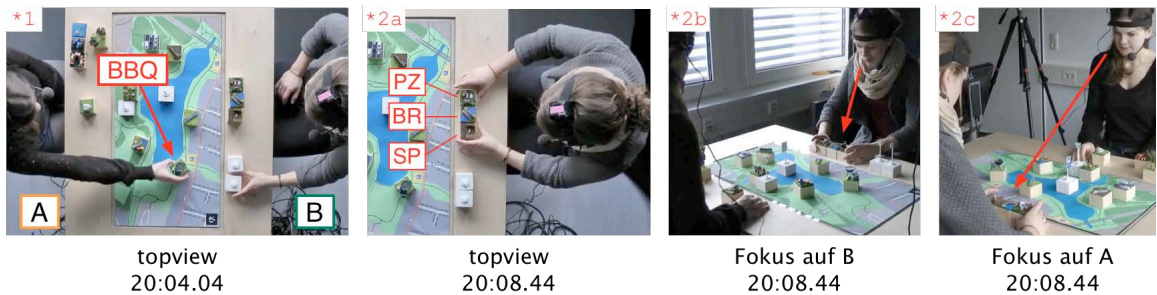
01 B-ver:  das wär in ORDnung=als |KOMPROmiss; |<<all>(XXX)>
    A-act:  |place(BBQ)... |@loc
    A-ver:  |mit GANZ viel
                                     *1

02 A-ver:  mülleimern u::nd-| |erstmal schau'n wie das klappt;
    B-ver:  |ja;|
    B-act:  |((greift zur Objektgruppe PZ+BR+SP))
    B-gaz:  |@PZ+BR+SP.....|
    A-gaz:  |>>>>>>>>>>PZ+BR+SP.....
                                               *2a+b+c

03 A-ver:  ((räuspern)) die BRÜCKE könntest du auch noch hier irgendwo

04 A-ver:  ins NATURschutzgebiet machen oder,

```



Das Transkript startet mit der Beendigung der Aushandlungsphase zum Bau eines Grillplatzes. Unter der Prämisse, für die Entsorgung des Mülls Sorge zu tragen, gibt B ihr Einverständnis, auf das A mit der Platzierung des Objekts auf dem zuvor ausgehandelten Lokalisationsort reagiert (01, *1). In sequentieller Nachfolge expandiert A ihre Objektplatzierung mit dem verbalsprachlichen Zugeständnis, Mülleimer aufzustellen, auf das B mit dem prosodisch markierten Abschlussformat „ja;“ antwortet (01+02). In simultanem Handlungsvollzug zu A's abschließendem Kommentar „erstmal schau'n wie das klappt;“, beginnt B eine neue Handlungsaktivität, in dem sie die zuvor arrangierte Objektgruppe „PettingZoo“ (PZ), „Bridge“ (BR) und „SkatePark“ (SP) manipuliert (*2a+b). In der Teilnehmerinterpretation A's markiert diese Handlung eine reflexiv sinnstiftende Aktivität, die auf B's vergangene Bemühungen zur Etablierung des Objekts „Bridge“ (BR) referiert (vgl. Fragment 13). Infolgedessen interpretiert sie die das Objekt „BR“ rahmende Greifaktivität (vgl. *2a) als Relevantsetzung und verbalisiert einen konzедierenden Vorschlag zur Platzierung (03+04).

2.3.9 Zusammenfassung

Ausgehend von den empirischen Beobachtungen zur „attrahierenden“ und „mediatisierenden“ Aufmerksamkeitsbindung (2.2) haben wir im vorliegenden Kapitel „mutual monitoring“-Aktivitäten in Bezug auf Lösungsstrategien von interaktionalen Fortsetzungsproblemen untersucht (2.3). Motiviert wurde der Untersuchungsgegenstand aus der zentralen Beobachtung, dass Orientierungsgestalten in ihrer aufmerksamkeitslenkenden Funktion unterschiedliche Qualitäten besitzen, die sich auf einem Kontinuum zwischen aufmerksamkeitsbindend (Attraktorfunktion) und aufmerksamkeitsvermittelnd (Mediatorfunktion) einordnen lassen⁴¹. Die Kategorien unterscheiden sich in ihrem Potential, die Summe an Fortsetzungsoptionen (situative Ambiguität), die interaktional begriffen ein Fortsetzungsproblem erhöhter koordinativer Anforderung und Aushandlungsrelevanz konstituieren, hinsichtlich einer faktischen interaktionalen Fortsetzung zu desambiguieren (situative Desambiguierung). In Übergangsphasen zwischen interaktionalen

⁴¹ Wir gehen von einem Kontinuum aus, da es sich um Ausführungsqualitäten einer Orientierungsgestalt handelt, die nicht strikt voneinander zu trennen sind. In Formgestalten mediatisierter Aufmerksamkeitsbindung lässt sich in der Onsetphase eine attrahierende Funktion feststellen. Die Aufmerksamkeit eines Ko-Partizipanten muss sozusagen erstmal in irgendeiner Form gebunden werden, um die Referenzentität inferentiell antizipieren zu können.

Aufgaben sind Praktiken der situativen Desambiguierung besonders relevant, da sich hier handlungsschematische Abläufe (Vorschlag – Aushandlung – Ratifizierung etc.) in einer Situation erschöpfen, in der die erforderlichen Handlungskonstituenten „wer?, wie?, was?“ unbestimmt sind (“What to do next?”, Garfinkel 1967). Die empirischen Beobachtungen aus 2.2 weisen darauf hin, dass sich die Ausführungsqualität situativer Desambiguierungsverfahren sowohl i) an ihrer präkontextuellen Einbettung als auch ii) an der Ostentativität der Formgestalt des multimodalen Orientierungsbündels bemisst.

Auf der Handlungsebene ließ sich für die attrahierende Aufmerksamkeitsbindung eine zeitliche Ausdehnung der Transitionsphase feststellen, in der ein Unklarheitsproblem vorliegt: Die Interaktanten haben keinen Anhaltspunkt, ob die vorige Aufgabe als beendet oder nicht beendet gilt und ein Übergang zur nächsten Aufgabenaktivität eine legitime oder illegitime Handlungsweise darstellt (vgl. Fragment 1). Die aus der situativen Ambiguität konstituierte Orientierungsgestalt zeichnet sich durch eine aufmerksamkeitsbindende Funktion aus, in der die Aufmerksamkeit des Ko-Partizipanten auf den Referenten gelenkt wird. Dabei erfolgt die Aufmerksamkeitsorientierung in simultaner Verlagerung zur gestischen Ausführung der Verweishandlung von der Onset- bis zur Strokephase (kein antizipatorisches Verhalten). Demgegenüber war für die mediatisierende Aufmerksamkeitsbindung ein Minimierungsanspruch der Transitionsphase erkennbar, in der eine ostentative Aufgabenbeendigung nahtlos an eine nächste Aufgabeneröffnung angeschlossen wurde. Obwohl oder gerade weil sich in dieser Übergangsform eine zeitliche Verdichtung und komprimierte Ressourcenkonvergenz der Orientierungsgestalt manifestiert, ließ sich antizipatorisches Verhalten des Ko-Partizipanten sowohl in Bezug auf den relevanten Suchraum (Fragment 2) als auch auf das „sprecherintendierte“ Referenzobjekt (Fragment 3) feststellen. In Fällen instantaner situativer Desambiguierung scheint die initiiierende Funktion der Orientierungsgestalt („Jetzt ist es relevant, sich zu orientieren“) die lenkende Funktion (vgl. Attraktorfunktion) zu überlagern bzw. in ihrer Relevanz zu mindern.

In den untersuchten Fällen zur „attrahierenden“ und „mediatisierenden“ Aufmerksamkeitsbindung ließen sich für den empirisch beobachtbaren Unterschied der Ausführungs- und Situierungsqualitäten von Orientierungsgestalten für die Interaktanten selbst nur eingeschränkt relevante Handlungsimplicationen erkennen. Zwar deutete eine „gelungene“ situative Desambiguierung darauf hin, dass ein zu orientierender Ko-Partizipant besonders „müheless“ einer Orientierungshandlung folgen kann, aber ob Interaktanten einen Vorteil daraus ziehen, den Referenten einer Orientierungshandlung schon vor der Strokephase des gestischen Verweises antizipieren zu können, lässt sich analytisch nicht ermitteln. Antizipatorisches Verhalten ist hier als Auswirkung einer besonders „gelungenen“ situativen Desambiguierung zu verstehen und verweist für sich genommen u.U. auf einen sprachökonomischen Optimierungsprozess.

Eine empirische Relevanz „mutual monitoring“-basierter Verfahren zur situativen Desambiguierung erschließt sich erst gänzlich in der systematischen Untersuchung von Lösungsstrategien interaktionaler Fortsetzungsprobleme und parallelemergenter, kinesischer Simultanaktivitäten (2.3.6, 2.3.7, 2.3.8). Die Analysen der empirisch begründeten Differenzierung zwischen „funktionaler“ und „konfligierender“ Simultaneität haben gezeigt, dass sich das Störpotential simultaner Aktivitäten daran bemisst, auf welche Art und Weise Interaktanten Probleme der interaktionalen Fortsetzung bearbeiten.

Auf Grundlage der empirischen Beobachtungen lassen sich die analytischen Ergebnisse abschließend – in Rückbindung zu den einleitend formulierten Prinzipien (2.3) und in Entsprechung des abgeleiteten Forschungsdesiderats (2.3.5) – zu einer Topographie kinesischer Simultanaktivitäten systematisieren, die ein phänomenologisches Äquivalent zur verbalsprachlichen “overlap“-Klassifikation nach Schegloff 2000 bildet (vgl. Abbildung 9).

Basisdifferenzierung: Kontingente zeitliche Verhältnisse vs. zeitlich konstituierte Relationen

Einen Ausgangspunkt der Differenzierung von Simultanaktivitäten bildete Stukenbrocks Hinweis auf die „erforderliche theoretische Unterscheidung zwischen zufällig zustande kommenden, kontingenten zeitlichen Verhältnissen und zeitlich konstituierten Relationen, [die nicht zufällig sind und denen eine Funktion zugeschrieben werden kann]“ (Stukenbrock 2012: 28; Abbildung 11). Die Aspekte „Nicht-Zufälligkeit“ und „Funktionalität“ sind dabei für die Kategorie der zeitlich konstituierten Relationen auszuklammern (vgl. 2.3.1), da insbesondere Formen „konfligierender Simultaneität“ Ereignisse darstellen, die zwar intrapersonell zeitlich konstituierte Relationen bilden, aber in ihrem interpersonellen Aufeinandertreffen Cha-

rakteristika der Kontingenz (i.S.v. Luhmann⁴²) aufweisen können (z.B. eine Zufälligkeit der konfligierenden Emergenz, eine prinzipielle Offenheit der interaktionalen Fortsetzung etc.). Ebenso verhält es sich mit der „Funktionalität“, die nicht per se in intra- und interpersonellen Simultanaktivitäten inskribiert ist, sondern von den Interaktanten selbst hergestellt und wechselseitig angezeigt wird. Das grundlegende Unterscheidungskriterium der Kategorien liegt somit nicht in primär etischen Merkmalen wie „Zufälligkeit“ und „Funktionalität“, sondern in der emischen Dimension der handlungsimplicativen Wirksamkeit der beiden Simultaneitätsformen.

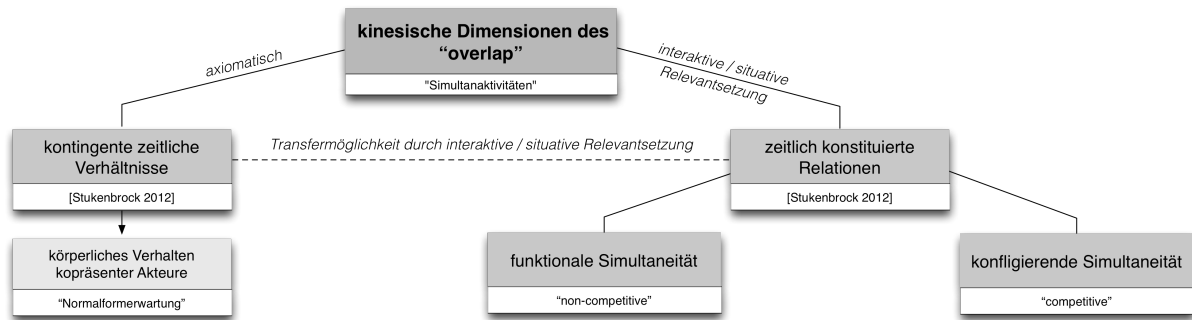


Abbildung 11: Basisdifferenzierung, basierend auf Stukenbrock (2012)

Wenn wir als Unterscheidungskriterium interpersoneller Simultanaktivitäten ihr Potential zur handlungsimplicativen Wirksamkeit anlegen, ist festzustellen, dass die genuin körperliche Anwesenheitsbedingung kopräsender Akteure als kontingente Relation zwar eine permanente Simultaneitätsbedingung in aufgabenfokussierter face-to-face Interaktion konstituiert, diese jedoch – sofern nicht anders markiert – keine Handlungsimplicationen erzeugt. Kontrastierend dazu verweisen die Formen „funktionaler“ und „konfligierender“ Simultaneität in zeitlich konstituierten Relationen auf eine handlungsimplicative Wirksamkeit, die sich in Veränderungen der Aufmerksamkeitsorientierung, suspendierten und elidierten Handlungsaktivitäten, Bearbeitungen interaktionaler Fortsetzungsoptionen, etc. manifestiert. Es erscheint somit relevant, unmarkiertes Verhalten (axiomatisch) von markiertem körperlichen Verhalten kopräsender Akteure (durch intra-/interpersonelle und/oder situative Relevantsetzung⁴³) analytisch zu trennen, da beiden Kategorien unterschiedliche Konstitutionslogiken zugrunde liegen.

Insofern lässt sich aus einer etischen Perspektive auf multimodale Interaktion das Interaktionsgeschehen zwar als permanente Gleichzeitigkeit von Aktivitäten charakterisieren, aus einer emischen, teilnehmerrekonstruktiven Perspektive ist diese jedoch nicht permanent relevant („Simultaneität ist permanent da, aber nicht permanent relevant“, vgl. 2.3.3.2). Unmarkiertes körperliches Verhalten bildet eine Normalformerwartung, die durch „mutual monitoring“-Aktivitäten begleitet wird, aber selbst keine Koordinierungsanforderungen stellt („Interpersonelle Simultaneität bildet eine Normalformerwartung“, vgl. 2.3.3.2). Die koordinative Anforderung besteht folglich nicht darin, sich permanent zu koordinieren, sondern zu markieren und zu erkennen, wann Koordinierungsrelevanzen bestehen⁴⁴. Die Verfahren zur Bearbeitung konfligierender Simultanaktivitäten deuten im Zuge dessen darauf hin, dass „monitoring“-Aktivitäten in der Konstitution intrapersoneller Handlungen als Kontrollmechanismus subsidiär „mitlaufen“, um auf koordinierungsrelevante Ereignisse, wie beispielsweise parallelemergente Handlungsansprüche, rechtzeitig reagieren zu können (prototypisches Beispiel: Fragment 8).

⁴² Luhmann 1984: 47: „Jeder komplexe Sachverhalt beruht auf einer Selektion der Relationen zwischen seinen Elementen, die er benutzt, um sich zu konstituieren und zu erhalten. Die Selektion placiert und qualifiziert die Elemente, obwohl für diese andere Relationierungen möglich wären. Dieses »auch anders möglich sein« bezeichnen wir mit dem traditionsreichen Terminus Kontingenz. Er gibt zugleich den Hinweis auf die Möglichkeit des Verfehlens der günstigsten Formung.“

⁴³ Als „situative Relevantsetzung“ bezeichnen wir Konfigurationen, in denen nicht die Interaktanten ihre Aktivitäten relevant setzen, sondern eine emergente Kontextbedingung konditionelle Relevanzen erzeugt – z.B. im konfligierenden Aufeinandertreffen simultaner Aktivitäten auf gleicher Fokus- und Modalitätsebene (prototypisches Beispiel: Fragment 6). Der Konflikt erzeugt eine Situation, die nur handlungsaktiv gelöst werden kann. Die Relevantsetzung der Aktivitäten erfolgt somit aus der kontextuellen Bedingung selbst.

⁴⁴ Watzlawicks prominent gewordenem Aphorismus „man kann nicht nicht kommunizieren“ wäre hier entgegenzusetzen „man kann kommunizieren, nicht zu kommunizieren“ und „man kann kommunizieren, wann es relevant ist zu kommunizieren“.

Funktionale Simultaneität

Im Eingangsbeispiel (Fragment 4) wurde eine phänomenologische Differenzierung von Simultanaktivitäten in Formen „funktionaler“ und „konfligierender“ Simultaneität empirisch begründet. Unter der Kategorie „funktionale Simultaneität“ (vgl. Abbildung 12) haben wir anhand von sechs Beispielen Verfahren analysiert, in denen

- 1) potentielle Fortsetzungsoptionen (situative Ambiguität), die interaktional begriffen ein Fortsetzungsproblem erhöhter koordinativer Anforderungen und Aushandlungsrelevanz konstituieren, zugunsten einer faktischen Fortsetzungsaktivität instantan und ohne Störausprägung elidiert wurden (situative Desambiguierung).
- 2) simultane Aktivitäten der Interaktanten im kontextuellen Zusammenhang von Überprüfungs-, Planungs- und Kooperationsphasen zur interaktionalen Fortsetzung standen (Konstituenten der Aufgabenbearbeitung).

Für die funktionale Simultaneität beider Verfahrensgruppen ist kennzeichnend, dass sie keine Beeinträchtigungen der kontemporären Aufgabenbearbeitung darstellt, sondern entweder die interaktionale Fortsetzung aktiv und für die Beteiligten nachvollziehbar gestaltet (1) oder eine Möglichkeit zur Überprüfung des gegenwärtigen Fortsetzungsstatus und daraus resultierender Handlungsimplicationen schafft⁴⁵ (2).

Während simultane Aktivitäten im kontextuellen Zusammenhang von Überprüfungs-, Planungs- und Kooperationsphasen als Konstituenten der Aufgabenbearbeitung (2) ein inhärent logisches Simultaneitätsformat bilden und daher mit Verweis auf die Fragmente (vgl. Fragment 12, Fragment 13, Fragment 14) hier in der Zusammenfassung ausgespart bleiben, stellen Verfahren der situativen Desambiguierung (1) ein bislang in der Forschung zur „Koordination“ und „Simultaneität“ kaum berücksichtigte Gestaltform simultaner Aktivitäten dar (vgl. 2.3.3). Dabei lässt sich situativen Desambiguierungsverfahren eine zentrale Relevanz in interaktiven Koordinierungsprozessen zur reibungslosen Emergenzentwicklung interaktionaler Fortsetzungsaktivitäten zuschreiben („smoothness of interaction“). Es handelt es sich um Verfahren, die prophylaktisch oder in situ wirksam werden, bevor sich Fortsetzungsprobleme tradieren. Im Hinblick auf kontextuelle Bedingungen, in denen ein Unklarheitsproblem der Art „what to do next?“ (Garfinkel 1967) vorliegt, bilden situativen Desambiguierungsverfahren ein konstitutives Prinzip zur Organisation der interaktionalen Fortsetzung.

In Bezug auf die Verfahrensqualität („Wie werden ambige Situation desambiguiert?“) hat sich die Ausdrucksassemblierung in Teilsimultaneität als grundlegendes Strukturprinzip herauskristallisiert (Fragment 4 und Fragment 5) („Intrapersonelle Simultaneität erfolgt nur in Teilsimultaneität“ (vgl. 2.3.3.2)). In der analytischen Rekonstruktion sprecherseitiger Aktivitäten haben wir dargelegt, wie phänomenologisch unterschiedliche, aber gestaltpsychologisch zusammenhängende Handlungsaktivitäten in sukzessiver Nominierung zu einer Fortsetzungsaktivität verdichtet werden und so dem Ko-Partizipanten eine „sequentielle Lesart“ der nächsten Aktivität vorgeben. Terminologisch lässt sich diese Art der Gleichzeitigkeit als „sequentielle Simultaneität“ konzeptualisieren – eine Simultaneitätsform, die in sequentiellen Teilsimultaneitäten aufgelöst wird.

In den Analysen ließen sich empirische Hinweise auf phänomenologisch unterschiedliche Ausdrucksassemblierungen zu „sequentiellen Lesarten“ finden: In Fragment 4 war die Lesart prospektiv ausgelegt, in dem die Erstaktivität (eine Objektidentifikation) die sich in teilsimultaner Ressourcenkonvergenz überlappende Zweitaktivität (eine projizierte Objektlokalisierung) in einem „multimodalen Verdichtungsraum“ (Stukenbrock 2012) spezifizierte. Die situative Ambiguität emergierte aus einem diskontinuierlichen Verhalten eines Akteurs, der eine etwaige Aufgabenkonstitution (Objektidentifikation und -manipulation) projizierte, jedoch keinerlei Orientierungshinweise auf die interaktive Relevanz als nächste Interaktionsaufgabe lieferte. Die Diskontinuitätsbedingung adressierend, reagierte der Ko-Partizipant durch eine Äußerungsgestalt in sequentieller Simultaneität, die das ambige Interaktionsgeschehen instantan linearisierte und so in ein für seinen Gesprächspartner (dem „Verursacher“ der Diskontinuitätsbedingung) geeignetes Interpretationsformat überführte. Mit einem Fokus auf die personenbezogenen Aktivitäten deutete das

⁴⁵ Was macht gerade der Ko-Partizipant, erzeugt die Aktivität unmittelbare Handlungsimplicationen und ist die Handlungsaktivität des Ko-Partizipanten (sofern sie keine konditionellen Relevanzen erzeugt) strategisch nutzbar, um eigene Handlungsaktivitäten vorzubereiten?

Beispiel auf ein Zusammenwirken intra- und interpersoneller Simultaneität: Die intrapersonelle Simultaneität der Ausdrucksassemblierung erfolgte in Reaktion, aber in teilsimultaner Überlappung zur intrapersonellen Simultaneität der diskontinuierlichen Aufgabenprojektion. Der interpersonellen Simultaneitätsemergenzz liegt somit eine Konstitutionslogik zugrunde, in der die intrapersonellen Aktivitäten der Interaktanten zeitlich und räumlich zur unmissverständlichen Projektion einer faktischen interaktionalen Fortsetzung relationiert werden. Das Verfahren ist somit durch eine direkte Konfrontationssystematik gekennzeichnet, in der einer Diskontinuitätsbedingung eine Kontinuitätsbedingung als interaktionales Fortsetzungsangebot gegenübergestellt wird.

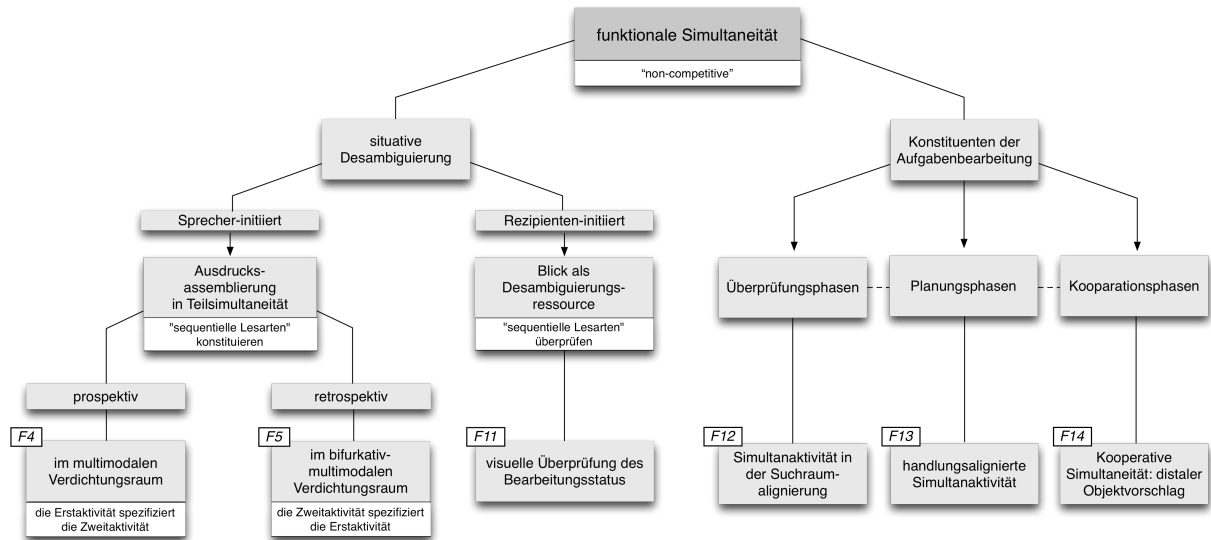


Abbildung 12: Formen funktionaler Simultaneität (Verweis auf die Fragmente: F+Nummerierung)

Gegenüber der prospektiv angelegten Lesart ließ sich in Fragment 5 eine Ausdrucksassemblierung in einer retrospektiven Lesart feststellen, in der die in teilsimultaner Ressourcenkonvergenz überlappende Zweitaktivität (eine linkshändige Objektmanipulation) die Erstaktivität (eine rechtshändige Objektretraktion) spezifizierte. Dieses Format der Ausdrucksassemblierung haben wir terminologisch mit dem Begriff des „bifurkativ-multimodalen Verdichtungsraums“ gefasst, da er das Phänomen gestalttypologisch („bi“: zweifach & „furca“: Gabel) und funktional durch eine Zustandsänderung abbildet, in der die Zweitkomponente der Äußerungsgestalt die Bedeutungskonstitution der Erstkomponente modifiziert. Im Unterschied zu Fragment 4, in der die prospektiv angelegte Ausdrucksassemblierung der Bearbeitung einer Diskontinuitätsbedingung diente, konstituierte die retrospektive Lesart der Ausdrucksassemblierung in Fragment 5 selbst eine Diskontinuitätsbedingung, auf die der Ko-Partizipient vorzeitig, unmittelbar nach Beendigung der Erstaktivität, jedoch in konfligierender Simultaneität zur Ausführung der Zweitaktivität, reagierte. Insofern erscheint die retrospektive Ausdrucksassemblierung im „bifurkativ-multimodalen Verdichtungsraum“ anfällig für Fehlinterpretationen zu sein, da das Format ein Risiko zur teilrezeptiven Äußerungsinterpretation birgt, wenn die mehrkomponentige Konstitutionslogik der Äußerungsgestalt nicht erkannt wird. Das Beispiel bildet somit nicht nur aufgrund der adversativen Ausdrucksassemblierung („prospektiv“ vs. „retrospektiv“) den phänomenologischen Konterpart zu Fragment 4, sondern auch aufgrund der unterschiedlichen Verfahrensqualität, gegenseitig wahrnehmbare Ordnungsstrukturen zur interaktionalen Fortsetzung zu schaffen.

Epistemisch relevant erscheint in diesem Zusammenhang die empirische Beobachtung, dass der Moment des konfligierenden Aufeinandertreffens (im Beispiel zwischen der Zweitaktivität des gegenwärtigen „Sprechers“ und der Reaktion des Ko-Partizipienten) zwar im Sinne der verbalsprachlichen overlap-Konzeption ein Störpotential für die interaktionale Fortsetzung bildet (vgl. „Marked Competition“, Abbildung 9), in der holistischen Aufgabenbearbeitung jedoch durchaus funktional sein kann, wenn die Bearbeitung der Störung („overlap resolution“) ein Ablaufformat nächster Handlungsschritte generiert. Im Beispiel wurde die Zweitaktivität des Sprechers zugunsten der vorzeitigen Reaktion des Ko-Partizipienten suspendiert und nachfolgend als legitime nächste Handlungsaktivität wieder aufgenommen. Hier zeigt

sich, dass die verbalsprachliche overlap-Konzeption als Störfall („competitive“, Jefferson 2004) – als physikalisches Resultat ko-okkurrender akustischer Signale – unter einer multimodalen Perspektive auf Simultaneität zu eng gefasst ist, da lokale Beeinträchtigungen globale Funktionen in der Konstitutionslogik interaktionaler Fortsetzungsaktivitäten einnehmen können.

Den Sprecher-initiierten Desambiguierungsverfahren der prospektiven und retrospektiven Ausdrucksassemblierung in Teilsimultaneität (Fragment 4 & Fragment 5) haben wir abschließend ein Rezipienten-initiiertes Verfahren gegenübergestellt, das ebenfalls im Kontext situativer Desambiguierung stand (vgl. Abbildung 12, Fragment 11). Gegenüber den „produktiven“ Konstitutionsformen zur Herstellung situativ eindeutiger Ordnungsstrukturen der interaktionalen Fortsetzung handelte es sich jedoch um ein visuelles Überprüfungsverfahren ohne eigenständigen Produktcharakter. Die Handlungsaktivität der Sprecherin war im Beispiel (Fragment 11) in Bezug auf ihre Fortsetzungsqualität vage und wurde von der Ko-Partizipanten durch eine visuelle Überprüfung des Bearbeitungsstatus der gegenwärtigen Interaktionsaufgabe adressiert.

Zusammenfassend deuten die empirisch rekonstruierten Verfahren der situativen Desambiguierung auf eine für die interaktionale Fortsetzungslogik funktionale und konstitutive Form der Simultaneität, in der die instantane Vermeidung und Bearbeitung von Unklarheits-/Fortsetzungsproblemen eine „interaktionale Triebfeder“ aller Beteiligten darstellt. Erst wenn diese Koordinierungsanforderung im Moment ihrer Emergenz unerkannt bleibt und nicht bearbeitet wird, kann konfligierende Simultaneität als interaktiv zu bearbeitender Störfall der interaktionalen Fortsetzung entstehen.

Konfligierende Simultaneität

Während funktionale Simultanaktivitäten durch instantane situative Desambiguierung Ordnungsstrukturen schaffen, in denen potentielle Fortsetzungshandlungen zugunsten einer faktischen Fortsetzungsaktivität elidiert werden, bilden konfligierende Simultanaktivitäten selbst eine Situation ambiger Fortsetzungsmöglichkeiten und wirken sich als Störung der kontemporären Aufgabenbearbeitung aus. Auf Grundlage einer Kollektion von sieben Fällen (Fragment 4 - Fragment 10) ließ sich das konfligierende Simultaneitätsformat als interpersonelle Konfiguration gerichteter Handlungsweisen konkretisieren, bei dem die Äußerungsgestalten der Interaktanten faktisch oder in einer teilsimultan nicht mehr auflösbaren Differenz miteinander konkurrieren. Erst die Reparaturverfahren konfligierender Simultanaktivitäten haben eine situationsdesambiguierende Funktion.

Analytisch lassen sich die untersuchten Phänomene konfligierender Simultaneität in drei konstitutive Bedingungen gliedern (vgl. Abbildung 13): (1) Emergenz, (2) Störausprägung und (3) Lösung des Fortsetzungsproblems:

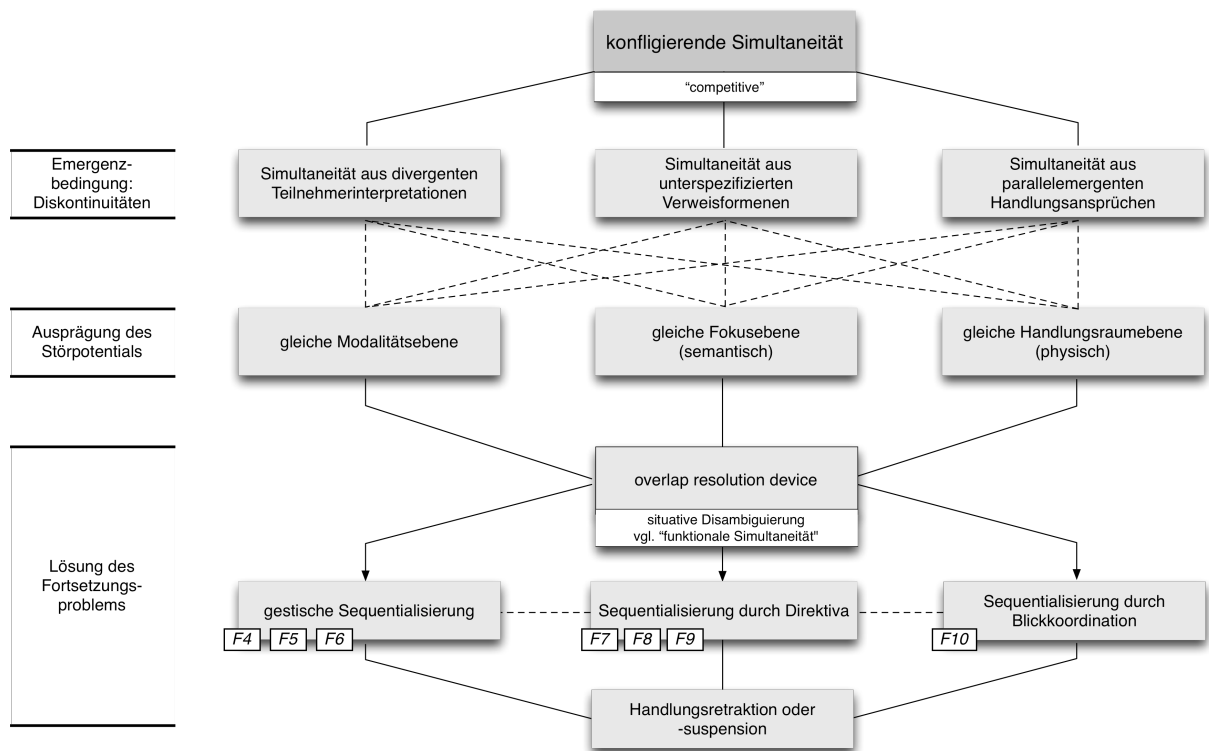


Abbildung 13: Konfigurierende Simultaneität – Emergenzbedingungen, Störausprägungen & Lösungsverfahren

(1) Emergenzbedingung: Diskontinuitäten

Als Ursachenbedingung konfligierender Simultaneität ließ sich in den untersuchten Fällen eine unterschiedliche Ausprägung an diskontinuierlichen Aktivitätszusammenhängen feststellen, in denen interaktionales Verhalten eine mangelnde oder ambige Zielorientierung aufwies. Diskontinuitäten können dabei sowohl etisch-deskriptiver (z.B. trajektorielle Diskontinuität) wie auch emisch-normativer Natur sein (z.B. eine Teilnehmerinterpretation als mangelnde Zielorientierung). In den Analysen ließen sich verschiedene mikrostrukturelle Formgestalten diskontinuierlichen Verhaltens als Ausgangsbedingung konfligierender Simultaneität rekonstruieren:

- a) mangelnde Projektionen von Fortsetzungsoptionen und ein Ausbleiben normativ erwartbarer Folgehandlungen (z.B. Fragment 4)
- b) mehrkomponentige Ausdrucksassemblierungen (Erst- und Zweitaktivität), in denen bereits die Erstaktivität auf verbalsprachlicher und gestaltpsychologischer Ebene eine geschlossene Turnkonstruktionseinheit bildet (z.B. Fragment 5)
- c) ein Ausbleiben erwartbarer Folgehandlungen bei einer unterspezifizierten Verweishandlung (z.B. Fragment 6)
- d) trajektorielle Diskontinuitäten, in denen Zögerlichkeiten projiziert werden (z.B. Fragment 7, Fragment 8)
- e) Pausen und Hesitationsvokalisationen, die dem Ko-Partizipanten Möglichkeiten zur Planung eines parallelemergenten Handlungsanspruchs eröffnen (z.B. Fragment 9, Fragment 10)

Die Verschiedenheit der mikrostrukturellen Formgestalten diskontinuierlichen Verhaltens weist darauf hin, dass es sich hier nur um einen Teil des gesamten Verhaltensrepertoires handeln kann, das es in der Folgeforschung weiter auszudifferenzieren gilt. Zusätzlich ist anzumerken, dass Diskontinuitätsbedingungen kaum in isolierter Form auftreten, sondern Diskontinuitätskonfigurationen ausbilden, in denen verschiedene Merkmale diskontinuierlichen Verhaltens konglomerieren. Sofern die Formgestalten der Dis-

kontinuität nicht per se einen handlungsschematischen Charakter haben, wie in etwa „trajektorielle Diskontinuitäten“, prägen sie sich in einem handlungsschematischen Zusammenhang aus. In Abbildung 13 sind aus diesem Grund die handlungsschematischen Kategorien als makrostrukturelle Repräsentanten der in (a)-(e) analytisch rekonstruierten Formgestalten abgebildet.

(2) Ausprägung des Störpotentials

Bei gleichen Konstitutionsbedingungen („diskontinuierliche Aktivitätszusammenhänge“ (1)) unterscheiden sich die Simultaneitätsformen („funktional“ und „konfligierend“) in der interaktiven Bearbeitung eines situativ emergierenden Fortsetzungsproblems. Während in funktionalen Simultaneitätsformen Verfahren der situativen Desambiguierung instantan greifen⁴⁶, ist für Instanzen konfligierender Simultaneität eine Störausprägung festzustellen, die von den Interaktanten bearbeitet werden muss, um interaktional fortsetzen zu können. Unsere Analysen haben gezeigt, dass die Ausprägung des Störpotentials variiert. Sofortiger Handlungsbedarf besteht in Situationen, in denen das Prinzip der Teilsimultaneität verletzt wird. Dies ist insbesondere dann der Fall, wenn die Aktivitäten der Interaktanten aus keinem eindeutigen Legitimationsanspruch auf das Rede- bzw. Handlungsrecht erfolgen, zeitgleich oder mit minimaler Verzögerung einsetzen („echte Simultaneität“) und auf der gleichen Modalitäts-, Fokus- und Handlungsraumebene konkurrieren, so dass sequentielle Abläufe der Aufgabenbearbeitung gestört werden (prototypisches Beispiel: Fragment 6). In der Konkurrenz richtungsäquivalenter Gestaltformen, in denen interpersonelle Äußerungsgestalten faktisch oder in einer teilsimultan nicht mehr auflösbaren Differenz „aufeinanderprallen“, ist ein kurzer Moment des „Reaktionsverlustes“ beobachtbar, in dem die Interaktanten die Situation kategorisieren müssen, um Fortsetzungsoptionen koordinativ auszuhandeln.

Fälle, in denen die verschiedenen Störkonstituenten zusammenwirken, scheinen jedoch eher ein aufschlussreiches Randphänomen zu bilden. Weitaus frequenter sind jene Fälle, in denen eine Aktivität teilsimultan, also in einer zeitlich auflösbaren Differenz, mit einer Aktivität des jeweiligen Ko-Partizipanten konkurriert (z.B. beim parallelemergenten Anspruch auf das Rede- und Handlungsrecht, vgl. 2.3.7.2). Im direkten Vergleich zu dem bei Schmitt (2005) fokussierten Fall parallelemergenter Melde- und Bittgesten einer Studentin in einem Unterrichtsdiskurs (vgl. 2.3.4) ist jedoch festzuhalten, dass die in 2.3.7 rekonstruierten Formen konfligierender Simultaneität kinesische Realisierungen des „competitive overlap“ (vgl. Abbildung 9) bilden, die sich phänomenologisch deutlich vom Beispiel Schmitts unterscheiden. Die phänomenologische Divergenz manifestiert sich in unseren Beispielen vor allem in der unmittelbaren, räumlich und zeitlich relationierten Aufmerksamkeitskonkurrenz der simultanen Teilnehmeraktivitäten, die eine Störung der sequentiellen Ordnung ausbildet, die behoben werden muss, um das Interaktionsgeschehen fortsetzen zu können. Insofern handelt es sich um kinesische Realisierungen, die sowohl im Minimierungsanspruch der Ko-Okkurrenz als auch im formulierten Prinzip „occurrences of more than one speaker at a time are common, but brief“ der verbalsprachlichen overlap-Konzeption ähneln.

(3) Lösung des Fortsetzungsproblems

Die Konvergenz räumlich und zeitlich relationierter Aufmerksamkeitskonkurrenzen interpersoneller Simultanaktivitäten manifestiert sich als Störung der sequentiellen Fortsetzungslogik und muss von den Interaktanten bearbeitet werden, um das Interaktionsgeschehen weiter kooperativ – und nicht in voneinander abgekoppelten Individualaktivitäten – voranzutreiben. In den empirischen Analysen zur Bearbeitung konfligierender Simultaneität (vgl. 2.3.7) haben wir Verfahren rekonstruiert, mit denen die Interaktanten das Problem der situativen Ambiguität ihrer in Konkurrenz stehenden Aktivitäten lösen. Dabei ließen sich empirische Hinweise auf die drei Verfahrensgruppen

- a) Gestische Sequentialisierung (Fragment 4 - Fragment 6)
- b) Sequentialisierung durch Direktiva (Fragment 7 - Fragment 9), und
- c) Sequentialisierung durch Blickkoordination (Fragment 10)

⁴⁶ Eine instantane situative Desambiguierung setzt voraus, dass die sich diskontinuierlich verhaltene Partei (vgl. (a)-(e)) die Äußerungsgestalt des Ko-Partizipanten als „Kontinuitätsangebot“ wahrnimmt, das die mangelnde Zielorientierung der spezifischen Situation adressiert.

mit strukturell unterschiedlichen Realisierungsmöglichkeiten zwischen Handlungsretraktion- und suspension aufzeigen. Auch hier ist anzumerken, dass es sich nur um eine Teilbeschreibung der Verfahrensvielfalt handeln kann, die es in der Folgeforschung weiter auszudifferenzieren gilt.

Das übergeordnete Prinzip der Bearbeitung konfligierender Simultaneität bildet die Linearisierung in sequentielle Ablaufformate, um situative Ambiguität in gerichtete und gegenseitig wahrnehmbare Fortsetzungsoptionen aufzulösen („Wenn interpersonelle Simultaneität sequentielle Abläufe der Aufgabenbearbeitung stört, wird Handlungsparallelität dem Prinzip der Sukzession unterworfen“, vgl. 2.3.3.2).

Phänomenologisch handelt es sich somit sowohl bei den untersuchten Desambiguierungsverfahren der funktionalen wie auch der konfligierenden Simultaneität um Mechanismen der Linearisierung komplexer Mehrdeutigkeitssituationen. Während die situativen Desambiguierungsverfahren der funktionalen Simultaneität jedoch eher proaktiv wirksam sind, indem einer situativen Diskontinuitätsbedingung ein konkretes Fortsetzungsangebot gegenübergestellt wird („sequentielle Lesarten“), ist bei den Verfahren zur Bearbeitung konfligierender Simultaneität eine eher reaktive Wirksamkeit festzustellen, die auf die interaktive Störausprägung einer initialen Diskontinuitätsbedingung reagiert. Insofern erscheint für die Bearbeitung konfligierender Simultaneität die Bezeichnung „overlap resolution“ durchaus angemessen (vgl. Abbildung 13). Dabei ist jedoch anzumerken, dass bei der konfligierenden Simultaneität neben die Möglichkeit der Störausprägung durch ko-okkurente Aktivitäten (z.B. aufeinandertreffende Gestenvektoren, vgl. Fragment 6) – wie sie in der verbalsprachlichen overlap-Konzeption als physikalische Ko-Okkurrenz akustischer Signale verankert ist – die Störausprägung durch die Emergenz des Fortsetzungsproblems tritt. Anders als bei dem von Schmitt (2005) behandelten Fall parallelemergenter Melde- und Bittgesten im Unterrichtsdiskurs (vgl. 2.3.4) lässt sich für die bei uns untersuchten Fälle konfligierender Simultaneität im Vergleich zum verbalsprachlichen overlap nicht unbedingt ein geringeres Störpotential, sondern vor allem eine andere Störqualität feststellen.

2.4 Mutual Monitoring in Augmented Reality-basierter Interaktion

In face-to-face Situationen konstituieren Gesprächsteilnehmer ihre verbalsprachlichen Redezüge überwiegend in sequentiell geordneter Form. Ausgehend vom “order at all points” Prinzip (Sacks 1984: 22), beschreibt Schegloff (2005), wie dieser in der konversationsanalytischen Forschung bekannt gewordene Aphorismus – losgelöst aus seinem Entstehungskontext⁴⁷ – eine missverständliche Deutung forciert.

“I do not myself believe that there is order at all points, nor do I think that Sacks believed this. If it were so, there would be no need to show that some particular manner of speaking was a “locus of order,” that it was oriented to by participants, etc. If there was order at all points, social life -- and talk-in-interaction in particular -- could always be examined productively at any point and in any respect.” (Schegloff 2005: 17)

Demgegenüber postuliert Schegloff die Lesart “order *possible* at any point” und unterstreicht “*no aspect of talk-in-interaction can be excluded a priori as a locus of order*”.

Die Analysen zur konfligierenden und funktionalen Simultaneität (vgl. 2.3.7 und 2.3.8) bestätigen dieses ursprünglich anhand verbaler Daten explizierte Prinzip. Während funktionale Simultanaktivitäten durch instantane situative Desambiguierung Ordnungsstrukturen zur interaktionalen Fortsetzung schaffen, bilden konfligierende Simultanaktivitäten ein Störpotential, das die kontemporäre Aufgabenbearbeitung beeinträchtigt. Dabei konnte gezeigt werden, dass Interaktanten in face-to-face Situationen über Verfahren verfügen, die es ihnen zu jeder Zeit ermöglichen (“order *possible* at any point”), konfligierende Simultanaktivitäten in sequentielle Ordnungsstrukturen aufzulösen (vgl. 2.3.7).

Konfligierende Simultaneität ist dabei nicht als chaotischer Ordnungszustand zu verstehen, der Situationen konstituiert, in denen die Teilnehmer ohne jegliche Regelmäßigkeit sich monologisierend in ihren Handlungsvollzügen stören. Es handelt sich vielmehr um einen Zustand, der sowohl in seiner präkontextuellen Konstitution als auch im simultanen Vollzug selbst regelhaften Mechanismen folgt, die jedoch in einem auf zeitlichen Abfolgen ausgerichteten Interaktionssystem nicht „verarbeitet“ werden können. Eine konsistente Ordnungsstruktur, in der die Teilnehmer zu jeder Zeit zielführend Abfolgen aufgabenrelevanter Bearbeitungssequenzen hervorbringen, würde dem sozialen Sprachgebrauch natürliche, endo- und exogene Emergenzkriterien absprechen und als reine Planaktivität beschreiben. Als natürliches Phänomen begriffen unterliegt sprachlichem Handeln jedoch ähnlichen Ordnungsprinzipien, wie sie in natürlichen Systemen auffindbar sind (z.B. die „Resilienz“ (Holling 1973) von Ökosystemen, nach der das System nach Störungen wieder in seinen Ausgangszustand zurückkehrt). Schegloffs Lesart ist infolgedessen nicht als maximale Kontrastierung zum „order at all points“ Prinzip zu verstehen, das Chaos von Regelmäßigkeit abgrenzt. Es ist eher als Kontinuum zu begreifen, in dem Ordnungsprinzipien anhand ihrer Verarbeitbarkeit in der jeweiligen kontextuellen Konfiguration bemessen werden.

Die auf permanentem “mutual monitoring” beruhende Fähigkeit, konfligierende Simultanaktivitäten in sequentielle Ordnungsstrukturen aufzulösen, ist in der AR Bedingung gestört. Simultanaktivitäten, in denen die Interaktanten aufgabendifferente Handlungen vollziehen, können in ihrer initialen Entstehung nicht erkannt und bearbeitet werden. Vor diesem Hintergrund verwenden wir den Terminus „Handlungsparallelität“, um simultan vollzogene Teilnehmeraktivitäten zu beschreiben, die sich nicht auf die gleiche interaktionale Aufgabe beziehen, sondern unterschiedliche Handlungsziele verfolgen.

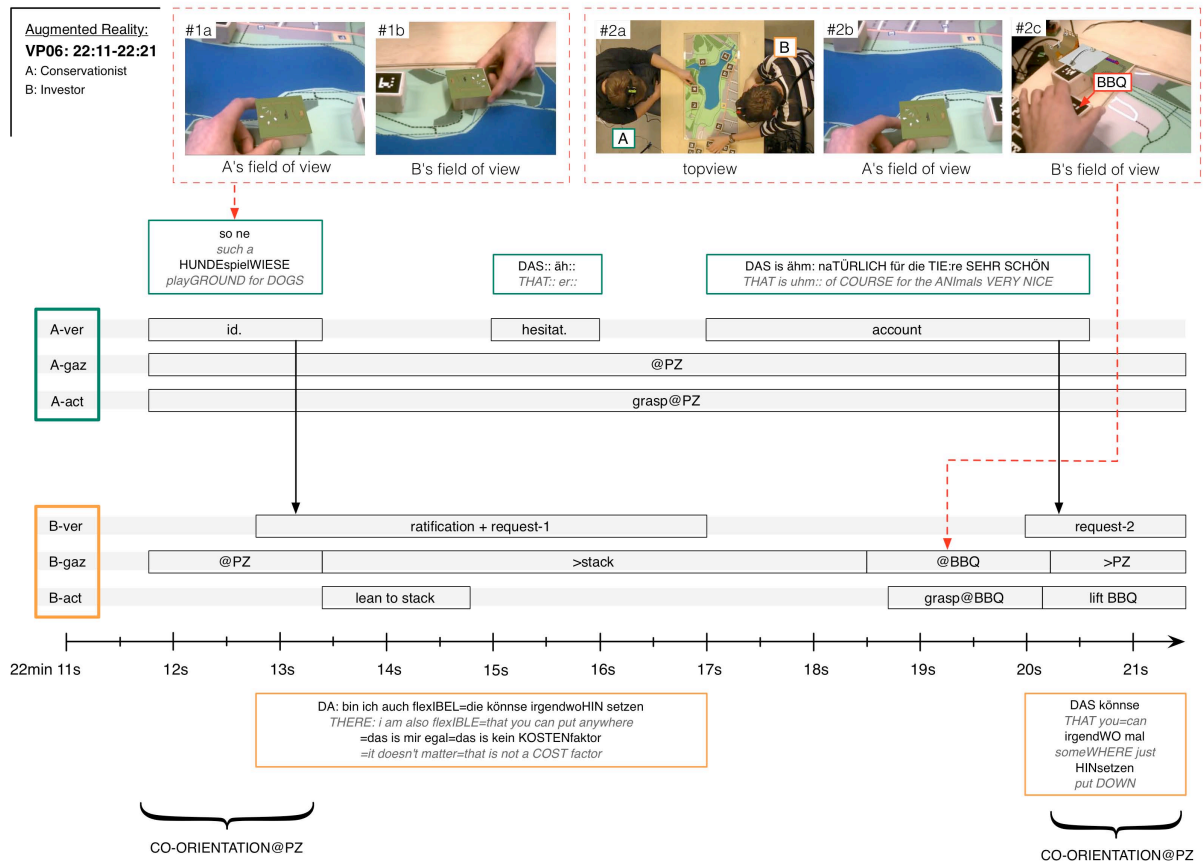
2.4.1 Emergierende Handlungsparallelität in der AR-Bedingung

Für die AR-Bedingung lässt sich feststellen, dass Partizipanten nur vermindert fähig sind, ihre Aktivitäten kleinschrittig zu koordinieren, um emergierender Handlungsparallelität entgegenzuwirken. Das Prinzip “order *possible* at any point” ist in seiner Wirksamkeit eingeschränkt und die Gesprächspartner interagieren unter der Prämisse, ihr Handeln werde vom jeweiligen Ko-Partizipanten wahrgenommen. Eine sich verlagernde Orientierung des Ko-Partizipanten bildet unter den restringierten Bedingungen des “mutual monitoring” kein stabiles Signalisierungssystem, das dem gegenwärtigen Sprecher erlaubt, Reparaturen zur

⁴⁷ “One possible misunderstanding [...] arises when persons disengage the aphorism from the context in which it was introduced, where it was offered as a contrast to the stance underlying a substantial range of work in the social sciences (Atkinson & Heritage, 21-27; Sacks, *ibid.*) to support a new initiative.” (Schegloff 2005: 17)

Rückgewinnung der Aufmerksamkeit einzuleiten. Infolgedessen entstehen Handlungsparallelitäten, in denen der gegenwärtige Sprecher simultan stattfindende Handlungsvollzüge des Ko-Partizipanten erst mit mehr oder minder großer Verzögerung erkennt. Das nachstehende Fragment aus unserer AR Bedingung veranschaulicht dieses Phänomen.

FRAGMENT 15:



Zu Beginn schlägt A das Objekt "Petting Zoo" (PZ) vor (hier von A als „Hundespielwiese“ benannt). Während er das Objekt mit „so ne HUNDEspielWIESE“ identifiziert, platziert er es mittig auf der Karte in einen gemeinsamen Wahrnehmungs- und Handlungsraum und konstituiert mit seiner am Objekt ruhenden Greifhand ein salientes Orientierungsangebot (#1a). B folgt A's Aktivität (#1b), so dass beide Interaktanten zu diesem Zeitpunkt auf das Objekt "PZ" ko-orientiert sind. Das Orientierungsangebot wird dabei von B mit einer sofortigen Ratifikation und Aufforderung zur Platzierung bearbeitet („DA: bin ich auch flexIBEL=die könne irgendwo HINsetzen=das is mir egal=das is kein KOSTENfaktor“). Simultan zur verbalsprachlichen Ratifikation und Platzierungsaufforderung ändert er seine Blick- und Körperorientierung auf die noch zu verhandelnden Objekte am Kartenrand (#2a). Der multimodalen Äußerungsgestalt B's ist zu entnehmen, dass für ihn die gegenwärtige Aufgabe – die Platzierungsaushandlung der Hundespielwiese – abgeschlossen ist. Die projizierte Bedeutungskonstitution der Äußerungsgestalt ist für A jedoch nicht in ihrer ganzheitlichen Form erkennbar: Aufgrund des eingeschränkten mutual monitoring steht ihm nur B's verbalsprachlicher Output zur Verfügung. Goodwins Konzeption sich wechselseitig elaborierender semiotischer Ressourcen folgend ("contextual configurations") bearbeitet A die multimodale Äußerungsgestalt B's bedeutungsdivergent zu ihrer ganzheitlichen Projektion. Die vermittelte „Dringlichkeit“ zum Aufgabenabschluss ist für ihn infolgedessen nicht erkennbar. Aus dieser Konfiguration heraus leitet B bereits ein nächstes Aushandlungsobjekt ein, indem er das Objekt "Barbecue" (BBQ) aus den noch zu verhandelnden Objekten am Kartenrand greift (#2c), während A weiterhin das Objekt „Hunde-

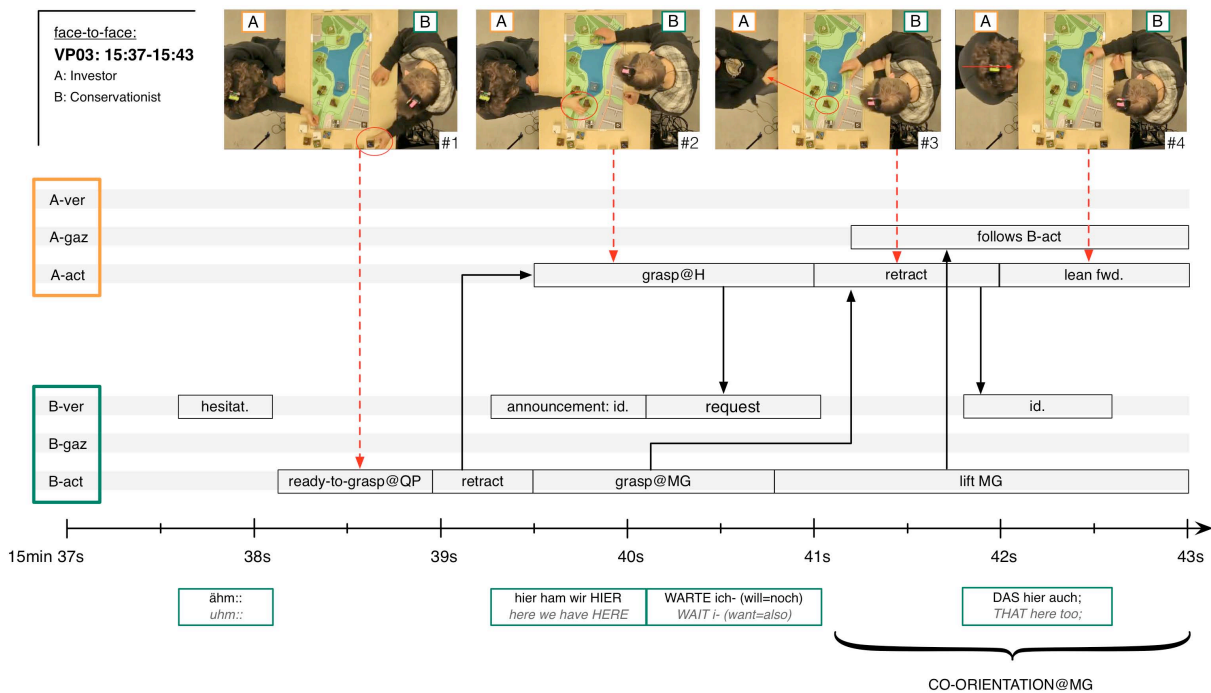
spielweise“ fokussiert (#2b) und in dieser Aufgabenbearbeitung fortfährt („DAS is ähm: naTÜRLICH für die TIE:re SEHR SCHÖN“).

Die Handlungsparallelität erfolgt hier nicht in einer Form, in der beide Interaktanten isoliert verschiedene Aufgaben bearbeiten. Vielmehr erfüllt B weiterhin die konditionellen Relevanzen der Aufgabenbearbeitung seines Ko-Partizipanten. In diesem Sinne betreibt er durch kontinuierliches Insistieren „task sharing“ („DAS könne irgendWO mal HINsetzen“). Die fortwährende Erfüllung konditioneller Relevanzen und ein kontinuierliches Angebot von „Bearbeitungsfragmenten“ zur gegenwärtigen Aufgabe konstituieren ein gesteigertes Anforderungsprofil an den Ko-Partizipanten. Diesem bietet sich kein Grund zur Annahme, dass sein Gegenüber bereits den Fokus der Aufgabenbearbeitung verlagert hat.

Kontrastierung zur face-to-face Bedingung

Mit Verweis auf die verschiedenen Verfahrensformen zur Bearbeitung konfligierender Simultaneität (vgl. 2.3.7) rekapitulieren wir nachfolgend einen Fall der Verfahrensgruppe „Sequentialisierung durch Direktiva“ (vgl. 2.3.7.2), um die Handlungsunterschiede zwischen der AR und F2F Bedingung in konziser Kontrastierung aufzuzeigen. Das Fragment zeigt ein Verfahren, in dem die emergierende Handlungsparallelität instantan und durch kleinschrittige Koordinierung in eine sequentielle Ordnungsstruktur überführt wird.

FRAGMENT 16 (ABBILDUNG DER SEQUENZSTRUKTUR VON FRAGMENT 8):



Der Beginn des Fragments markiert einen Zeitpunkt, in dem B das Rederecht innehat und versucht, seinen Ko-Partizipanten A auf ein neues Objekt zu orientieren. Die Auswahl eines potenziellen Planungskonzeptes zur Aufwertung der Oberseeregion geschieht dabei sehr zögerlich. Diese Zögerlichkeit wird sowohl auf der verbalsprachlichen Ebene durch das Hesitationssignal „ähm:“ als auch kinesisch durch die in Greifbereitschaft geführte, aber nicht greifende Hand zum Objekt „Quadpark“ (QP) projiziert (#1). In der Tat zeigt sich wenig später, dass B seine potentielle Objektauswahl „QP“ aufgibt, sich zur rechten Kartenseite orientiert, dort das Objekt „Miniature Golf“ (MG) greift und die Objektidentifikation mit „hier ham wir HIER“ ankündigt (#2). A reagiert währenddessen auf B's projizierte Zögerlichkeit und die damit verbundene Situation einer nicht festgelegten nächsten Kooperationsaufgabe, indem er das Objekt „Hotel“ (H) greift, um es als potentiell bearbeitungsrelevantes Objekt einzuführen (#2). Zu diesem Zeitpunkt ist erkennbar, dass beide Interaktanten auf unterschiedliche Bereiche der Karte orientiert sind (#2). Dennoch zeigt B's Aufforderung „WARTE ich- (will=noch)“, als unmittelbar auf die emergierende Hand-

lungsparallelität geäußerte Reaktion, dass auch in Situationen, in denen Handlungsaktivitäten des Gegenüber nur peripher wahrgenommen werden können, eine permanente gegenseitige Beobachtung und Einschätzung der jeweiligen Aktivitäten erfolgt (“peripheral monitoring”, Heath et al. 1995).

B’s Aufforderung wird von A mit dem Abbruch der Planungsaktivität bearbeitet, indem er seine Greifhand vom Objekt in eine Position unterhalb des Tisches verlagert (#3). Die Positurveränderung projiziert hier sein Verständnis der Handlungsaufforderung B’s, sich erneut in die Rolle des Beobachters zu begeben. Während B nun das Objekt mit “DAS hier auch;” identifiziert und es salient nach oben in Richtung seines Ko-Partizipanten hält, manifestiert A seine ihm zugewiesene und eingenommene Beobachterrolle, indem er die räumliche Distanz zum Objekt durch eine Oberkörperverlagerung verkürzt (#4).

2.4.2 Auflösung konkurrierender Aktivitäten

Aus einer Kontrastierung der Bedingungen in 2.4.1 leitet sich die Feststellung ab, dass die beiderseitige Orientierung an interaktionalen Fortsetzungsoptionen sehr unterschiedlich abläuft. In der AR-Bedingung können individuelle Interessen zu einer sequentiellen Abkopplung vom kooperativen Handlungsvollzug führen. Während sich in der F2F-Bedingung aus solchen sequentiellen Separationen ein Fortsetzungsproblem der Art „Welche Handlungsoption soll verfolgt werden?“ konstituiert, folglich außerplanmäßige Individualhandlungen durch “mutual monitoring” -Aktivitäten in einem kohäsiven Verhalten münden, wird die Emergenz des Fortsetzungsproblems in der AR-Bedingung erst mit zeitlicher Verzögerung gegenseitig wahrnehmbar. Daher stellt sich die Frage, wie die Auflösung konkurrierender Aktivitäten (“overlap resolution”) unter der Beschränkung des “mutual monitoring” interaktiv gehandhabt wird. Lassen sich vergleichbare Verfahren der Bearbeitung zu den in 2.3.7 rekonstruierten Phänomenen konfligierender Simultaneität der F2F-Bedingung beobachten? Wie wirkt sich die verzögerte Wahrnehmung parallelemerger Handlungen aus?

Bezugnehmend auf die in Fragment 15 dargelegte Okkurrenz emergierender Handlungsparallelitäten in AR betrachten wir nachfolgend die interaktionale Fortsetzung des Falls, um empirische Hinweise auf die Lösungsstrategien parallelemerger Handlungsstränge zu gewinnen. Das Beispiel setzt mit dem account A’s „das is ähm: naTÜRLICH für die TIE:re SEHR SCHÖN“ und der zweiten Handlungsaufforderung B’s „das können=se irgendWO mal HINsetzen“ ein (vgl. Fragment 15). Zu diesem Zeitpunkt hat A noch keinerlei Hinweise auf die parallelemergente Handlungsplanung seines Ko-Partizipanten, die sich durch eine Aufmerksamkeitsverlagerung und der Auswahl des Objekts “Barbecue” (BBQ) abzeichnet. Insofern lässt sich von einer asymmetrischen Informationsverteilung ausgehen. Ein analytisch zugängliches Informationsdefizit liegt bei A vor, dessen Handlungsaktivitäten der Prämisse unterliegen, den gegenwärtig relevanten Handlungsstrang zu bearbeiten. Analytisch weniger transparent ist im vorliegenden Fall das Informationsdefizit B’s, der unter der Voraussetzung der Anwesenheitsbedingung von Ko-Orientierung davon ausgehen muss, sein parallelemergender Handlungsanspruch werde vom Ko-Partizipanten wahrgenommen. Da sich im vorliegenden Fall die gegenseitige Unterstellung von Ko-Orientierung nicht gänzlich zeigt, widmen wir uns vorab der analytischen Ausarbeitung des Handlungsprinzips (2.4.2.1), da es für die im zweiten Teil erfolgende Analyse zur Auflösung konkurrierender Aktivitäten (2.4.2.2) grundlegend ist.

2.4.2.1 Die Unterstellung von Ko-Orientierung

Das Postulat der gegenseitigen Unterstellung von Ko-Orientierung soll nachfolgend durch zwei strukturell unterschiedliche Hinweise analytisch gestützt werden. Während Hinweis 1 der Frage nachgeht, inwieweit die Unterstellung von Ko-Orientierung ein grundlegendes Handlungsprinzip einander zugewandter Akteure in aufgabenfokussierter face-to-face Interaktion bildet, werden in Hinweis 2 empirische Beobachtungen der AR-Bedingung aufgezeigt, in denen das Handlungsprinzip analytisch zugänglich ist:

- Hinweis 1 stellt die zentralen Beobachtungen zur funktionalen und konfligierenden Simultaneität aus der face-to-face Bedingung argumentativ in den Kontext phylogenetischer Tradierung. Als inskribierte Handlungspraxis kopräsender und einander zugewandter Akteure verstanden, ist eine gegenseitige Unterstellung von Ko-Orientierung auf die AR-Bedingung übertragbar.

- Hinweis 2 widmet sich empirischen Beobachtungen aus der AR-Bedingung, in denen eine gegenseitige Unterstellung von Ko-Orientierung analytisch zugänglich wird. Als „Belegfunktion“ wäre dieser Hinweis ausreichend. Eine adäquate Bewertung kann jedoch nur unter dem Verständnis erfolgen, dass die Unterstellung von Ko-Orientierung kein exklusives Phänomen der AR-Bedingung darstellt, sondern auch ein Konstitutionsmerkmal einander zugewandter Akteure in aufgabenfokussierten face-to-face Situationen bildet.

In der analytischen Rekonstruktion des Phänomens fällt auf, dass die Unterstellung von Ko-Orientierung unter eingeschränkten „mutual monitoring“ -Bedingungen in AR analytisch leichter zu fassen ist. Umgekehrt ist das Postulat in der face-to-face-Bedingung kaum zugänglich, da unter der Persistenz des „mutual monitoring“ davon auszugehen ist, dass Orientierungsverfahren unter der visuellen Beobachtung des Ko-Partizipanten gestaltet werden. Eine analytische Ausdifferenzierung zwischen der Unterstellung von Ko-Orientierung und dem Recipientdesign in der Äußerungskonstitution ist daher erschwert.

Hinweis 1: Phylogenetische Tradierung einer kommunikativen Praxis

In den Analysen zur funktionalen Simultaneität (2.3.8) und Bearbeitung konfligierender Simultaneität (2.3.7) ist aufgefallen, dass die Sequentialisierungsverfahren sowohl auf Sprecher- wie auch Rezipientenseite in ihren subtilen Gestaltformen hochfunktionale Mechanismen sprachlicher Kompetenz bilden. Für funktionale Simultaneitätsformen ist festzustellen, dass die sukzessive Nominierung einzelner Ausdrucksressourcen einen kontinuierlichen Prozess an Bedeutungsveränderungen und -erweiterungen in Gang setzt, der eine online-Interpretation des jeweiligen Ko-Partizipanten erfordert (für ein prototypisches Beispiel vgl. Fragment 4).

Die teilrezeptiven Inferenzprozesse einer holistischen Äußerungsgestalt verweisen im Zuge dessen nicht nur auf kommunikativ-kognitive Kompetenzen des Rezipienten, sondern auch des Produzenten, für den die sukzessive Nominierung konvergierend zusammenlaufender Modalitäten sowohl ein kommunikatives Potential als auch eine komplexe intrapersonelle Koordinierungsanforderung darstellt: Einerseits eröffnet das Prinzip die visuelle Darstellungsoption, einzelne Ausdrucksressourcen in einer konnektierten Form einer ganzheitlichen Wahrnehmung zuzuführen. Insofern kann die sukzessive Nominierung in Teilsimultaneität als phylogenetische Tradierung der wahrnehmungsspezifischen Anforderung verstanden werden, die Umwelt nur in zeitlich linearen Formen erschließen zu können (Fixationen und Sakkaden). Andererseits zeigt sich jedoch auch, dass die Aufgabe konnektierter Ausdrucksassemblierung nicht immer optimal gelöst werden kann. Am Beispiel des „bifurkativ-multimodalen Verdichtungsraums“ (Abbildung 10) haben wir dargelegt, dass bei einer unzureichenden Projektion der Konnektivität zeitlich dispersiver, aber im Sinne der Einheitenbildung bedeutungskonstitutiver Ausdrucksressourcen, die Gefahr einer auf Teilrezeption beruhenden Reaktion besteht, die syntaktisch vor der eigentlichen Terminierung der Äußerungsgestalt einsetzt. Im Vergleich wahrnehmbarer und nicht-wahrnehmbarer Konnektierungen einzelner Ausdrucksressourcen zu Äußerungsgestalten lässt sich somit feststellen, dass die Gestaltform in ihrer deiktischen Qualität strukturell sehr unterschiedlich organisiert sein kann.⁴⁸

Aus den empirischen Beobachtungen zur funktionalen Simultaneität manifestiert sich die Erkenntnis, dass die Perzeption von Handlungsweisen kein Randphänomen bildet („ich sehe, was du tust“), sondern genuin mit dem Anspruch einer Primärbedingung erfüllt wird („ich sehe, was du tust und verfolge präzise wie du es tust“). Die Anwesenheitsbedingung der Ko-Orientierung ist dabei nicht als eine faktische Permanenz gemeinsamer Orientierung zu verstehen, sondern als eine gegenseitige Unterstellung, die dem Handeln kopräsender, einander zugewandter Interaktanten in „face engagements“ zugrunde liegt („ich sehe, was du tust, und verfolge präzise, wie du es tust, und gehe davon aus, dass deinem Verhalten der gleiche Anspruch zugrunde liegt“).

Im Rahmen der Spezifik der vorliegenden Situation wird diese Unterstellung vor allem in der Emergenz und Bearbeitung konfligierender Simultaneität auf unterschiedlicher Fokusebene sichtbar. Die Emer-

⁴⁸ In theoretischer Hinsicht bildet der Informationsvorsprung des Äußerungsproduzenten eine mögliche Ursache. Während sich die Assemblierung der Ausdrucksressourcen auf Seite des Rezipienten als modularer Inferenzprozess mit hypothetischen Annahmen zur Bedeutungskonstitution darstellt, ist sie für den Produzenten in erster Linie Mittel zum Zweck. Das heißt, in den meisten Fällen weiß die äußerungsgestaltene Partei, was sie tut und steht damit nicht vor der Anforderung, Bedeutung und Zweck zu inferieren, sondern wahrnehmbar zu bündeln, mit den Zuhörersignalen des Ko-Partizipanten abzustimmen und ggf. zu verändern.

genz von Handlungsparallelitäten stellt hier kein zufälliges Interaktionsprodukt dar, sondern ein strategisches Element, das eine Orientierung an Diskontinuitäten in der kontemporären Aufgabenbearbeitung der primären Handlungsperson („Sprecher“) erkennen lässt (vgl. z.B. Fragment 8). Die strategische Platzierung zeigt, dass ein Handlungsanspruch gemeinhin nicht erschlichen wird („mein Gesprächspartner schaut gerade nicht hin, deshalb kann ich ein Planungsvorhaben realisieren“), sondern unter der Voraussetzung seiner Wahrnehmung einem Kalkül unterliegt, in dem von einer realistischen Chance auf die Rederechts-/Handlungsübernahme ausgegangen wird. Dabei lässt sich eine Orientierung an einem Progressionsanspruch feststellen („Wenn mein Gesprächspartner nicht liefert, bin ich an der Reihe“). Demgegenüber verweisen die Beabreitungsverfahren konfligierender Simultaneität auf die Ratifikation des Kalküls. Im „self-monitoring“ des diskontinuierlichen Handlungsvollzugs kann die gegenwärtig das Handlungsrecht beanspruchende Person auch im „peripheral monitoring“ die emergente Aktivität des Ko-Partizipanten als Handlungsanspruch erkennen und instantan unterbinden (vgl. z.B. Fragment 8). Die Möglichkeit der instantanen „Störabwehr“ setzt ein Wissen voraus, dass die Emergenz in diskontinuierlichen Produktionsphasen als Anspruch auf das Handlungsrecht und nicht als kooperativer Beitrag zur kontemporären Aufgabenbearbeitung zu werten ist.

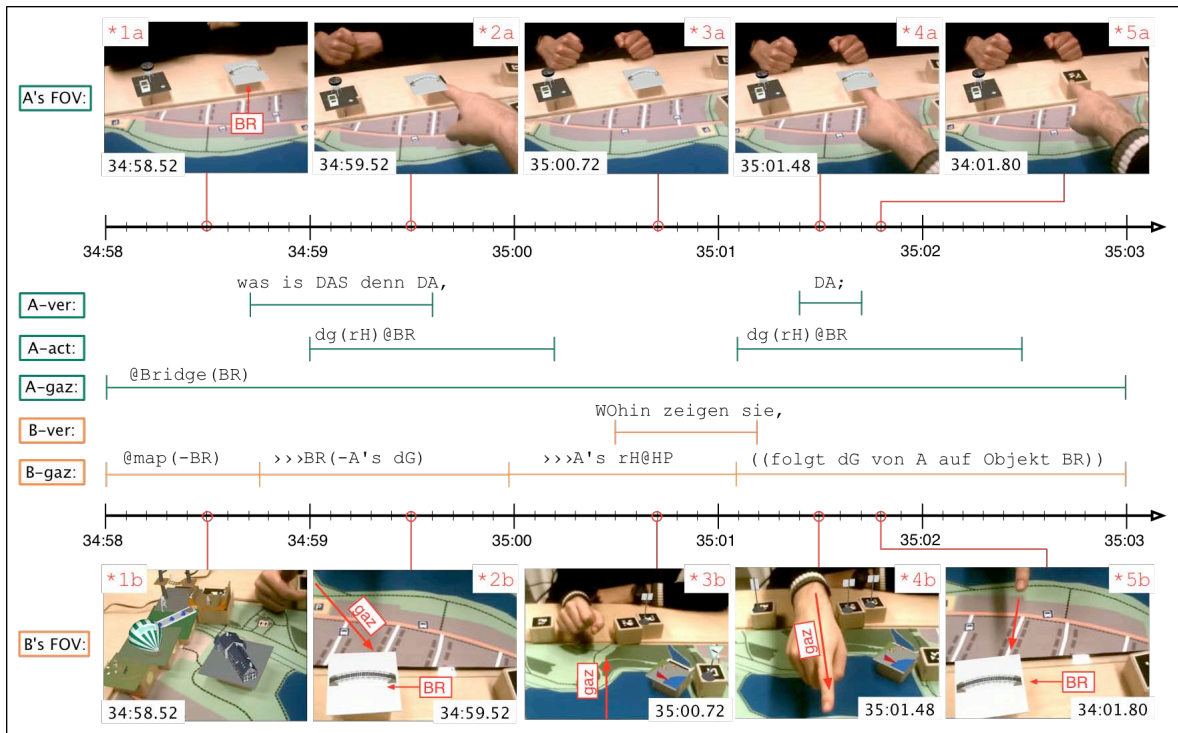
Die in der face-to-face Bedingung analytisch rekonstruierten Formen interaktionaler Wahrnehmung verweisen auf eine hochdifferenzierte Komplexität, die nur unter der Anwesenheitsbedingung, bzw. der gegenseitigen Unterstellung von Ko-Orientierung praxeologisch tragfähig wird. Dabei ist davon auszugehen, dass es sich um phylogenetische Tradierungen einer kommunikativen Praxis handelt, in der wahrnehmungsspezifische Anforderungen im menschlichen Kommunikationsverhalten inskribiert sind. Vor diesem Hintergrund ließe sich die gegenseitige Unterstellung von Ko-Orientierung in der AR-Bedingung als grundlegendes Handlungsprinzip voraussetzen.

Hinweis 2: Empirische Beobachtungen aus der AR-Bedingung

Obgleich sich der vorige Hinweis aus der empirischen Beobachtung ableiten ließ, dass die analytisch rekonstruierten Sequentialisierungsverfahren einen hochfunktionalen Mechanismus sprachlicher Kompetenz im Rahmen der gegenseitigen Unterstellung von Ko-Orientierung bilden, ist die Annahme einer phylogenetischen Tradierung und Übertragbarkeit auf die AR-Bedingung letztendlich nicht belegbar und an eine inhärente Logik und Plausibilität der Darstellung gebunden. Insofern lässt sich hier nicht zuverlässig ermitteln, ob die Unterstellung von Ko-Orientierung auch der AR-Bedingung als Handlungsprinzip zugrunde liegt.

Auf Basis empirischer Beobachtungen wird das Postulat jedoch analytisch zugänglich. Für gestische Verweishandlungen stellen wir fest, dass Interaktanten keine visuelle Überprüfung der Aufmerksamkeitsorientierung des Ko-Partizipanten vornehmen (s.a. Fragment 15, Fragment 19, Fragment 20). In der Unterstellung von Ko-Orientierung wird die Suffizienz des Verfahrens vorausgesetzt (vgl. Fragment 17).

FRAGMENT 17 (AR-BASELINE, VP-GRUPPE 01, VP-A: NATURSCHÜTZER / VP-B: INVESTOR):



Fragment 17 zeigt eine deiktische Verweishandlung im Rahmen eines Objektidentifikationsproblems. Zu Beginn des Fragments ist A auf das unmittelbar vor seinem Ko-Partizipanten befindliche Objekt "Bridge" (BR) orientiert (*1a), während B's Aufmerksamkeitsorientierung auf einem räumlich anderen Teilbereich der Karte liegt, der in der vorigen Interaktionsaufgabe relevant war (*1b). Mit ausgestrecktem Zeigefinger verbalisiert A das Objektidentifikationsproblem im "first-pair part" mit „was is DAS denn DA,“ (*2a). Die deiktische Verweishandlung weist eine kontinuierliche Formgestalt auf, die sich über einen Zeitraum von 1,2s von der Onset- bis zur Retraktionsphase erstreckt und mit einem Stroke von ca. 400ms keinerlei Expansionsauffälligkeiten erkennen lässt.

Die Anschlussreaktion B's ist insofern bemerkenswert, da sie i) in minimaler Verzögerung, teilsimultan zur emergenten Verbalisierung des Objektidentifikationsproblems erfolgt (vgl. Annotation »BR(-A's dG)) und ii) eine Orientierung an einem vektoriellen Suchraum erkennen lässt, ohne diesen anhand der deiktischen Lokalisation seines Ko-Partizipanten räumlich erschließen zu können (A's Zeigehandlung ist für B nicht sichtbar)(*2a). In beiden Punkten manifestiert sich ein antizipatorisches Verhalten, das jedoch auf unterschiedlicher Funktionsebene wirksam ist: Während i) auf eine Orientierung am handlungsschematischen Zusammenhang verweist, in der aus dem kooperativen Abschluss der vorigen Aufgabe eine sequentiell nachfolgende Aufgabenkonstitution antizipierbar wird, veranschaulicht ii) die situative Antizipation einer multimodalen Orientierungsgestalt: Aus dem verbalsprachlichen Verweis wird eine verkörperte Zeigehandlung inferiert (eine retrospectiv-analytische Bestätigung manifestiert sich in der anschließenden Rückfrage „Wohin zeigen sie,“). Dieses Verhalten ist deswegen so bemerkenswert, da es die Vorstellung moderner Deixistheorie, eine Zeigegeste eröffne einen Suchraum, in dem das konkrete Referenzobjekt erschlossen werden muss (vgl. Stukenbrock 2012: 282: "Vektor-Ziel-Problem"; s.a. Goodwin 2000b: "domain of scrutiny"; Kranstedt et al. 2006: "pointing cone"; Rieser 2004: "region-pointing"), dahingehend präzisiert, dass der Suchraum unter bestimmten kontextuellen Bedingungen nicht durch die Zeigegeste eröffnet wird, sondern durch antizipatorische Inferenzprozesse. Das heißt, die Orientierung am Suchraum erfolgt in der Erwartung, die Hypothesen zum spezifischen Referenzobjekt anhand einer antizipierten gestischen Verweishandlung im Suchraum selbst konkretisieren zu können.

Im vorliegenden Fall lässt sich erkennen, dass dieser Inferenzprozess eine Aufmerksamkeitsverlagerung auf den relevanten Suchraum ermöglicht, in dem das intendierte Referenzobjekt liegt (*2b). Die Hypothese des relevanten Suchraums kann jedoch nicht ratifiziert werden, da die deiktische Verweishandlung A's für B nicht sichtbar ist. Unter der Unterstellung gegenseitiger Ko-Orientierung wird die Insuffizienz der Orientierungsverfahrens jedoch nicht in Betracht gezogen. So lässt sich sequentiell nachfolgend beobachten, dass A weiterhin das Referenzobjekt fokussiert und keine visuelle Überprüfung der Aufmerksamkeitsorientierung seines Ko-Partizipanten vornimmt (*3a). Währenddessen ändert B seine Strategie, indem er den Suchraum aufgibt und seinen Blick in Richtung A wendet (*3b). Erst nachdem er A's Verweishand in ihrer "home position" (HP) vorfindet (»A's rH@HP), leitet er verbalsprachlich im "second pair part" mit „WOhin zeigen sie,“ eine Reparatur ein, die von A durch eine zweite, formäquivalente Orientierungsgestalt aus gestischem und verbalsprachlichem Verweis („DA;“) bearbeitet wird (*4a). Von der deiktischen Geste ausgehend (*4b) folgt B dem Orientierungshinweis seines Ko-Partizipanten auf das Referenzobjekt "Bridge" (BR)(*5a+b).

Im Beispiel manifestiert sich die Unterstellung von Ko-Orientierung in zweifacher Weise – einerseits in der Aufrechterhaltung der Unterstellung bis emergierende Probleme in der Verweishandlung sichtbar werden, andererseits in der Problembearbeitung selbst, in der ebenfalls von einer Wahrnehmbarkeit der ausgeführten Reparaturhandlung ausgegangen wird⁴⁹. Dass die Unterstellung von Ko-Orientierung im gegenwärtigen Handlungsgeschehen folgenreich sein kann, zeigt sich in Fällen, in denen eine Orientierungsgestalt fehlinterpretiert wird (Fragment 21 & Fragment 22). Mit Verweis auf die analytischen Ausführungen der Fragmente sei hier nur kurz auf das Phänomen multipler Orientierungsangebote referiert, da es die Problematik konzise veranschaulicht (vgl. Fragment 22): Im dargelegten Fall verhandeln die Gesprächspartner die Platzierung eines Objekts. Während A das Objekt in der linken Hand salient nach oben hält, verweist B auf einen möglichen Lokalisationsort. In Reaktion auf das Angebot formuliert A einen Gegenvorschlag und deutet mit der rechten Hand auf einen alternativen Platzierungsort, während er das Objekt in seiner linken Hand akustisch wahrnehmbar abstellt. Diese Situation schafft akzidentell ein Orientierungsangebot, das B mit der verbalsprachlichen Ebene des alternativen Platzierungsvorschlags seines Ko-Partizipanten A in Verbindung bringt. Ohne Kenntnis auf den faktisch vorgeschlagenen Lokalisationsort drängt B auf die Realisierung seines eigenen Vorschlags. Die Fehlinterpretation bleibt dabei unbemerkt, da A das Insistieren B's als Reaktion auf seinen faktischen Platzierungsvorschlag deutet.

2.4.2.2 Auflösung konkurrierender Aktivitäten unter der Anwesenheitsbedingung von Ko-Orientierung

Die in 2.4.2.1 gewonnenen Erkenntnisse bilden die theoretische Basis der nachfolgenden Analyse zur Auflösung konkurrierender Aktivitäten. Das Beispiel zeigt die interaktionale Fortsetzung der in Fragment 15 dargelegten Okkurrenz emergierender Handlungsparallelitäten in AR und setzt mit dem account A's „das is ähm: naTÜRLICH für die TIE:re SEHR SCHÖN“ und der zweiten Handlungsaufforderung B's „das können=se irgendWO mal HINsetzen“ ein (vgl. Fragment 15). Unter Verweis auf die empirischen Beobachtungen in 2.4.2.1 interagieren beide Gesprächspartner unter der Annahme, ihr Handeln werde vom Gegenüber wahrgenommen. Im vorliegenden Fall bedeutet dies, dass A davon ausgehen muss, seine Handlungsaktivitäten beziehen sich auf die einzige und gegenwärtig relevante Interaktionsaufgabe, während B unter der Annahme agiert, sein parallelememgender Handlungsanspruch werde in seiner markierten Dinglichkeit wahrgenommen.

⁴⁹ Die Frage, warum Ko-Orientierung unterstellt wird, ist nur theoretisch zu klären. Einerseits haben wir darauf hingewiesen, dass die Unterstellung von Ko-Orientierung auch ein grundlegendes Handlungsprinzip einander zugewandter Akteure in aufgabenfokussierten face-to-face Situationen darstellt, wobei das Ausmaß der Unterstellung aufgrund der Persistenz des "mutual monitoring" analytisch nicht zugänglich ist. Andererseits lassen sich für die AR-Bedingung mögliche Hypothesen aufzeigen, die das Verhalten in Erweiterung einer phylogenetisch tradierten Handlungspraxis begründen: Ein verweisender Akteur könnte annehmen oder seinem Verhalten die praxeologische Erfahrung zugrunde legen, dass i) seine Orientierungshandlung sequentiell nachfolgend verbalsprachlich ratifiziert wird, ii) die exklusive Wahrnehmung auf Kopf- und Körperorientierung in der Beurteilung der Aufmerksamkeitsorientierung unzureichend ist (Augen sind durch HMDs verdeckt), iii) durch Ausbleibendes körperliches Feedback irritiert sein (der Blick zum Gesprächspartner kann nicht erwidert werden) oder iv) den aktiven Prozess der visuellen Überprüfung lästig empfinden (bei einer Aufmerksamkeitsorientierung auf die Karte muss die Axialorientierung des Kopfes aktiv verändert werden, da das periphere Sichtfeld vollständig ausgeblendet ist).

FRAGMENT 18 (BEZUGNEHMEND AUF FRAGMENT 15):

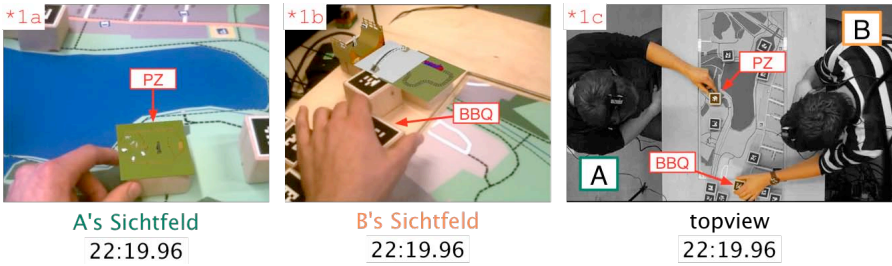
AR-Baseline:	VP-Gruppe 06 (VP-A: Naturschützer / VP-B: Investor)
Zeit:	22:17 - 22:43

```

01 A-ver: |DAS is ähm: |natÜRLICH für die TIE:re |SEHR SCHÖN
A-lh:   |@PZ((mit 2 Fingern)).....
A-gaz:  |@PZ.....
B-ver:                                |das können=se irgendWO
B-act:                                |grasp>>>>>>>>>>>>@BBQ...|lift>>>>>>>>hold....
B-gaz:  |@stack...>>>|@BBQ.....|>>>map.....

```

*1a-c



Zu Beginn des Fragments ist A im “post-proposal account” (Houtkoop-Steenstra 1990) der gegenwärtigen Vorschlagskonstitution auf das Objekt “PettingZoo” (PZ) fokussiert (*1a). Nachdem B eine initiale Platzierungsaufforderung verbalisiert hat (vgl. Fragment 15), orientiert er sich auf die übrigen Objekte am Kartenrand und greift das Objekt “Barbecue” (BBQ) (*1b). Während parallelemergente Handlungsansprüche in der face-to-face Bedingung wahrgenommen und bearbeitet werden (vgl. 2.3.7), bleibt das Phänomen unter der Beschränkung des “mutual monitoring” in AR unbemerkt. Insofern liegt eine Situation vor (*1c), in der konkurrierende Handlungsaktivitäten üblicherweise in einem interaktiven Aushandlungsprozess münden, hier jedoch in sequentieller Abkopplung voneinander emergieren.

```

02 A-ver:                                |und ähm:::
A-lh:   .....|grasp@PZ...lift>>>>>>>place@Kartenrand
A-gaz:   .....|>>>BBQ...>>>>>>>((searching))>>>>>>>
B-ver:   mal HINsetzen |SETZEN sie das DA: =das können=se gerne
B-act:   .....

```

*2 *3a-c *4a+b



Nachdem B das Objekt “BBQ” aus den noch zu verhandelnden Objekten am Kartenrand greift und über die Karte führt, erscheint das Objekt am Ende der zweiten Platzierungsaufforderung „das können=se irgendWO mal HINsetzen“ erstmalig in A’s Sichtfeld (01+02) (*2). An dieser Stelle manifestiert sich für A eine doppelte Koordinierungsanforderung: (1) Auf verbalsprachlicher und kinesischer Ausdrucksebene lässt sich ein Bearbeitungsanspruch der Platzierungsaufforderung erkennen, die in ihrer wiederholten Ausführung die konditionelle Relevanz einer Reaktion erhöht. Verbalsprachlich wird die interaktionale Fortsetzung als planerische Aktivität ausgewiesen („und ähm::“), die gestisch in Form einer akuten Handlungsbereitschaft projiziert wird: Im Vergleich der Screenshots *2 und *3b ist erkennbar, wie die am

Objekt ruhende Hand in eine das Objekt umschließende Greifhand transformiert wird. (2) Die zweite Koordinierungsanforderung äußert sich in A's initialer Wahrnehmung der parallelemergenten Handlungsaktivität seines Ko-Partizipanten: Er löst seinen Blick von der Primäraufgabe „PZ“ (*2) und orientiert sich zur aufmerksamkeitskonkurrierenden Aktivität B's (*3a+c), wobei ihm die Augmentierung durch die partiell verdeckende Greifhand des Markers nicht angezeigt wird (*3a).

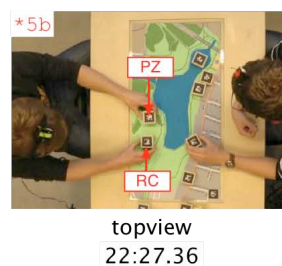
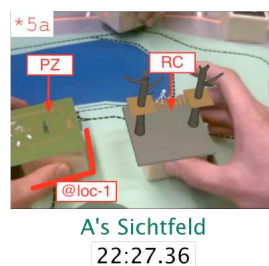
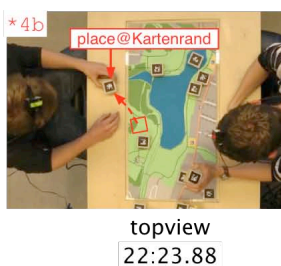
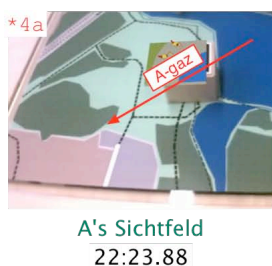
Kontrastierend zur face-to-face-Bedingung, fällt hier die zeitliche Dauer der voneinander getrennten Aktivitäten auf. Vom Zeitpunkt der initialen Abwahl der Handlungsrelevanz „PZ“ – markiert durch B's erste Aufforderung zur Platzierung⁵⁰ und Orientierungsverlagerung auf die übrigen Objekte am Kartenrand um ca. 22:14 min (vgl. Fragment 15) – vergehen knapp acht Sekunden, bis A den parallelemergenten Handlungsanspruch erkennen kann (22:22 min). Die Möglichkeit der initialen „Störabwehr“ ist damit annulliert, da die physische Präsenz des Objekts „BBQ“ im unmittelbaren Handlungsraum akute Zugzwänge konstituiert, die prozessual mit der gegenwärtigen Handlungsplanung „PZ“ konfligieren.

Der Konflikt zwischen der fortgeschrittenen Handlungsplanung und der Einlösung konditioneller Relevanzen manifestiert sich sequentiell nachfolgend. Simultan zur angekündigten Handlungsplanung A's („und ähm:“) formuliert B eine dritte, konkretisierte Aufforderung zur Platzierung des Objekts: „SETZEN sie das DA: =das können=se gerne auf ne GRÜNfläche setzen;“ (02+03). Seine Aufmerksamkeitsorientierung auf das in A's ruhender Greifhand befindliche Objekt „PZ“ (*3b) lässt erkennen, dass sich die verbaldeiktische Angabe („DA:“) seiner Handlungsaufforderung auf die als Wald ausgewiesene Kartenfläche beziehen soll, auf der das Objekt bereits halbseitig steht (*3b). Dieser Handlungsaufforderung steht das Voranschreiten der eigenen Planungsaktivität A's gegenüber: Simultan zur dritten Platzierungsaufforderung sucht er selbst nach einem geeigneten Platzierungsort (*4a) und stellt das Objekt vorerst – gleichsam emblematisch zum gegenwärtig indefiniten Verhandlungsstatus – an den Kartenrand (*4b).

```

03 A-ver:      |<<p>das is mir relativ->
A-lH:  .....place@loc-1|((leichtes Anheben & Manipulieren))
A-rH:  HP.....grasp>>>>>@RC.....
B-ver:  auf ne GRÜNfläche setzen; |
B-act:  .....
                                     *5a+b

```



A's eigenes Planungsvorhaben wird aus dem indefiniten Verhandlungsstatus konkretisiert: Vom Kartenrand platziert er das Objekt „PZ“ zurück auf die gleiche Stelle (*5a), formuliert „autosuggestiv“ einen nicht näher bestimmbar oder interaktiv relevanten Gedankengang (<<p>das is mir relativ->) und stellt durch die rechthändige Manipulation des Objektes „RopesCourse“ (RC) visuell eine Relevanz zwischen den Objekten her (*5a+b).

```

04 A-ver:  also HIER ham wa=ja den |KLETTERPARK im WALD °hh
A-lH:  @PZ.....
A-rH:  ...lift>>>>>>>>place-back>>>HP.....
B-ver:  |also wir ham uns ja EBEN darau-
B-act:  ...place>>>>>>>>@map...|>>>>>>>>>>>>>>>>>>HP.....
                                     *6a+b                                     *7

```

⁵⁰ „DA: bin ich auch flexIBEL=die könnse irgendwo HINsetzen=das is mir egal=das is kein KOSTENfaktor“ (vgl. Fragment 15)

Aus der „regulären Rederechts-/Handlungsübernahme“ spezifiziert B seinen Vorschlag: Im „pre-proposal account“ verweist er auf einen bereits als „Park“ ausgehandelten Bereich („loc-2“) und führt emblematisch – durch eine den Objektumriss skizzierende Verweisgeste (*8a+b) – den Vorschlag ein, dort einen Grill aufzustellen (07+08). Aus der Konfiguration des beobachtenden Akteurs kann A den Orientierungshinweisen seines Ko-Partizipanten folgen (*8a+b). Erst nachdem B den vorgeschlagenen Platzierungsort markiert hat, greift er das in der Parallelemergenz der vorigen Handlungsaktivitäten suspendierte Objekt „Barbecue“ (BBQ) wieder auf und führt es in Ko-Orientierung zum vorgeschlagenen Platzierungsort (*9-10,*10).

2.4.3 Zusammenfassung

Aus den empirischen Beobachtungen zur Emergenz (2.4.1) und Auflösung (2.4.2) konkurrierender Aktivitäten in AR lassen sich strukturelle Unterschiede im kontrastiven Vergleich zur face-to-face Bedingung ableiten (vgl. Abbildung 14). Im Rahmen parallelemergenter Rederechts-/Handlungsansprüche ist für die face-to-face Bedingung eine Orientierung an Diskontinuitäten in der kontemporären Aufgabenbearbeitung feststellbar. Neue Handlungsinitiativen (z.B. durch die Auswahl eines neuen Objekts) erfolgen aus einem Progressionsanspruch, der mit der wahrnehmbaren Diskontinuität konfligiert und pauschalisiert als Verhaltensregularität „Wenn mein Ko-Partizipant nicht liefert, bin ich an der Reihe“ zusammengefasst werden kann. Faktisch handelt es sich jedoch eher um eine situative Abwägung, in der die den Handlungsanspruch stellende Partei eine mangelnde Zielrichtung im Verhalten des Gegenüber erkennt und von einer realen Chance ausgehen kann, sich selbst durch ein zielgerichtetes Handeln als legitimer „Sprecher“ zu etablieren. Der Terminus „zielgerichtet“ ist an dieser Stelle nicht mit „intentional“ gleichzusetzen. Unter Verweis auf die empirischen Erkenntnisse zur funktionalen Simultaneität (vgl. 2.3.8) verstehen wir zielgerichtetes Verhalten auf praxeologischer Ebene als eine auf dem Prinzip der Teilsimultaneität beruhende visuelle Darstellungsoption, in der einzelne Ausdrucksressourcen durch sukzessive Nominierung in „sequentiellen Lesarten“ konnektiert werden. Im Gegensatz dazu ist im diskontinuierlichen Verhalten eine sequentielle Lesart, in der eine Zielrichtung erkennbar wird, meist nicht gegeben.

Diskontinuitäten in einer gegenwärtigen Aufgabenbearbeitung können dabei in ihrer multimodalen Formgestalt unterschiedlich konstituiert sein. Die empirischen Analysen der face-to-face Bedingung verweisen beispielsweise auf mangelnde verbalsprachliche Projektionen von Fortsetzungsoptionen (z.B. Fragment 4), trajektorielle Diskontinuitäten, in denen eine projizierte Zögerlichkeit wahrnehmbar wird (z.B. Fragment 7, Fragment 8) oder auf Pausen und Hesitationsvokalisationen, die dem jeweiligen Ko-Partizipanten eine Planungsmöglichkeit eröffnen, einen parallelemergenten Handlungsanspruch semantisch an die kontemporäre Aufgabenbearbeitung anzuschließen (z.B. Fragment 9) (vgl. 2.3.9, „Emergenzbedingung: Diskontinuitäten“). Nicht immer liegen Diskontinuitäten jedoch im Verhalten des gegenwärtig legitimen „Sprechers“ begründet. Während trajektorielle Diskontinuitäten etc. sowohl aus einer etischen wie einer emischen Perspektive beobachtbar sind, lassen sich auch Fälle identifizieren, in denen Diskontinuitäten eine interpretative Dimension einnehmen, in der die Anwartschaft auf das Rede-/Handlungsrecht aus einem Verkürzungsanspruch erfolgt. Der gegenwärtigen Interaktionsaufgabe wird in diesen Fällen eine generelle oder situative Bearbeitungsrelevanz abgesprochen (z.B. Fragment 11).

Zentral erscheint uns der Hinweis, dass ein Anspruch auf das Rede-/Handlungsrecht nicht „erschlichen“, sondern ihm unter den situativen Bedingungen der Diskontinuität und Voraussetzung der Wahrnehmbarkeit eine reelle Chance eingeräumt wird. Empirische Hinweise darauf lassen sich aus den Reaktionen auf einen parallelemergenten Handlungsanspruch ableiten. Fälle, in denen ein gegenwärtig legitimer „Sprecher“ eine „Störabwehr“ initiiert, bevor sich der parallelemergente Handlungsanspruch als faktisch störend erweist (z.B. lediglich eine Fokusverlagerung und/oder Manipulation eines Objekts vorliegt, z.B. Fragment 8), deuten auf ein implizites Wissen, dass Simultanaktivitäten in diskontinuierlichen Aufgabenbearbeitungen einen Anspruch auf die legitime „Sprecherrolle“ darstellen. Hier zeigt sich, dass sich sowohl die das Rede-/Handlungsrecht beanspruchende Partei wie auch der gegenwärtig „legitime Sprecher“ selbst an situativen Diskontinuitäten orientieren. Die beiderseitige Orientierung ist in Abbildung 14 durch bidirektional gerichtete Pfeile dargestellt. Aus der gemeinsamen Orientierung erfolgt die „overlap resolution“, in der die Interaktanten die konkrete interaktionale Fortsetzung durch unterschiedliche Sequentialisierungsverfahren aushandeln (in der Abbildung als disjunkte „oder“-Verknüpfung dargestellt) (vgl. 2.3.9).

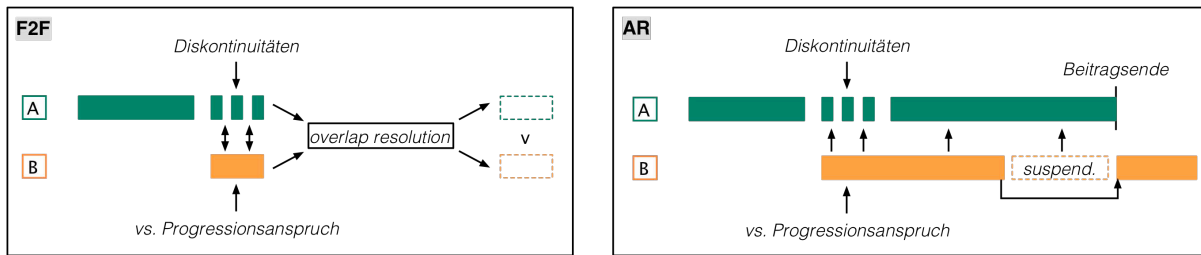


Abbildung 14: Emergenz & Auflösung konkurrierender Aktivitäten im kontrastiven Vergleich (F2F vs. AR)

Kontrastierend verweisen die empirischen Beobachtungen zur AR Bedingung nicht nur auf strukturelle Unterschiede in der Emergenz und Auflösung konkurrierender Aktivitäten, sondern liefern ebenso eine theoretische Basis für die analytische Bewertung der rekonstruierten Verfahren der face-to-face Bedingung. Im direkten Vergleich der Bedingungen wird deutlich, dass sich die Parallelemergenz von Handlungsaktivitäten zwar punktuell als Störfaktor ausprägen kann (insbesondere wenn die parallelemergenten Aktivitäten auf gleicher Fokus- und Modalitätsebene konkurrieren), die Interaktanten bei intaktem mutual monitoring jedoch auf zahlreiche Verfahren zur situativen Desambiguierung potentieller Fortsetzungsoptionen zurückgreifen können, die es Ihnen ermöglichen, Unklarheiten der interaktionalen Fortsetzung instantan oder mit minimaler Verzögerung zu beseitigen. In diesem Sinne bilden konkurrierende Handlungsaktivitäten lokale Beeinträchtigungen ohne globale Störimplikationen. „Probleme“ entstehen vielmehr, wenn parallelemergente Handlungsansprüche nicht erkannt werden, wie sich anhand der AR Bedingung zeigen lässt.

Das Auftreten parallelemergenter Rederechts-/Handlungsansprüche ist hier ebenfalls an situative Bedingungen der Diskontinuität geknüpft. Im aufgeführten Fallbeispiel (Fragment 15 (Emergenz)/ Fragment 18 (Auflösung)) hatte sie eine interpretative Dimension, in der die Perpetuierung der kontemporenen Interaktionsaufgabe in ihrer Handlungsrelevanz abgestuft wurde. Im Unterschied zur face-to-face Bedingung verweist das Beispiel jedoch auf eine einseitige Orientierung an Diskontinuitäten (vgl. Abbildung 14, AR: markiert durch unidirektional gerichtete Pfeile), da der gegenwärtige „Sprecher“ durch die Fokussierung der eigenen Aufgabenbearbeitung den Rederechts-/Handlungsanspruch des Ko-Partizipanten nicht in seiner holistischen Gestaltform wahrnehmen kann (zu divergenten Teilnehmerinterpretationen s.a. Fragment 21, Fragment 22). Im Hinblick auf Goodwins Konzeption der „contextual configurations“ (2000a, 2003) ist eine teilrezeptive Äußerungsinterpretation fehleranfällig, da die holistische Äußerungsgestalt in ihrem Zusammenwirken deutliche Projektions- und Bedeutungsdivergenzen zu ihren Einzelkomponenten aufweist⁵¹ (Goodwin 2011). Im vorliegenden Fall ließ sich feststellen, dass die konditionelle Relevanz zur Aufgabenbeendigung nicht initial, durch die multimodale Gestalt aus verbalsprachlicher Aufforderung, Aufmerksamkeitsverlagerung und Auswahl eines neuen Objekts erzeugt wurde, sondern erst in der interaktionalen Fortsetzung durch repetitive Verbalanweisungen hergestellt werden konnte („DA: bin ich auch flexibel =die können=se irgendwo HINsetzen [...] das können=se irgendWO mal HINsetzen SETZEN sie das DA:=das können=se gerne auf ne GRÜNfläche setzen;“).

Das Problem, das aus einer verzögerten Wahrnehmung entstehen kann, lässt sich wie folgt beschreiben: Beide Interaktanten setzen ihre individuellen Aktivitäten in unterschiedlicher Ausprägung fort (A in sequentieller Abkopplung, B unter fortwährender Bearbeitung konditioneller Relevanzen der „Primäraufgabe“ (markiert durch unidirektional gerichtete Pfeile, vgl. Abbildung 14)). Mit dem Fortschritt der Individualaktivitäten wird die Möglichkeit einer „Störabwehr“ oder der neuen Aushandlung des Rede-/Handlungsrechts aufgehoben, da ein geeigneter Zeitpunkt (Orte der Diskontinuität) durch die Progression der Aktivitäten nicht mehr gegeben ist. Im vorliegenden Beispiel erzeugte die physische Präsenz des in den Handlungsraum geführten Objekts akute Zugzwänge, die prozessual mit der Handlungsplanung des gegenwärtig legitimen „Sprechers“ konfligierten. Der Konflikt, der sich in erhöhten koordinativen Anforderungen manifestiert, konstituiert sich somit zwischen einem Fortschreiten der auf die gegenwärtige

⁵¹ Ergebnisse aus den AR-Studien zeigen in diesem Zusammenhang, dass die funktionale Relation der Komponenten nicht immer eine obligate Prämisse zum Verständnis des Handlungsvollzugs bildet. Interaktanten sind unter bestimmten kontextuellen Bedingungen in der Lage, teilrezipierte Äußerungsgestalten zu rekonstruieren, ihren semantischen Gehalt zu inferieren und die erzeugte Präferenzstruktur sinngemäß zu bearbeiten (vgl. 3.2).

Interaktionsaufgabe bezogenen Handlungsplanung und den sich tradierenden Zugzwängen eines parallelemergenten Handlungsanspruchs des Ko-Partizipanten. Die Komplexität der Situation erscheint dabei schwer auflösbar. Im Fallbeispiel wurden die Zugzwänge „ausgehalten“, woraufhin B – in Reaktion auf die ausbleibende Bearbeitung seiner Aufforderungen – den parallelemergenten Handlungsanspruch suspendierte (vgl. Abbildung 14) und das Objekt vorläufig abstellte. Erst nach Beendigung der kontemporären Interaktionsaufgabe konnte der Handlungsanspruch wieder aufgenommen werden.

Im kontrastierenden Vergleich der Bedingungen (F2F vs. AR) wird deutlich, dass konkurrierende Handlungsaktivitäten oder emergierende Handlungsansprüche auf unterschiedlicher Fokusebene nicht ausschließlich als Störfall oder Bedrohung des gegenwärtigen Rederechts zu bewerten sind (“minimize overlap”), sondern je nach Diskontinuitätsausprägung bei intaktem mutual monitoring legitime Möglichkeiten darstellen, Handlungsansprüche an geeigneter Stelle im interaktionalen Geschehen zu stellen und diese kooperativ auszuhandeln. Unter der verzögerten Wahrnehmung im eingeschränkten “mutual monitoring” hingegen erscheint das Potential der kooperativen Aushandlung konkurrierender Aktivitäten gehemmt.

Interpersonelle Koordinierungs- und Verständigungsprobleme?

Im Vergleich zur F2F-Bedingung verweisen parallelemergente Handlungsansprüche in der AR-Bedingung auf zeitlich verzögerte und gesteigerte Koordinierungsanforderungen konkurrierender Aktivitäten, da die Beschränkung des “mutual monitoring” die Emergenz und Fortsetzung individueller Handlungsvollzüge – in unterschiedlich ausgeprägter sequentieller Entkopplung voneinander – begünstigt. Aus dieser empirischen Beobachtung ließe sich die Hypothese ableiten, dass interpersonelle Koordinierungs- und Verständigungsprobleme den konstitutiven Normalfall der AR-Bedingung bilden und sich die Einschränkung des mutual monitoring in unterschiedlichen Störausprägungen der Bearbeitung interaktionaler Aufgaben abzeichnet. Die analytischen Ergebnisse des nachfolgenden Kapitels (Kapitel 3) liefern jedoch empirische Hinweise auf eine „Robustheit interaktiver Koordinierungsressourcen“, die auch unter den erschwerten Interaktionsbedingungen die Möglichkeit zur kooperativen Bearbeitung interaktionaler Aufgaben weitestgehend sicherstellen und ein „Zusammenbrechen“ des Interaktionssystems verhindern.

3 Robustheit interaktiver Koordinierungsressourcen

Die empirischen Beobachtungen zur Emergenz und Auflösung konkurrierender Aktivitäten im kontrastiven Vergleich der AR- und F2F-Bedingung (vgl. 2.4) haben ein phänomenologisches Unterscheidungskriterium hervorgebracht, das die interaktionalen Auswirkungen des eingeschränkten *mutual monitoring* beschreibt: Während Interaktanten unter der Persistenz gegenseitiger Beobachtung in der F2F-Bedingung über Verfahren verfügen, die es ihnen bereits im initialen Stadium ermöglichen, konfligierende Simultanaktivitäten in sequentielle Ordnungsstrukturen zu überführen, lässt sich für die AR-Bedingung feststellen, dass die Partizipanten nur vermindert fähig sind, ihre Aktivitäten kleinschrittig zu koordinieren, um emergierender Handlungsparallelität entgegenzuwirken. Eine sich verlagernde Orientierung des Rezipienten bildet unter den restringierten Bedingungen im *mutual monitoring* kein stabiles Signalisierungssystem, das dem gegenwärtigen Sprecher erlaubt, Reparaturen zur Rückgewinnung der Aufmerksamkeit einzuleiten. Infolgedessen emergieren Handlungsparallelitäten, in denen individuelle Handlungsvollzüge episodisch – in unterschiedlich ausgeprägter sequentieller Entkopplung voneinander – das situative Handlungsgeschehen bestimmen.

Auch wenn die empirischen Beobachtungen zur Emergenz sequentiell entkoppelter Handlungsparallelitäten ein phänomenologisches Charakteristikum einer Interaktionsbedingung bilden, in der die Interaktanten ihre jeweiligen Handlungsaktivitäten nicht kontinuierlich beobachten und feingraduell aufeinander abstimmen können, präsupponieren sie in gewisser Weise interpersonelle Koordinierungs- und Verständigungsprobleme als konstitutiven Normalfall der AR-Bedingung. Wie wir nachfolgend zeigen werden, ist dies jedoch nicht der Fall. Die skizzierten Störausprägungen, in denen sich teilnehmerrelevante Probleme in der Bearbeitung interaktionaler Aufgaben abzeichnen, manifestieren sich eher in singulären Randphänomenen, die jedoch nicht den konstitutiven Normalfall beschreiben. Vielmehr lässt sich beobachten, dass die Partizipanten lokale Interaktionsaufgaben kooperativ bearbeiten und das Handlungsgeschehen mit Blick auf die globale Aufgabenbearbeitung sequentiell und interaktiv vorantreiben. Aus dieser empirischen Beobachtung leitet sich das zugrundeliegende Forschungsdesiderat des vorliegenden Kapitels ab, in dem auf Grundlage systematischer Analysen die folgenden Forschungsfragen adressiert werden: Mit welchen multimodalen Verfahren bearbeiten die Interaktanten die gesteigerten Koordinierungs- und Orientierungsanforderungen der AR-Bedingung? (Wie) tragen die verwendeten Verfahren dazu bei, Distorsions- und Störeffekte zu minimieren? Welche Rückschlüsse lassen sich daraus auf die Robustheit interaktiver Orientierungs- und Koordinierungsressourcen/-verfahren im Speziellen und interaktionaler Systeme im Allgemeinen ziehen?

Das Themenfeld ‚interaktionale Robustheit‘ bildet einen Untersuchungsgegenstand, zu dem bislang kaum empirisch fundierte Erkenntnisse vorliegen. Während im Rahmen der interaktionalen Modellierung von Agenten- und Robotersystemen der Begriff zum Teil recht inflationär gebraucht wird, um auf die Zuverlässigkeit und das ‚Funktionieren‘ der implementierten Verfahren im „hybriden Austausch“⁵² (Krummheuer 2010) zwischen Mensch und Maschine zu verweisen, liegen auch in der konversationsanalytischen Forschung nur vereinzelte Perspektiven vor, die den Gegenstand aus einer etischen Beschreibungsdimension skizzieren. Eine zentrale Perspektive bildet das Prinzip der „Kontingenz“ als epistemische Begründung zur empirisch beobachtbaren Robustheit interaktionaler Systeme:

“Contingency – interactional contingency – is not a blemish on the smooth surface of discourse, or of talk-in-interaction more generally. It is endemic to it. It is its glory. It is what allows talk-in-interaction the flexibility and robustness to serve as the enabling mechanism for the institutions of social life.” (Schegloff 1996b: 22)

Interaktionale Kontingenz bezeichnet somit eine Initialbedingung, die auf die prinzipielle Offenheit und Unbestimmtheit sozialer Interaktion verweist: „There is virtually nothing in talk-in-interaction which can get done unilaterally, and virtually nothing which is thoroughly pre-scripted“ (ebd.). Dieses „auch anders

⁵² Krummheuer verwendet den Begriff „hybrider Austausch“, um auf die Diskrepanz zwischen menschlicher Face-to-Face-Interaktion und mediatisierter Mensch-Maschine-Interaktion am Beispiel des virtuellen Agenten „Max“ hinzuweisen und die Frage, ob es sich bei der letzteren um eine soziale Interaktionsform handelt, als empirische Frage ‚unvoreingenommen‘ stellen zu können: „Der hybride Austausch zeichnet sich im Gegensatz zur Face-to-Face-Interaktion durch seine Zweiweltlichkeit aus. Es sind zwei unterschiedliche Wesen (Agent und Mensch), die sich in unterschiedlichen Welten [virtuelles Bildschirm-bild, anfassbare, physikalische Welt im Kaufhaus] befinden.“ (Krummheuer 2010: 318f.)

möglich sein“, wie es Luhmann (1984: 47) umschreibt, konstituiert eine grundlegende Flexibilität im interaktiven Handlungsvollzug, in der die genuine Vielfalt an Fortsetzungsoptionen robuste interaktionale Strukturen schafft. Neben dem Kontingenzprinzip wird interaktionale Robustheit auch in den Kontext hochdifferenzierter Verfahren zur Reparatur interaktionaler Probleme gestellt (z.B. Deppermann 2010):

„Krisen und Zusammenbrüche (z.B. in Form von Missverständnissen) sind [...] weniger um ihrer selbst willen [...] interessant, sondern weil sie über die im funktionierenden Fall stumm und unsichtbar bleibenden Normalitätserwartungen und die Relevanz unscheinbarer Praktiken Aufschluss geben, die erst im Fall ihres Versagens bzw. Fehlens deutlich werden. Dies heißt nun keineswegs, dass mangelnde Aufmerksamkeit, schlechtes Hören, Missverständnisse, Dissens, Koordinations- und Sprachproduktionsprobleme, strategisches oder provokatorisches Handeln und andere Fälle, in denen die Ordnung der Routine versagt, außerordentlich selten sind. Im Gegenteil, die Robustheit der alltäglichen wie auch institutionellen Interaktion besteht gerade darin, dass für wiederkehrende Probleme auch ebenso routinisiert einsetzbare Reparaturverfahren zur Verfügung stehen, die dazu führen, dass Probleme schnell gelöst und als solche meist gar nicht bewusst wahrgenommen werden (Schegloff, Jefferson & Sacks 1977; Egbert 2009).“ (Deppermann 2010: 651f.)

Interaktionale Robustheit wird hierbei einerseits im Sinne der “deviant case analysis” *ex negativo* als ein Zustand definiert, der sich erst im ‚Versagen‘ und ‚Fehlen‘ unscheinbarer Praktiken manifestiert, die das interaktionale Geschehen im funktionierenden Fall stabil halten. Andererseits gründet die interaktionale Robustheit gerade auch auf der Möglichkeit, rekurrent auftretende Probleme aus der Abwesenheit oder situativen Dysfunktionalität jener subtilen Praktiken durch routinisierte Verfahren der Problembearbeitung zu beheben.

Auch wenn beide Perspektiven (Kontingenz- und Reparaturprinzipien) zentrale Stabilitätskonstituenten der sozialen Interaktion abbilden, haben sie gemein, dass sie das analytisch schwer greifbare Desiderat der interaktionalen Robustheit retrospektiv aus einer systemischen Sicht skizzieren, in der die empirische Robustheit aus zentralen Funktionsmechanismen der Interaktion deduziert wird. Aber wie wirkt sich die interaktionale Robustheit in situ aus? Wie manifestiert sie sich in einzelnen Orientierungsgestalten und der Koordinierung gemeinsamer Orientierung? Wie ‚profitieren‘ die Interaktanten selbst von ihr? Um empirische Hinweise auf diese Fragen zu erhalten, erweitern wir die beiden dargelegten Perspektiven auf interaktionale Robustheit um eine weitere, vollzugsrekonstruktive Perspektive, in der genau jene AR-spezifischen Interaktionsbedingungen, die zur Emergenz sequentiell entkoppelter Handlungsparallelitäten beitragen, methodisch und analytisch gewinnbringend gewendet werden: Während die Okkurrenz multimodaler Gestalten auf der Äußerungsebene unter der gegenseitigen Unterstellung von Ko-Orientierung nicht beeinträchtigt zu sein scheint, ist die Rezeption jener Gestalten durch eine fragmentarische Wahrnehmung gekennzeichnet („Teilrezeption“), in der dem Rezipienten zu unterschiedlichen Zeitpunkten je unterschiedliche semiotische Einzelinformationen aus der Gesamtgestalt zur Verfügung stehen. Insofern ermöglicht das AR-System im analytischen Zugang zu den visuell-semiotischen Ressourcen der Teilnehmerwahrnehmungen, einzelne Ausdrucksmodalitäten der holistischen Gestaltform mit den entsprechenden Reaktionen des jeweiligen Ko-Partizipanten in Verbindung zu bringen und das „reagieren auf“ zu präzisieren und einzugrenzen.

Auf Grundlage dieses „Sezierungsparadigmas“ lassen sich einerseits jene „unscheinbaren Praktiken“ (s.o.) spezifizieren, die im Falle ihrer Nicht-Rezeption zu einer analytisch wahrnehmbaren Divergenz zwischen der Bedeutungskonstitution und –interpretation einer Äußerungsgestalt führen. Andererseits wird in den Analysen deutlich, dass ein „Fehlen“ (s.o) semiotischer Ressourcen und Konstituenten in der holistischen Gestaltform sich nicht unbedingt in revisions- und reparaturbedürftigen Bearbeitungsstrategien manifestiert, sondern im interaktiven Zusammenwirken aus Produktion und Teilrezeption kompensiert werden kann. Diese Kompensation ist dabei keinesfalls mit schlichter Redundanz gleichzusetzen. Im Rahmen von Fragment 4 haben wir analytisch dargelegt, dass die McNeill’schen Regeln zur „semantischen und pragmatischen Synchronie“, in denen Sprache und Gestik koinzident dieselbe Bedeutung ausdrücken (“idea unit”) und dieselbe pragmatische Funktion erfüllen, kein universelles Prinzip bilden. Weder die Sprache noch die Gestik sind begleitende Ausdrucksressourcen der jeweils anderen, sondern konstituierende Ressourcen eines Handlungsvollzugs. In den nachfolgenden Analysen wird vielmehr deutlich, dass ein „Fehlen“ semiotischer Ressourcen und Konstituenten folgenlos bleiben kann, da sie in einem Funktionscluster zwischen gegenseitiger Elaboration (Goodwin 2009; 2010), wechselseitiger Stabilisation und partieller Redundanz zusammenwirken.

Das folgende Kapitel ist in vier empirische Analysebereiche aufgeteilt, denen einleitend ein theoretischer Überblick zum Forschungsfeld der Deixis vorangestellt ist (3.1). Der theoretische Überblick widmet sich der interaktiven Konzeption der Deixis, stellt zentrale analytische und konzeptionelle Beschreibungsinstrumente vor und skizziert die Teilnehmeranforderungen deiktischer Orientierungsverfahren in der face-to-face-Interaktion. Im ersten analytischen Kapitel (3.2) werden diese Teilnehmeranforderungen unter der Spezifik der fragmentarischen Wahrnehmung (Teilrezeption) deiktischer Orientierungsverfahren in der AR-Bedingung untersucht. In der methodischen Anwendung des „Sezierungsparadigmas“ liegt der analytische Fokus auf lokaldeiktischen Ausdrücken und deiktischen Gesten. Die empirisch beobachtbare Robustheit interaktiver Koordinierungsressourcen, die sich in suffizienten Verfahren zur Herstellung von Ko-Orientierung unter den restringierten Bedingungen im *mutual monitoring* manifestiert, wird daraufhin in den beiden nachfolgenden Kapiteln analytisch ausdifferenziert. Zum einen wird im zweiten analytischen Kapitel deutlich (3.3), dass die Interaktanten unter einem gesteigerten Toleranzverhalten agieren, in der die Präferenzorganisation weitaus offener gehandhabt wird als in der F2F-Bedingung. Zum anderen kann im dritten analytischen Kapitel (3.4) gezeigt werden, dass den Teilnehmern ein vielfältiges, semiotisch differentes und multimodales Repertoire an Orientierungsressourcen zur Verfügung steht, das eine stabile Herstellung von Ko-Orientierung ermöglicht. Im abschließenden analytischen Kapitel (3.5) werden die gewonnenen empirischen Erkenntnisse zur interaktionalen Robustheit mit empirischen Beobachtungen aus der AR-Bedingung im Exhibition-Design-Szenario kontrastiert, in der die Interaktionsumgebung durch die Instabilität der Referenzobjekte eine herausfordernde „ecology of sign systems“ (Goodwin 2003a) konstituiert.

3.1 Deixis im Verweisraum – Räumliche Orientierung

“If language is basic to human sociality, deixis is basic to language in its capacity to constitute both subjects and objects.” (Hanks 2005: 191)

Als „klassisches linguistisches Phänomen“ (Mondada 2007c: 85) blickt der Untersuchungsgegenstand „Deixis“ auf eine gut hundertjährige Forschungsgeschichte zurück⁵³ (vgl. Klein & Jungbluth 2002: 6), aus der sich gleichermaßen eine Vielfalt wissenschaftlicher Erkenntnisse wie auch die Schwierigkeit einer all-gemeingültigen Definition ergeben. Die definitorische Lücke resultiert aus einer Uneinigkeit darüber, was dem Phänomenbereich der „Deixis“ zuzuordnen ist (vgl. Stukenbrock 2012: 3). Vor diesem Hintergrund ist es relevant darauf hinzuweisen, dass der vorliegende Überblick zum Forschungsfeld die „interaktive Konzeption der Deixis“ (Mondada 2007c: 85) fokussiert. Eine definitorische Abgrenzung des Phänomenbereichs ist nur insofern relevant, als sie zur Bestimmung eines Basissets deiktischer Teilnehmerverfahren nützlich ist. Unter einer praxeologischen Perspektive erscheint eher eine Definition sinnvoll, die sowohl der Kontextsensitivität als auch der Interaktivität deiktischer Orientierungsverfahren Rechnung trägt:

„Deixis ist [...] nichts Anderes als das Sichtbarmachen des Sehens; die visuelle Wahrnehmung eines bestimmten Details wird selbst unmissverständlich wahrnehmbar gemacht.“ (Hausendorf 2010: 177)

Im Rahmen einer phänomenologischen Fokussierung auf das „Sichtbarmachen des Sehens“ ist der Deixisbegriff (griech. *deiknymi* „ich zeige“) in einem „krass metaphorischen Sinn“ zu verstehen, „in dem er eigentlich jeden Erklärungswert verliert“ (vgl. Klein & Jungbluth 2002: 7). Insofern stellen auch die in 2.3 analysierten Desambiguierungsverfahren zur Auflösung situativer Mehrdeutigkeiten („Fortsetzungsprobleme“) für uns im Kern deiktische Prozeduren dar: Interaktanten konstituieren in teilsimultaner Ausdrucksassemblierung „sequentielle Lesarten“, die einem in der Unklarheitsbedingung „verirrten“ Ko-Partizipanten als perzeptorisches Angebot dienen, einen definiten Übergang zur nächsten Aktivität mitzuverfolgen (z.B. Fragment 4).

⁵³ Ein konziser Überblick zu deutschen und angloamerikanischen Strömungen innerhalb der Deixisforschung findet sich bei Fricke (2007: 13ff.). Den Kernunterschied der Forschungstraditionen beschreibt sie folgendermaßen: „In der deutschen Tradition wird in Anlehnung an Bühler Deiktizität zumeist an den Begriff der Origorelativität gebunden, während in der angloamerikanischen Tradition unter dem Einfluss der sprachanalytischen Philosophie Deiktizität mit Kontextrelativität bzw. Indexikalität gleichgesetzt wird. Auf die angloamerikanische Tradition geht auch die Dichotomie zwischen „deiktisch“ und „intrinsic“ zurück. Als deiktisch werden nur solche Phänomene klassifiziert, bei denen die Origo dem Sprecher zugeordnet ist. Damit liegt in den angloamerikanischen Deixistheorien tendenziell eine starke Sprecherzentriertheit vor“ (ebd.: 17).

Ausgangspunkt einer solchen interaktiven und perzeptorischen Deixisauffassung bildet die in Bühlers „Sprachtheorie“ (1934) formulierte Konzeption der *Zweifelderlehre*, der *Origo* und der deiktischen Dimensionen *Raum* (Lokaldeixis), *Zeit* (Temporaldeixis) und *Person* (Personaldeixis) (vgl. Abbildung 15). In seiner „Zweifelderlehre der Sprache“ geht Bühler von einer phänomenologischen Unterscheidung in zwei Kategorien sprachlicher Zeichen aus: den „Symbolen“ und den „Zeigwörtern“. Konzeptuell werden sie in zwei verschiedenen „Umfeldern“ (Bühler 1934/1982: 154) verortet, in denen sie ihre „Bedeutungserfüllung“ (ebd.: 80) erfahren: Im „Symbolfeld“ siedelt Bühler sog. „Nennwörter“ an, die ihre Bedeutung aus konventionalisierten Bezeichnungen⁵⁴ und „dem syntaktischen und lexikalischen Moment der Sprache“ (ebd.: 151) beziehen. Demgegenüber erhalten die im „Zeigfeld“ situierten Wörter ihre Bedeutung „von Fall zu Fall“ und dem, „was das Zeigfeld den Sinnen zu bieten vermag“ (ebd.: 90). Obwohl Bühlers Sprachtheorie in etlichen Punkten von einer interaktionslinguistischen Perspektive auf Kommunikation abweicht (z.B. im „Sender-Empfänger-Paradigma“ und dem Austausch von „Signalen“ (s.a. Bühlers „Organonmodell“)), ist ihr eine interaktionslinguistische Relevanz zuzusprechen: Einer situations- und subjektentbundenen Begriffsbezogenheit im „Symbolfeld“ wird eine dynamische, subjekt- und situationsbezogene Handlungskomponente im „Zeigfeld“ gleichberechtigt gegenübergestellt, die für die nachfolgende Deixisforschung richtungweisend ist.

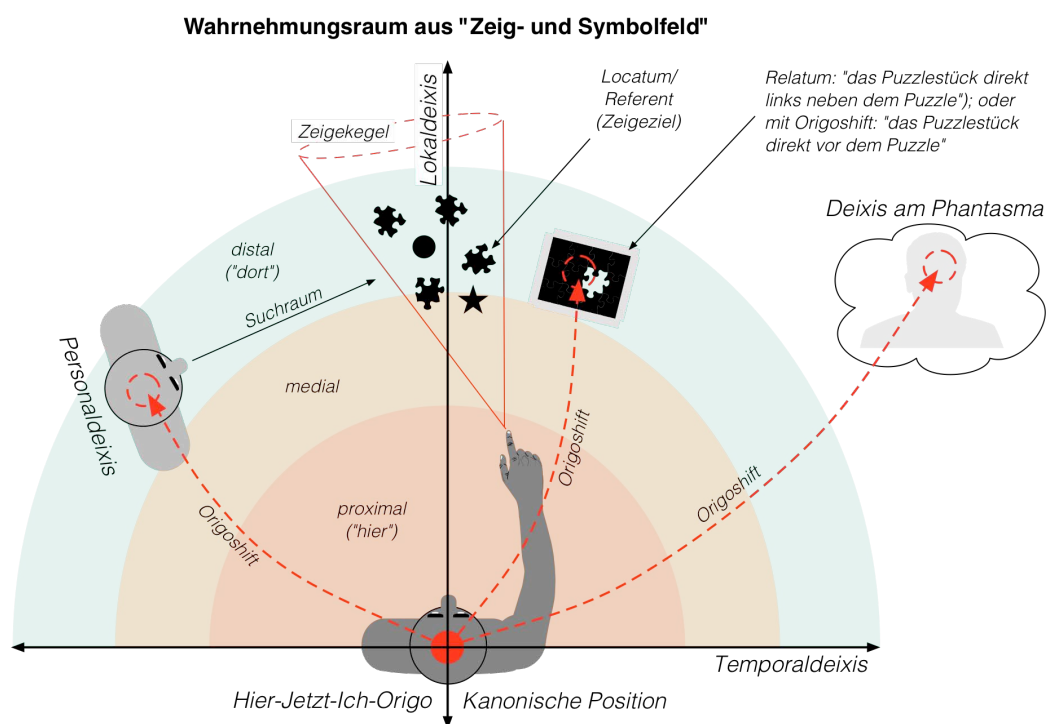


Abbildung 15: Schematische Darstellung der deiktischen Dimensionen basierend auf Bühlers Konzeption der Origo (1934), des „Vektor-Ziel-Problems“ (Stukenbrock 2012) und der Zusatzannahme einer „Kanonischen Position“ der Origo-Instanz (Klein 1990)

Der besondere Stellenwert der „Sprachtheorie“ für die deixistheoretische Konzeption der Folgeforschung manifestiert sich in den organisationalen Dimensionen *Raum* (Lokaldeixis), *Zeit* (Temporaldeixis) und *Person* (Personaldeixis), die Bühler als Konstitutionsbedingung des Zeigfeldes annimmt. Sie bilden ein kartesisches Koordinatensystem im dreidimensionalen Raum, dessen Nullpunkt terminologisch mit dem Begriff der „Origo“ (lat. Herkunft, Ursprung) gefasst und als deiktisches Zentrum der „subjektiven Orientierung“ (ebd.: 102) konzeptualisiert wird.

⁵⁴ Ein Nennwort hat nach Bühler die Eigenschaft, „daß es im Munde jedes und aller als Symbol für *denselben* Gegenstand verwendet wird [...]“. Nennwörter definieren sich aus einer „Zuordnungskonstanz von Sprachsymbolen und Gegenständen“ (Bühler 1934/1982: 103f).

„Zwei Striche auf dem Papier, die sich senkrecht schneiden, sollen uns ein Koordinatensystem andeuten, O die Origo, den Koordinatenausgangspunkt: Ich behaupte, daß drei Zeigwörter an die Stelle von O gesetzt werden müssen, wenn dies Schema das Zeigfeld der menschlichen Sprache repräsentieren soll, nämlich die Zeigwörter *hier, jetzt* und *ich*“. (Bühler 1934/1982: 102)

Zentral ist hierbei die Relativität der Bezugsgrößen (*Raum, Zeit, Person*): Deiktische Orientierungshandlungen sind an ein orientierendes Subjekt gebunden und beziehen sich auf eine konkrete Äußerungssituation – einem Bezugssystem, in dem räumliche und zeitliche Parameter bekannt sind, unterstellt werden können oder unmissverständlich in der Orientierungshandlung selbst vermittelt werden (z.B. durch Lokal- und Temporaldeiktika). Umgekehrt ist aus Sicht des Adressaten die korrekte Interpretation der Bezugsgrößen und ihrer Relation zueinander relevant, um die Orientierungshandlung sinngemäß verstehen und ihrer Anforderung zur Aufmerksamkeitsverlagerung nachkommen zu können. Bühler nimmt dabei die egozentrische Orientierung (Hier-Jetzt-Ich Origo) – die sogenannte „demonstratio ad oculos et ad aures“ – als unmarkierten Standardfall an:

„Das Zeigfeld der Sprache im direkten Sprachverkehr ist das hier-jetzt-ich-System der subjektiven Orientierung; Sender und Empfänger leben wachend stets in dieser Orientierung und verstehen aus ihr die Gesten und Leithilfen der demonstratio ad oculos.“ (Bühler 1934/1982: 149)

Darüber hinaus beschreibt Bühler jedoch auch Formen der Origoversetzung, in der die egozentrische Perspektive (ich-zentriert) zugunsten einer allozentrischen (fremd-zentriert) aufgegeben wird. Konzeptuell werden sie bei Bühler mit den Begriffen „topomnestische Orientierung“⁵⁵ und „Deixis am Phantasma“ gefasst. Während topomnestische Verfahren als Bezugssystem der Orientierung eine gemeinsame Kenntnis über geographische Positionen (bei Bühler „Landmarken“, S. 146) und Richtungen voraussetzen⁵⁶, wie es zum Beispiel in Wegbeschreibungen der Fall ist, beschreibt die Deixis am Phantasma Orientierungsverfahren, in denen auf Entitäten verwiesen wird, die der sinnlichen Wahrnehmung nicht zugänglich sind – also in einem Vorstellungsraum imaginiert werden müssen.

Die Bühlersche Deixiskonzeption ist für ein interaktionlinguistisches Verständnis grundlegend, das deiktische Verfahren nicht auf den „Zeigeakt“ reduziert, sondern als relationales Ausdrucksensemble verschiedener semiotischer endo- und exogener Ressourcen begreift. Diese Komplexität semiotischer Wechselseitigkeit beschreibt Goodwin am Beispiel deiktischer Zeigehandlungen („pointing“):

“Pointing is not a simple act, a way of picking out things in the world that avoids the complexities of formulating a scene through language or other semiotic systems, but is instead an action that can only be successfully performed by tying the point to the construals of entities and events provided by other meaning making resources. Rather than being a stepping stone to language, pointing presupposes it. [...] pointing is constituted as a meaningful act through the mutual contextualization of a range of semiotic resources including at least 1) a body visibly performing an act of pointing; 2) talk which both elaborates and is elaborated by the act of pointing; 3) the properties of the space that is the target of the point; 4) the orientation of relevant participants toward both each other and the space that is the locus of the point; and 5) the larger activity within which the act of pointing is embedded.” (Goodwin 2003b: 2)

Die komplexen Wirkungszusammenhänge – speziell zur interaktiven Konstitution und Deutung von Raumentitäten (“tying the point to the construals of entities”) – lassen sich anhand der schematischen Darstellung in Abbildung 15 veranschaulichen und als interaktionale Aufgaben deiktischer Orientierungshandlungen ausdifferenzieren:

Um einen Ko-Partizipanten auf eine spezifische Raumentität orientieren zu können, muss eine Situation reziproker Zweifachorientierung hergestellt werden, die einerseits aus der gegenseitigen Orientierung der Interaktanten und andererseits aus der gemeinsamen Orientierung in den verweisenden Handlungsraum besteht. Die koordinative Anforderung liegt folglich darin, eine Aktivität als ‚orientierende Aktivität‘ erkennbar zu machen, die sich in ihrer syntaktischen und situativen Einbindung gestalttypologisch von der

⁵⁵ „Topomnestische Bezugssysteme sind Beispiele für allozentrische Bezugssysteme [...], und zwar solche, die auf Gedächtnisrepräsentationen der Umgebung beruhen. Die Fähigkeit zu Wegbeschreibungen hängt wesentlich von der allozentrischen Repräsentation der Umgebung und deren Aktivierung in der Beschreibungssituation ab, jedoch auch von der Beherrschung der Mittel zur sprachlichen Umsetzung (vgl. Klein, 1983)“ (Vorweg 2001: 82).

⁵⁶ Levinson (1996) spricht hier von einem „absoluten Bezugssystem“.

vis-à-vis Situation abgrenzt. Derartige Mehrfachorientierungen zwischen Sprech- und Kontextsituationen und die damit verbundene subjektive Raumstrukturierung und -vermittlung erfordern ein dynamisches Raumkonzept, wie es in neueren Untersuchungen zur multimodalen Raumkonstitution vorgeschlagen wird (z.B. Schmitt & Deppermann 2010; Hausendorf 2010; Goodwin 2003b; Mondada 2007c, 2009). In analytischer Präzisierung der Bühlerschen Unterscheidung zwischen Wahrnehmungs- („demonstratio ad oculos“) und Vorstellungsraum („Deixis am Phantasma“) wird von einer multidimensionalen Raumkonzeption aus „physikalischem Umgebungsraum“, „Interaktionsraum“, „individuellem Verhaltensraum“ und „imaginärem Raum“ ausgegangen, die von den Interaktanten organisiert und aufeinander bezogen werden (vgl. Schmitt und Deppermann 2010: 228). Eine übergeordnete Relevanz kommt dabei der interaktiven Transformation und permanenten Rekonfiguration (i.S.d. „interactive achievement“) eines sozial kodierten Umgebungsraumes in einen sozial elaborierten Interaktionsraum zu:

„In ihrer Interaktion konstituieren die Beteiligten aus dem Umgebungsraum ein für die Zwecke des gemeinsamen Handelns konstituierten sozialen Relevanzrahmen, den wir als „Interaktionsraum“ bezeichnen [...].“
(Schmitt und Deppermann 2010: 228)

Die zentrale Teilnehmeranforderung besteht folglich darin, aus einer visuell ganzheitlich verfügbaren materiellen Umgebungsstruktur, die gleichermaßen Angebote zur Nutzung bereitstellt („Affordanz“ der Raumstruktur, Gibson 1982) wie auch Nutzungsbeschränkungen vorgibt, interaktional relevante und gegenseitig wahrnehmbare „Raumausschnitte“ zu enactieren und sie sukzessiv mit dem Fortschreiten der Bearbeitung lokal anstehender Teilaufgaben zu rekonfigurieren. In diesem Sinne ist der Interaktionsraum sowohl ein „orientierter Raum“ als auch ein „prozessualer, aktivitätskonstituierter Raum, der nicht unabhängig vom Handeln der Beteiligten ‚in ihm‘ existiert [...]“ (ebd.: 228f.).

Die Herstellung eines interaktional relevanten „Raumausschnitts“ bildet jedoch lediglich eine Voraussetzung zur referentiellen Orientierung auf eine in ihm befindliche Raumentität. Der Orientierungserfolg hängt hier im Wesentlichen davon ab, ob die eingesetzten referentiellen Verfahren (z.B. Benennungen, deiktische Gesten, lokale deiktische Ausdrücke etc.) eine Angemessenheit in Bezug auf die materielle Raumstruktur aufweisen, Relationen innerhalb dieser eindeutig abbilden und eine perspektivische Wahrnehmung auf die Raumsituation vermitteln (ego- oder allozentrisch). Im Rahmen der Verwendung von Zeigegesten liegt dabei nur selten eine Salienzkonfiguration vor, in der das Zeigeziel a priori bestimmt werden kann – wie in etwa bei einem vorbeifliegenden Kranichzug, der sich kontextuell als ‚Ereignis‘ vom Hintergrund absetzt. Vielmehr weist die distale Orientierung durch Zeigegesten ein multireferentielles Angebotsprofil auf, das bei Stukenbrock als „Vektor-Ziel-Problem“ konzeptualisiert wird. Danach bilden Zeigegesten keinen „punktgenauen Vektor [...], dessen imaginäre lineare Verlängerung automatisch zum Zeigeziel führt“, sondern verweisen wie ein „Scheinwerferkegel“ auf einen „Suchraum“⁵⁷, in dem das Zeigeziel vom Ko-Partizipanten erst aufgefunden werden muss (Stukenbrock 2012: 282; s.a. Goodwin 2000b: „domain of scrutiny“⁵⁸; Kranstedt et al. 2006: „pointing cone“; Rieser 2004: „region-pointing“).

Zeigegesten sind infolgedessen in der Regel auf zusätzliche Orientierungshinweise angewiesen. Dabei ist selbst der vermeintlich einfache Fall – die ‚Benennung‘ eines Objektes – nicht immer trivial, da keine invarianten Zuordnungen zwischen ‚Wörtern‘ und ‚Bedeutungen‘ bestehen. Am Beispiel einer umweltpolitischen Diskussion zeigt Deppermann (2001), wie Interaktanten durch „metasemantische Handlungen“ die Ausdrucksbedeutung kontextsensitiv und lokal konstituieren und so „Wörtern Interpretationen verleihen“ (ebd.: 73; s.a. Herrmann (1994): „Wort-Konzept-Relation“⁵⁹; Brennan & Clark (1996): „conceptual pacts“). Oftmals werden jedoch auch raumreferentielle Ausdrücke zur Orientierung verwendet, die sich relationaler Abbildungen der materiellen Raumstruktur bedienen („links/rechts neben...“) und eine Perspektivierung der Orientierung vermitteln – sei es explizit („von dir aus...“) oder, wie es von Bühler (1934/1982: 149) oder auch Stukenbrock (2012: 6) für die egozentrische Orientierung angenommen wird, implizit unterstellend („Wenn die Perspektivierung nicht expliziert wird, gilt das System der subjektiven Orientierung“). Die koordinativen Anforderungen in der Verwendung und Rezeption raumreferentieller

⁵⁷ „Während der *Verweisraum* eine die koordinative Selbstwahrnehmung betreffende sowie die Aufmerksamkeitsorientierung des Adressaten kalkulierende Angelegenheit des Zeigenden darstellt, ist der *Suchraum* eine perzeptorische Angelegenheit des Adressaten“ (Stukenbrock 2012: 71, Hervorhebung im Original).

⁵⁸ “[...] rather than transparently locating a **target**, a point specifies a **domain of scrutiny** a region where the addressee should begin to search for something that might count as a target” (Goodwin 2000b: 73, Hervorhebung im Original).

⁵⁹ „Bedeutung ist keine Eigenschaft eines Wortes, sondern seine Relation zu Begriffen.“ (Herrmann 1994: 7)

Ausdrücke stehen somit – nicht in ausschließlicher, jedoch gravierender Weise – in Abhängigkeit zum Komplexitätsgrad der materiellen Raumstruktur und ihrer konkreten situativen Nutzung. Dieser Aspekt, aus dem sich unterschiedliche ‚Orientierungsqualitäten‘ raumreferentieller Verfahren ableiten lassen, kann abschließend anhand unserer schematisierten Beispielsituation verdeutlicht werden (vgl. Abbildung 15).

Vor dem Hintergrund, dass die lineare Extrapolation der Zeigegeste nicht automatisch auf den Referenten verweist (Stukenbrock 2012; Goodwin 2000b), stellt die Benennung des Zeigezieles ‚Puzzlestück‘ aufgrund der Vielzahl an Puzzlestücken eine unterspezifizierte Verweisform dar. Durch die situative Ambiguität der Raumstruktur ist dieses Orientierungsformat problematisch. Aus der subjektiven Orientierung des verweisenden Akteurs in „kanonischer Position“⁶⁰ (Klein 1990, 1991, 1994) wäre ebenfalls eine lokaldeiktische Raumreferenzierung möglich, die sich des eigenen Körperschemas bedient („das rechte Puzzlestück“) oder zusätzlich gestalttypologische Prinzipien angesichts der rautenförmigen Anordnung der Puzzlestücke zugrunde legt („das Puzzlestück rechts Vorne“). Eine hierbei zu berücksichtigende Problematik besteht jedoch darin, dass die dem Verweis zugrundegelegte Körperorientierung in der Wechselseitigkeit des Blickaustausches zwischen verweisendem und rezipierendem Akteur kein stabiles Bezugssystem bildet. Daraus ergibt sich eine rezipientenseitige Anforderung, in der gewissermaßen eine ‚Basisorientierung des Verweisgebers‘ interpoliert werden muss, um der Raumreferenz Sinnhaftigkeit zu verleihen. Weitere Möglichkeiten der Referenzherstellung resultieren aus den Wahloptionen unterschiedlicher Relata („das Puzzlestück direkt links neben dem Puzzle“ vs. „das Puzzlestück zwischen dem Ball und dem Puzzle“) und diversen Versetzungsoptionen der Origo-Instanz. Eine dieser Optionen, in der die Interdependenz zwischen materieller Raumstruktur im „physikalischen Umgebungsraum“ und raumreferentieller Handlung im „Interaktionsraum“ besonders anschaulich wird, ist die Origoversetzung in ‚unbelebte‘ Raumentitäten: Während „das Puzzlestück rechts vom Kreis“ aufgrund der fehlenden intrinsischen Orientierung eines Kreises auf potentiell alle Puzzlestücke verweist, verweist die Orientierungsangabe „rechts vom Stern“ deutlich präziser auf das Zeigziel, obgleich einem ‚Stern‘ ebenfalls keine intrinsische Orientierung zugeordnet werden kann. An diesem Beispiel wird erkennbar, dass das ‚Vorne‘ und ‚Hinten‘ des Objekts durch die Sprachverwendung selbst erzeugt wird, in dem der Stern in Relation zu einem relevanten (die übrigen Objekte) und irrelevanten Raumbereich (ohne Objekte) gesetzt wird. Insofern besteht eine zentrale Anforderung bei deiktischen Verweishandlungen darin, eine kontextuelle Situation mit seiner Raum-, Sprach- und Beteiligungsstruktur zu kategorisieren und in Abstimmung zur gegenwärtigen Rezipientenorientierung abzuwägen, welches Orientierungsverfahren die situativen Bedingungen adäquat aufnimmt.

3.2 Die Relevanz lokaldeiktischer Ausdrücke („hier“) und Rekonstruktion deiktischer Gesten bei der Herstellung von Ko-Orientierung

Um die visuelle Wahrnehmung eines bestimmten Details im Wahrnehmungsraum sichtbar zu machen (vgl. Hausendorf 2010), verwenden Interaktanten referenzielle Verfahren, mit denen die Aufmerksamkeit des Ko-Partizipanten gebunden und auf einen spezifischen Verweisort gelenkt werden kann. Als Voraussetzung zur Herstellung von Ko-Orientierung sind sie grundlegend, um einen an sich intrinsisch-kognitiven Wahrnehmungsprozess wahrnehmbar und damit interaktional anschlussfähig zu gestalten. Ein gängiges Verfahren der Referenzherstellung im sinnlich erfahrbaren Handlungsraum stellen deiktische Zeigegesten dar, die üblicherweise aus den sich gegenseitig elaborierenden semiotischen Ressourcen [Körperpositur] + [deiktische Geste] + [Sprache/Deiktika] + [Blick] + [Kontext/Rahmenaktivität] konstituiert werden (vgl. 3.1, Goodwin 2003b).

⁶⁰ „Die Vertikale ist durch die kanonische Kopf-Fuß-Orientierung gegeben, die Horizontale durch die Körperseiten-Asymmetrie (links ist, wo bei den meisten das Herz ist)“ (Klein 1991: 95). Die Relevanz dieser Zusatzannahme kommt in Beispielen zum Ausdruck, die darauf hinweisen, dass nicht „objektive“ (z.B. Schwerkraft, Himmelsrichtungen), sondern „subjektive Raumeigenschaften“ für die Raumreferenz maßgeblich sind: „Nicht auf die Körperposition bezogene Raumeigenschaften können intervenieren. So läßt sich die Vertikale auch durch die Schwerkraft definieren. Unten ist demnach, wohin die Steine fallen, und oben, woher sie kommen, wenn sie fallen. Normalerweise entsprechen sich nun der Ort, wohin die Steine fallen, und der Ort wo die Füße sind. Aber das muß nicht sein; im Schwimmbad, im Bett und in der Schwerelosigkeit sind sie dissoziiert. In diesen Fällen setzt sich in der Regeln die Körperorientierung durch; unten im Bett ist, wo die Füße zu sein pflegen; [...]“ (Klein 1990: 23).

Unter der Sichtfeldeinschränkung von 42° in der AR-Bedingung (anstelle der annähernd 180° natürlicher Perzeption), entstehen Situationen, in denen die Interaktanten die visuell-semiotischen Ressourcen einer Orientierungshandlung in unterschiedlich ausgeprägter Fragmentierung wahrnehmen. Die gegenseitige Orientierung am Blickverhalten des Gegenüber fällt dabei als semiotische Orientierungsressource durch das Tragen von Head-Mounted-Displays gänzlich weg. Dementsprechend lassen sich nur selten Situationen beobachten, in denen Blickkontakt, bzw. aufgrund der Augenverdeckung im wörtlichen Sinne face-to-face-Kontakt hergestellt wird. Ausnahmen bilden insbesondere die Interaktionseröffnung, gemeinsames Lachen, persönliche Angriffe und „metaperspektivische“ Problemaushandlungen, die keiner unmittelbaren Objektmanipulation bedürfen. Die Sichtfeldeinschränkung bedingt dabei auch eine Einschränkung an Reaktivität, da der Blick zum Ko-Partizipanten nicht erwidert werden kann, wenn dieser im Rahmen der Aufgabenfokussierung auf die Karte orientiert ist. Die hochfunktionalen Mechanismen des „peripheral monitoring“ (Heath et al. 1995), wie wir sie im Rahmen der Auflösung simultaner Aktivitäten in der face-to-face Bedingung analytisch dargelegt haben (2.3.6), sind hier außer Kraft gesetzt. Das ausbleibende körperliche Feedback mag ein weiterer empirischer Grund dafür sein (neben der strikten Aufgabenfokussierung), dass auch die Kopforientierung des jeweiligen Gegenüber als semiotische Ressource tendenziell kaum genutzt wird. Demgegenüber konstituieren die übrigen visuell-semiotischen Ressourcen – vor allem die Positur des Oberkörpers, die Hand-Arm-Orientierung und deiktische Gesten – oftmals nur fragmentarisch wahrnehmbare Orientierungshinweise.

Am Beispiel von Telekollaborationssystemen haben Luff et al. (2003: „GestureMan“; s.a. Hindmarsh et al. 1998, 2000b: „Furniture-World“) Koordinierungsprobleme aufgedeckt, die auf Inkohärenzen zwischen den mediatisierten Teilnehmerumgebungen zurückzuführen sind („fractured ecology“, vgl. 2.1). Da die Koordinierung der Teilnehmerperspektiven und -positionen im Raum gestört ist, stellt die gemeinsame Referenzherstellung auf Objekte ein zentrales Problem dar:

“Despite the opportunities afforded by the system, the materials reveal the difficulties participants face in accomplishing object-focused actions and the ways in which they attempt to establish and sustain common frames of reference; the system fractures the environments of action and inadvertently undermines the participants’ ability to produce, interpret, and coordinate their actions in collaboration with each other.” (Luff et al. 2003: 53)

Derartige Ko-Orientierungsprobleme, in denen unkoordinierte Teilnehmerorientierungen zu Fehldeutungen und unterschiedlichen Handlungsinterpretationen orientierender Aktivitäten führen, sind innerhalb der AR-Bedingung im Obersee-Szenario jedoch nur äußerst selten zu verzeichnen. Unsere Analysen verweisen vielmehr auf spezifische interaktionale Fähigkeiten, die es den Interaktanten erlauben, lokal anstehende Interaktionsaufgaben kooperativ zu bearbeiten und das Handlungsgeschehen mit Blick auf die globale Aufgabenbearbeitung interaktiv voranzutreiben. Insofern ist es relevant, die Konzeption der „fractured ecology“ (Luff et al. 2003) von den empirischen Beobachtungen einer „fragmentarischen Wahrnehmung“ zu unterscheiden, um sowohl diejenigen Parameter interaktionaler Bedingungen aufzudecken, die neben endogenen Faktoren der interaktionalen Emergenz Koordinierungsprobleme begünstigen als auch diejenigen, die ihnen entgegenwirken können. Obgleich beide Bedingungen (Telekollaboration vs. AR-Kollaboration im „ARbInI System“, 1.2.3) Einschränkungen im „mutual monitoring“ der Interaktanten implizieren, bestehen für die Koordinierung relevante Unterschiede vor allem in der physischen Ko-Präsenz und der definiten räumlichen Dimension sich am „Verhandlungstisch“ gegenüberstehenden Interaktanten in der AR-Bedingung. Im Gegensatz zur Telekollaboration, in der die raumstrukturellen Dimensionen samt ihrer Relationen zwischen Raumentitäten und wechselnden Teilnehmerpositionierungen in vollmediatisierter, haptisch unzugänglicher und perspektivisch verzerrter Weise abgebildet werden, unterliegen diese Parameter in der AR-Bedingung einer perzeptuell unmittelbaren Erfahrbarkeit (haptisch, auditiv, olfaktorisch, m.E. visuell). Entscheidend ist hierbei, dass in der erfahrbaren Definitheit der räumlichen Dimension und Teilnehmerpositionierung fundierte Hypothesen zur Lokalisation von Handlungsaktivitäten des Ko-Partizipanten möglich sind, wenn diese aufgrund der Sichtfeldeinschränkung gegenwärtig nicht wahrgenommen werden.

Unter den geschilderten interaktionalen Bedingungen ist festzustellen, dass die lokaldeiktischen Ausdrücke innerhalb einer Verweishandlung eine wichtige, koordinierungsrelevante Funktion bei der Herstellung einer gemeinsamen Orientierung übernehmen: In Situationen, in denen die Teilnehmer nicht ko-orientiert sind, setzen Lokaldeiktika die Orientierungsverlagerung des Ko-Partizipanten konditionell rele-

vant. Während die Abbildungseigenschaften lokaldeiktischer Ausdrücke für proximale (z.B. „hier“) und distale Relationen (z.B. „dort“) innerhalb unseres Settings eher von untergeordneter Relevanz sind, bildet ihre Funktion als „attention getter“ (Mondada 2012; Jakobson 1971: „shifter“) ein konstitutives Merkmal bei der Herstellung von Ko-Orientierung. Unter den spezifischen Bedingungen physisch ko-präsent und vis-à-vis zum Verweisraum positionierter Akteure liegt der Verwendung lokaldeiktischer Ausdrücke dabei eine zweifache Orientierungslogik zugrunde: Einerseits fungieren sie als ‚zeitliche Orientierungsmarker‘, um auf ‚Wahrnehmung als relevante Interaktionsmodalität gleichsam umzuschalten‘ (Hausendorf 2010: 177). – Andererseits als ‚verweisende Orientierungsmarker‘, die nicht nur potentiell relevante Objektlokalisierungen anzeigen, sondern auch – und dies ist in der AR-Bedingung von zentraler Bedeutung – auf weitere körperliche Ressourcen verweisen (z.B. deiktische Gesten), die in der holistischen Gestaltform der multimodalen Orientierungshandlung zum Ausdruck kommen. Eine Re-Orientierung erfolgt hier aus dem situativen Handlungswissen, dass der Referent der Orientierungsgestalt bei einer rechtzeitigen Aufmerksamkeitsverlagerung ermittelt werden kann.

Auf empirischer Grundlage der vollannotierten Versuchspersonengruppen der AR-Bedingung (AR-05, AR-06, AR-10) lassen sich drei phänomenologisch unterschiedliche Fallgruppen ausdifferenzieren:

- 1) Direkt erfolgreiche Ko-Orientierung durch lokaldeiktische Ausdrücke (71%, 22 Vorkommen):
Beispiel: Fragment 19
- 2) Indirekt erfolgreiche Ko-Orientierung (nach einer Reparatur) durch lokaldeiktische Ausdrücke (23%, 7 Vorkommen):
Beispiel: Fragment 20
- 3) Unterschiedliche Teilnehmerinterpretationen lokaldeiktischer Ausdrücke (6%, 2 Vorkommen):
Beispiel: Fragment 21 & Fragment 22

In den Angaben zur numerischen Häufigkeit der Fallgruppen wurden solche Situationen erfasst, in denen die Interaktanten initial nicht ko-orientiert sind und verbaldeiktische Ausdrücke eine erkennbare Reaktion des jeweiligen Ko-Partizipanten zur Aufmerksamkeitsverlagerung hervorrufen. Vor dem Hintergrund der in den drei relevanten Versuchspersonengruppen verfügbaren Anzahl von je 18 Mediationsobjekten mögen die Vorkommen in den Fallgruppen relativ gering erscheinen. Hierbei ist jedoch zu berücksichtigen, dass i) in der räumlichen Dimensionierung der Tischsituation die Anwesenheit von Ko-Orientierung einen Standardfall markiert, ii) Objektvorschläge zum Teil aus bereits etablierten Ko-Orientierungen semantisch relationiert werden („dann können wir auch Objekt XY dort aufstellen“) und iii) eine Vielzahl interaktionaler Verfahren in multimodalen Gestaltformen aktiv bei der Herstellung von Ko-Orientierung beteiligt sind (vgl. 3.4). Insofern beschreibt die Herstellung von Ko-Orientierung durch lokaldeiktische Ausdrücke solche Fälle, in denen verweisende Akteure Ko-Orientierung unterstellen (vgl. 2.4.2), diese jedoch in der spezifischen Situation nicht etabliert ist. In diesem Sinne stellen lokaldeiktische Ausdrücke neben aktiven Orientierungspotentialen gleichermaßen Re-Orientierungspotentiale bereit, die in der Sprachverwendung selbst ‚vorbeugend‘ angelegt sind.

Um die numerische Distribution adäquat einordnen zu können, sei an dieser Stelle auch auf den anhaltenden Diskurs zur Problematik der Quantifizierung in der Konversationsanalyse verwiesen (Schegloff 1993; Heritage 1995). Im Vergleich zu strikt experimentellen Settings, in denen sich Phänomenbereiche auf einzelne Untersuchungsvariablen reduzieren lassen, handelt es sich bei interaktionalen Phänomenen stets um multifaktorielle Zusammenhänge semiotischer Teilnehmer- und Kontextressourcen, die nicht stabil sind, sondern sich zeitlich – im Sinne des Emergenzprinzips – mit der fortschreitenden Bearbeitung interaktionaler Aufgaben verändern. Die nichtlineare Dynamik des Interaktionsgeschehens entzieht sich gewissermaßen einer Operationalisierung, die in der Lage wäre, das situative Ensemble interaktionaler Aktivitäten in seiner genuinen Charakteristik abzubilden. Trotz des in der Komplexität reduzierten Phänomenbereichs (‚lokaldeiktische Ausdrücke‘) und des methodisch vorteilhaften Ansatzes, einen analytischen Zugang zu den visuell-semiotischen Ressourcen der Teilnehmerwahrnehmungen zu haben, der es erlaubt, das ‚reagieren auf‘ zu präzisieren und einzugrenzen (vgl. 1.2.3), kann es sich bei der numerischen Distribution nur um absolute Werte einer konversationsanalytisch fundierten Rekonstruktion, jedoch um relative Werte einer interaktionalen Komplexität multimodaler Orientierungsgestalten handeln. Danach bildet die numerische Verteilung lediglich ab, ob lokaldeiktische Ausdrücke mit einer sequentiell nachfolgenden Reaktion des Ko-Partizipanten zusammenfallen, nicht aber inwieweit diese Reaktion beispielsweise durch syntaktisch-semantische Projektionen oder subtile Orientierungshinweise (z.B. Geräusche der

Objektmanipulation) begünstigt oder gar hervorgerufen wird. Insofern sind die Angaben zur Häufigkeit der Phänomenklassen als Darstellungsoption und Versuch zur Präzisierung der „informellen Quantifizierung“⁶¹ (Schegloff 1993: 118f.) zu betrachten, die in ihrer Aussagekraft eine empirische Tendenz widerspiegeln.

3.2.1 Erfolgreiche Ko-Orientierung durch verbalsprachliche Lokaldeixis

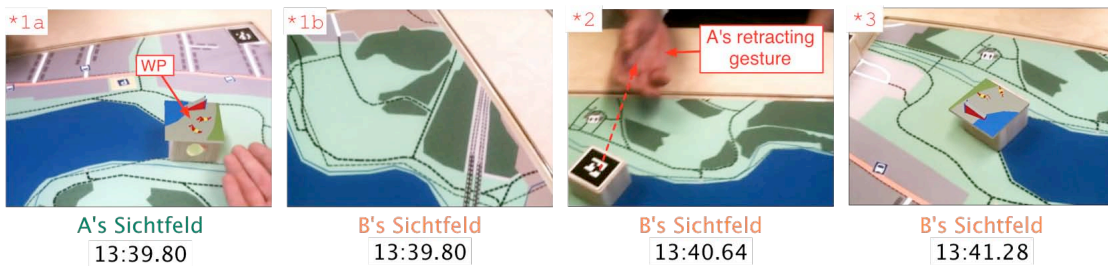
Das folgende Fragment 19 repräsentiert mit 22 Vorkommen den *default*-Fall verbalsprachlich (re-)initiiertes Ko-Orientierung durch lokaldeiktische Ausdrücke. Dementsprechend führen in 71% der Fälle Lokaldeiktika zu einer „direkt erfolgreichen Ko-Orientierung“ (3.2, Punkt 1) und erfüllen die Kriterien „keine Erfordernis zur Reparatur“ und „keine abweichende Interpretation zwischen intendierter und rezipierter Aufmerksamkeitsorientierung“. Dabei greifen oftmals zwei koordinative Mechanismen ineinander: Während das Lokaldeiktikum eine Blickrichtungsänderung des Ko-Partizipanten konditionell relevant setzt, konstituiert die sich in der Retraktionsphase befindliche Zeigegeste des verweisenden Sprechers einen suffizienten Orientierungshinweis auf das Zeigeziel: Der Ko-Partizipant nutzt die Fähigkeit, den Referenten retrospektiv aus der retrahierenden Geste des Verweisgebers zu inferieren. In dieser Hinsicht agieren die Teilnehmer „opportunistisch“, da sie dem “principle of least joint effort”⁶² (Clark 1996; Clark & Brennan 1991) folgend – unter der gegenseitigen Unterstellung von Ko-Orientierung (vgl. 2.4.2) – alltägliche Orientierungsroutinen nutzen, die für die praktischen Zwecke lokal anstehender Aufgaben zwar suffizient sind, jedoch eine geringere Stabilität und Reliabilität als in „störungsfreien“ face-to-face Situationen aufweisen (vgl. 3.2.4).

FRAGMENT 19:

AR-Baseline:	VP-Gruppe 06 (VP-A: Naturschützer / VP-B: Investor)
Zeit:	13:36 - 13:41

```
01 A-ver: |zur entspannung können=äh: die leute |dann ja HIER
    A-act:                                |dG>>>>>>>>>WP
    A-gaz: |@WP.....
    B-gaz: |@map.....
                                          *1a+b

02 A-ver:  zu ihrem wunderBAR=zu ihrer wunderbaren RUTsche|
    A-act:  >>>>>>>>>>>>>@HP.....
    B-gaz:  >>>>@WP.....
              *2      *3
```



⁶¹ “Informal quantification is the product of a quite different – but nonetheless methodological – orientation to empirical materials. Terminology such as *occasionally* or *massively* reports an *experience* or *grasp* of frequency, not a count; an account of an investigator’s sense of frequency over the range of a research experience, not in a specifically bounded body of data; a characterization of distribution fully though tacitly informed by the analytic import of what is being characterized” (Schegloff 1993: 118f., Hervorhebungen im Original).

⁶² Clark & Krych (2004: 64, Hervorhebungen im Original): “According to that principle, people are *opportunist*: they try to select from the available methods the ones they think take the least effort for the two of them jointly – the least cost in time, resources, errors etc. [...] If people are opportunist, they should generally opt for the grounding methods that are most efficient (Clark & Brennan 1991).”

Zu Beginn des Fragments sind beide Interaktanten auf unterschiedliche Entitäten orientiert – A fokussiert das Objekt “WaterPark” (WP) (*1a), während B auf einen Kartenabschnitt orientiert ist, der sich auf der gegenüberliegenden Seite des Sees befindet (*1b). Dabei verweist A in einer multimodalen Äußerungs-gestalt aus deiktischer Geste und lokaldeiktischem Ausdruck „HIER“ auf das Objekt “WP” (*1a). Auch wenn A’s gestischer Verweis B zu diesem Zeitpunkt als Orientierungsressource nicht zur Verfügung steht (*1b), reagiert er zeitlich präzise, unmittelbar nach der Äußerung des Lokaladverbs, und leitet eine Blick-richtungsänderung ein (*2). Währenddessen nimmt er die sich in der Retraktionsphase befindliche Geste A’s wahr und erschließt den Referenten (*3), noch bevor dieser verbalsprachlich als „RUTsche“ identi-fiziert wird.

Das Fragment liefert empirische Hinweise auf zwei relevante, miteinander verschränkte Aspekte der Koordinierung: Zum einen konsolidiert das Beispiel den empirischen Befund, dass die Interaktanten unter einer gegenseitigen Unterstellung von Ko-Orientierung agieren. – Äquivalent zu Fragment 17 (vgl. 2.4.2) sind keinerlei Expansionsauffälligkeiten im bedeutungskonstitutiven *Stroke* der Zeigegeste erkennbar, da unter den limitierten Bedingungen des *mutual monitoring* der verweisende Akteur sein Verhalten nicht mit der Aufmerksamkeitsorientierung des Ko-Partizipienten koordinieren und ggf. optimieren kann (z.B. durch zeitliche Expansionen oder salientere Trajektorien der Zeigegeste).

Zum anderen jedoch manifestieren sich genau aus dieser Konstellation empirische Hinweise auf die Robustheit interaktiver Koordinierungsressourcen oder – aus einer etischen Perspektive – interaktionaler Systeme. Das analytisch rekonstruierte Verfahren dieser Fallgruppe deutet darauf hin, dass nicht nur der *Onset* einer deiktischen Verweishandlung einen vektoriellen Suchraum eröffnet, in dem das Zeigeziel auf-gefunden werden muss (vgl. Stukenbrock 2012; Goodwin 2000b), sondern auch die *Retraktion*, die es dem jeweiligen Ko-Partizipienten unter spezifischen Bedingungen erlaubt, den vektoriellen Suchraum retro-spektiv zu erschließen und inferentiell das Zeigeziel aus dem Handlungskontext zu ermitteln. Dies könnte darauf hinweisen, dass deiktische Gesten in ihrer zeitlichen und figurativen Gestaltform holistisch memo-riert werden, so dass die fragmentarische Wahrnehmung konstitutiver Gestenphasen nicht unbedingt ein interaktionales Problem darstellt, sondern eine erhöhte Anforderung zur ganzheitlichen Rekonstruktion. Vor diesem Hintergrund erscheinen deiktische Gesten in ihrer orientierenden Funktion relativ robust zu sein. Darüber hinaus ist festzustellen, dass Clarks Annahme für deiktische Ausdrücke:

“they require both speakers and addressees to establish that the things indexed are in their joint attention [...]. That, in turn, requires speakers to monitor what their addressees are doing, and addressees to show what they are doing.” (Clark 2003)

nicht ein generelles Prinzip unter limitierten Bedingungen des *mutual monitoring* darstellt. Insofern kann die der Gestaltpsychologie entlehnte methodologische Annahme einer holistischen Produktion und Rezeption multimodaler Gestalten dahingehend präzisiert werden, dass Interaktanten versuchen, die bedeutungskon-stituierenden Elemente der holistischen Gestaltform zu rekonstruieren, wenn diese ihrer Wahrnehmung nicht zugänglich sind. Dieses Rekonstruktionspotential lässt sich jedoch ebenso auf gestalttypologische Prinzipien zurückführen. Wie die Fragmente der folgenden Fallgruppen zeigen, resultiert daraus jedoch nicht die indirekte Bestätigung der methodologischen Annahme, da die Rekonstruktion holistischer Ge-staltformen kein universelles Prinzip darstellt, sondern an spezifische Bedingungen geknüpft ist.

3.2.2 Reparatur von Ko-Orientierung: Wenn lokaldeiktische Ausdrücke insuffizient sind

Äquivalent zum vorigen Fragment (vgl. 3.2.1, Fragment 19) zeigt das nachfolgende Fragment 20 eine Si-tuation, in der ein verbalsprachlicher Ausdruck des Verweisgebers eine Orientierungsanpassung des Ko-Partizipient initiiert, in der er lediglich die retrahierende Zeigegeste als semiotische Orientierungsressource nutzen kann, um auf den intendierten Referenten zu schlussfolgern. Während dieses Orientierungsver-fahren in den meisten Fällen suffizient ist (71%, 22 Vorkommen, vgl. 3.2.1), verweisen einige Phänomenin-stanzen auf interaktionale Bedingungen, in denen der jeweilige Ko-Partizipient nicht in der Lage ist, das Zeigeziel inferentiell zu erschließen und infolgedessen eine Reparatur einleitet (23%, 7 Vorkommen). Dabei ist zu beobachten, dass die Erfordernis zur Reparatur eng an die physikalisch-materiellen Merk-malsspezifikationen des Referenten und seines jeweiligen Aushandlungsstatus gekoppelt ist. In Fällen, in denen das Zeigeziel eine saliente, in seiner figurativ-räumlichen Gestalt klar umrissene Orientierungs-markierung konstituiert („Landmarke“ bei Bühler), wie sie in etwa durch die AR-Würfel repräsentiert wird

‚aufladen‘ und ihm ggf. eine ‚Orientierungsrelevanz‘ zuweisen. Die Suffizienz deiktischer Verweishandlung in der AR-Bedingung ist dementsprechend in erheblichem Maße voraussetzungsreich, was sich nicht nur in der relativ hoch erscheinenden Anzahl reparaturbedürftiger Verweishandlungen in Fällen der lokaldeiktisch initiierten Ko-Orientierung abzeichnet (knapp $\frac{1}{4}$), sondern vor allem im numerischen Gesamtvergleich der deiktischen Gesten zwischen der AR- und F2F- Bedingung niederschlägt (vgl. Abbildung 16).

Abbildung 16 zeigt, dass die Annotationszeitwerte deiktischer Gesten auf Objekte (dG@object(s)) und deiktischer Gesten auf die Karte (dG@map) in der AR-Bedingung um etwa die Hälfte reduziert sind. Interaktanten der AR-Bedingung wenden demnach im Durchschnitt nur etwa halb so viel Zeit auf deiktische Verweisgesten auf wie die Interaktanten der F2F-Bedingung. Die durchschnittliche Ausführungszeit von der Onset- bis zur Retraktionsphase (Ausführungsendpunkt ist durch die *home position* definiert) einzelner Gesten beider Kategorien ist dabei jedoch in der AR-Bedingung leicht erhöht (dG@object(s): AR=1,86s/dG; F2F=1,35s/dG // dG@map: AR=2,31s/dG; F2F=1,89s/dG). Die Auffälligkeiten in der Distribution manifestieren sich vor allem in der numerischen Anzahl deiktischer Gesten (absolute Anzahl), die in der F2F-Bedingung für die Kategorie „dG@object(s)“ um den Faktor 3,1 und für die Kategorie „dG@map“ um den Faktor 2,5 erhöht ist.

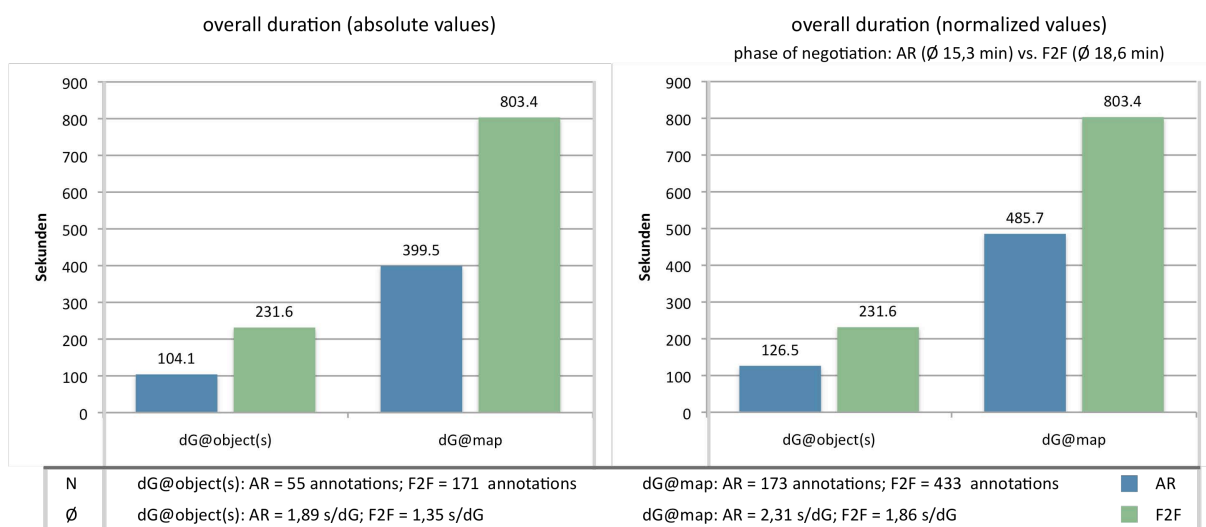


Abbildung 16: Absolute (links) und normalisierte Annotationszeitwerte (rechts) deiktischer Gesten auf Objekte (@object(s)) und Bereiche der Karte (@map) zwischen der AR- und F2F- Bedingung. Dargestellt sind die summierten Werte der jeweils drei vollannotierten Versuchspersonengruppen (VP-Gruppen AR: 05+06+10; VP-Gruppen F2F: 03+04+05). Die absoluten Annotationszeitwerte wurden in der rechten Darstellung extrapoliert, da die Aushandlungsphase der AR-Bedingung mit 15,3 min durchschnittlich kürzer ist als die der F2F-Bedingung mit 18,6 min.

Vor dem Hintergrund der einführend dargestellten Problematik zur Quantifizierung in der Konversationsanalyse (Schegloff 1993; Heritage 1995, vgl. 3.2) erscheint es uns nicht angemessen, die distributiven Auffälligkeiten unter dem Stichwort der „Signifikanz“ zu bewerten. Auch wenn es sich hier um einen kontrastiven Vergleich zwischen verschiedenen Interaktionsbedingungen handelt, dessen Erkenntnisinteresse nicht in der sequentiellen Einbettung und situativen Ausgestaltung eines Verfahrens, sondern in der Distribution der klar definierbaren Handlungsweise „deiktische Geste“ besteht, muss diese Verteilung nicht zuletzt auch im Zusammenhang der interaktionalen Emergenz und der individuellen Spezifik der Personenkonstellationen betrachtet werden. Die Ableitung universeller Prinzipien der Art „unter *mutual monitoring* einschränkenden Bedingungen verwenden die Teilnehmer weniger deiktische Gesten“ erscheint daher nur bedingt sinnvoll⁶³. Vordergründig soll die statistische Verteilung als holistische Darstellungsoption und Abbildung einer Tendenz innerhalb der hier fokussierten Teilnehmerinteraktionen

⁶³ Dies bedeutet nicht, dass Universalien dieser Art auszuschließen sind. Eine fundierte Bewertung erfordert jedoch eine fundamental breitere Annotationsbasis, in der sich interaktionale Emergenzeffekte und individuelle Unterschiede zwischen Versuchspersonen und Versuchspersonenkonstellationen in der untersuchten Gesamtheit ‚neutralisieren‘.

dienen, in der die empirischen Beobachtungen der qualitativen Fallanalyse Erklärungsmodelle zur quantitativen Distribution liefern.

Unter dieser Perspektive lassen sich die folgenden empirischen Beobachtungen im analytischen Vergleich beider Interaktionsbedingungen als Erklärungsansätze zur quantitativen Distribution der deiktischen Gesten formulieren:

- 1) Die Interaktionen der AR-Bedingung verweisen auf eine ‚entschleunigte Kommunikation‘ („*sloving-down-communication*“), in der interaktional abgebildete Relationen zwischen unterschiedlichen Objekten und Entitäten der Karte vermindert auftreten. Koordinierungsanforderungen zur Beendigung einer gegenwärtigen Aufgabe und parallelemergente Handlungsansprüche werden nicht instantan, sondern mit sequentieller Verzögerung vom jeweiligen Ko-Partizipanten erkannt (vgl. Fragment 15 und Fragment 16).
- 2) Zeitlich flüchtige Verweisgesten auf Raumentitäten, die nur schwerlich durch interaktionale Salienzindikatoren mit einer ‚Orientierungsrelevanz‘ belegt werden können (z.B. unspezifische Orte der Karte), sind oftmals in der situativen Referenzierung auf Reparaturen angewiesen, wenn die *stroke*-Phase der Verweisgeste nicht erkannt oder inferentiell ermittelt werden kann (vgl. Fragment 19 und Fragment 20).

Vor dem Hintergrund dieser Erkenntnisse, die sich in einer regressiven Distribution deiktischer Gesten in der AR-Bedingung niederschlagen, stellt sich die Frage, inwieweit die interaktive Kollaboration von interaktionalen Problemen geprägt ist? Obgleich verbaledeiktische Ausdrücke überwiegend eine suffiziente Ressource zur direkten Herstellung von Ko-Orientierung bilden (vgl. Fragment 19), weist knapp $\frac{1}{4}$ der Phänomeninstanzen eine Notwendigkeit zur Reparatur auf (vgl. Fragment 20). Dies hält die Interaktanten jedoch nicht davon ab, die Bearbeitung lokal anstehender Interaktionsaufgaben zielgerichtet – im Fokus der globalen Aufgabebearbeitung – voranzutreiben.

Eine zentrale Relevanz kommt hier der strikten Trennung zwischen einer etisch-analytischen Bewertung und der emisch-teilnehmerrekonstruktiven Analyse zu. Aus einer etischen Perspektive lässt sich im kontrastiven Vergleich der Bedingungen anmerken, dass die Einschränkungen im *mutual monitoring* ‚kostenintensiv‘ sind. Der multireferentielle Angebotscharakter der F2F-Bedingung, der sich in der Vielzahl hergestellter relationaler Bezüge zwischen Objekten und etablierten Raumstrukturen manifestiert, ist in der AR-Bedingung dergestalt nicht beobachtbar. Im Hinblick auf die numerische Anzahl deiktischer Gesten auf Objekte (*dg@object(s)*), ergibt sich aus den 55 annotierten Instanzen (Summe der drei voll-annotierten VP-Gruppen) eine durchschnittliche Anzahl von 18,3 objektreferentiellen Zeigegesten pro VP-Gruppe. Setzen wir diese Anzahl in Bezug zu den 18 verfügbaren Mediationsobjekten im Experiment, findet die empirische Beobachtung in ihrer distributiven Verteilung eine Bestätigung. Die Einschränkung des *mutual monitoring* wirkt sich gewissermaßen in einer Tendenz aus, in der ‚objektpluralistische‘ Bearbeitungsstrategien der F2F-Bedingung unter den gesteigerten Koordinierungsanforderungen in der AR-Bedingung ‚objektmonistisch‘ stabilisiert werden.

Aus einer etischen Perspektive ist ebenfalls auf die Auswirkung hinzuweisen, dass fast jede vierte durch lokaldeiktische Ausdrücke initiierte Orientierungsverlagerung nicht direkt im Zielzustand der Ko-Orientierung mündet, sondern von den Teilnehmern repariert werden muss. Im kontrastiven Vergleich zur F2F-Bedingung, in der orientierungsrelevante Aktivitäten selbst im „*peripheral monitoring*“ instantan erkannt und bearbeitet werden, kann die relativ ausgeprägte Revisionsbedürftigkeit in der AR-Bedingung als Beeinträchtigung der ‚zeitlichen Effizienz‘ (‚entschleunigte Kommunikation‘) gewertet werden.

Aus einer emisch-teilnehmerrekonstruktiven Perspektive ist jedoch anzumerken, dass diese vermeintlichen Einschränkungen und Störparameter keine teilnehmerrelevanten Kategorien bilden. Das Verhalten der Interaktanten orientiert sich primär daran, ob interaktionale Aufgaben kooperativ bearbeitet werden können und nicht daran, ob diese Bearbeitung effizient ist oder welche Effizienz ihr unter anderen Bedingungen zukäme. Reparaturen von Ko-Orientierung beschreiben hierbei aus der emischen Teilnehmerperspektive eher unproblematische Fälle, da sie eine reziproke Erkenntnis voraussetzen, dass das, was der verweisende Akteur zu Orientierungszwecken sichtbar gemacht hat, und das, was davon wahrgenommen wurde, nicht kohärent ist. Eine festgestellte Inkohärenz kann jedoch relativ leicht und direkt bearbeitet werden (vgl. Fragment 20), so dass sie nicht Gefahr läuft, sich in interpersonell abweichenden Wissensrepräsentationen eines Sachverhalts (i.S.d. „*common ground*“) zu fossilieren.

Betrachten wir jene seltenen Fälle, in denen die Teilnehmer divergente Interpretationen einer Orientierungsgestalt entwickeln, wird einerseits deutlich, dass sie als ‚normale‘ Missverständnisse von den Interaktanten gedeutet und durch eine sofortige Reparatur aufgelöst werden. Dies zeigt das folgende Fragment 21, in dem divergente Interpretationen genau wie in Fragment 20 durch eine objektmanipulative Handlung korrigiert werden. Andererseits verweist Fragment 21 darauf, dass divergente Interpretationen nicht notwendigerweise von den Interaktanten erkannt werden und im interaktionalen Folgegeschehen unbemerkt bleiben. Auch wenn sich solche Situationen im Anwendungsfall der Technik konfliktär auswirken könnten, stellen sie im Rahmen der interaktionalen Aufgabenbewältigung für die Teilnehmer selbst kein Konfliktpotential dar.

3.2.4 Abweichende Fälle

Wie wir in der vorigen Analysesektion gezeigt haben, ist das Referenzobjekt lokaldeiktischer Ausdrücke nicht immer vom Ko-Partizipienten eindeutig zu identifizieren. Etwaige interaktionale Probleme in der Bestimmung des intendierten Referenten stehen dabei sowohl in Abhängigkeit zum spezifischen Affordanzcharakter des Referenten als auch zur Votalität der deiktischen Verweisgeste, die aufgrund der Einschränkung im *mutual monitoring* und resultierender visueller Dispersionseffekte nicht mit dem Verhalten des Ko-Partizipienten in situ koordiniert werden kann. Infolgedessen geht die verweisgebende Partei unter der Annahme und reziproken Unterstellung von Ko-Orientierung als *default*-Fall davon aus, dass die rezipierende Partei das referentielle Verfahren erkennen und deuten kann.

In der folgenden Analysesektion präsentieren wir zwei abweichende Fälle (i.S.d. „deviant case analysis“, z.B. Schegloff 1968; Maynard & Clayman 2003: 177ff.), in denen die Einschränkung des *mutual monitoring* zu divergenten Interpretationen zwischen der intendierten Referenzierung des Verweisgebers und ihrer rezipierten Interpretation des Ko-Partizipienten führen. Auch wenn fehlgedeutete Interpretationen lokaldeiktischer Ausdrücke seltene Phänomene innerhalb unseres Korpus, bzw. der hier numerisch ausgewerteten drei VP-Gruppen bilden (6%, 2 Vorkommen), sind diese Fälle epistemisch besonders aufschlussreich, da sie Erkenntnisse zum Einfluss des *mutual monitoring* auf die Teilnehmerverfahren zur Koordinierung ihrer Handlungen liefern. Im Speziellen legt die Analyse offen, dass divergente Interpretationen auf eine wahrnehmungsspezifische Rezeptionsproblematik zurückzuführen sind, in der relevante Desambiguierungshinweise in der multimodalen Orientierungsgestalt der verweisenden Partei vom Rezipienten nicht erkannt werden.

3.2.4.1 Desambiguierungsprobleme – Divergenz von Bedeutungskonstitution und -interpretation einer Orientierungsgestalt

Der erste Fall zeigt, wie Teilnehmer divergente Interpretationen eines lokaldeiktischen Ausdrucks durch die Wahl unterschiedlicher Bezugspunkte der Raumorientierung entwickeln (*Relatum* nach Bühler 1934). Dabei wird deutlich, dass sowohl der Referenzierung des Verweisgebers als auch der Referenzdeutung des Rezipienten implizite Annahmen zugrunde liegen, wie die referentielle Praxis vor dem Hintergrund relevanter Raumentitäten der Interaktionsgeschichte zu verstehen ist. Im nachfolgenden Fragment kann die von der Verweiskonstitution des Sprechers abweichende Adressateninterpretation jedoch sofort bearbeitet werden, so dass mitunter interaktional folgenreiche Informationsdefizite und Abweichungen im „common ground“ (Clark et al. 1983, 1991) der Interaktanten abgewendet werden und sich nicht in der weiteren prozessualen Interaktionsentwicklung tradieren.

FRAGMENT 21:

AR-Baseline:	VP-Gruppe 10 (VP-A: Investor / VP-B: Naturschützer)
Zeit:	19:22 - 19:32

Präkontext: Aushandlung der Objektplatzierung "Hot-Air-Balloon" (HAB):



B: dann aber vielleicht auch HIER am rand; =das das hier HINTEN alles frei bleibt;

Im Vorfeld des Fragments (ca. 28 Sekunden zuvor) handeln die Interaktanten die Platzierung des Objekts "Hot-Air-Balloon" (HAB) aus. Im Zuge dessen schlägt B vor, das Objekt im Randbereich der Karte zur platzieren (*1), damit eine Fläche frei bleibt, die er gestisch als die komplette, ihm zugewandte Uferseite des Sees markiert (*2a). Ein Vergleich der Screenshots *2a und *2b zeigt, dass B's gestisch-expandierter Verweis zu diesem Zeitpunkt in A's Sichtfeld nicht erkennbar ist. In Orientierung auf das kontemporär relevante Objekt "HAB" (*2b) erlangt A infolgedessen keine präzisen Informationen über die exakte Ausdehnung der Freifläche. In Bezug auf den nachfolgend dargestellten Orientierungsversuch manifestiert sich die spezifische Relevanz der von B entworfenen Freifläche: Als Relikt der Interaktionsgeschichte beeinflusst die vermeintlich ausgehandelte Region den interpretativen Modus eines sequentiell nachfolgenden Orientierungsangebots.

Teil A: Phase der Emergenz:

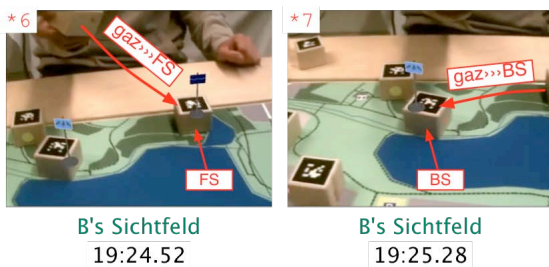
19:22
 |
 01 A-ver: |das KANN man ja vielleicht noch **HIER** auf der anderen seite-
 B-gaz: |@loc-0.....>>>@MG..
 A-act: |lift (MG).....dG...hold..
*3a+b *4a+b *5



Unmittelbar im Anschluss an die Platzierungsaushandlung des Objekts "HAB" schlägt A das Objekt "Miniature Golf" (MG) vor, hält es angehoben in Richtung seines Ko-Partizipanten B und verbalisiert mit „das KANN man ja vielleicht noch HIER auf der anderen seite-“ ein Platzierungsangebot (01). Wie schon in den vorigen Fragmenten beobachtbar, fungiert das Lokaladverb „HIER“ als "attention getter" (fett markiert). Im direkten Vergleich der Sichtfelder beider Teilnehmer zu Beginn des Vorschlags (*3a+*3b) lässt sich beobachten, dass A's initiale Handlung der Objektsalienzsteigerung durch Anheben und ein im Neigungswinkel dem Ko-Partizipanten zugewandtes Präsentieren (*3a) für B nicht wahrnehmbar ist (*3b). In analytischer Fokussierung auf die ganzheitliche Proposition „HIER auf der anderen seite-“ fällt

dabei eine zweifache Indizierung der „anderen seite-“ in A’s multimodaler Orientierungsgestalt auf: Einerseits verweist er auf den Referenten (die grau markierte „Wohnfläche“) durch seine Blickorientierung, andererseits durch die physische Objektausrichtung in Kombination mit einer deiktischen Geste, die durch ein “pointing” der materiellen und augmentierten Ecke des Objektes (durch vor- und zurückbewegen) auf das Zeigeziel realisiert wird (*4a, markiert durch rote Balken). Das Sichtfeld B’s zeigt jedoch, dass ihm die doppelte Indizierung der „anderen seite-“ als semiotische Orientierungsressourcen zu diesem Zeitpunkt nicht zur Verfügung steht (*4b), da seine Aufmerksamkeitsverlagerung auf das Aushandlungsobjekt “MG” etwas zu spät erfolgt (*5). Die fehlende Augmentierung (*5) (evt. verdeckt A hier den AR-Marker) wird dabei nicht thematisiert, da A bereits initial, noch vor der hier dargestellten Vorschlagshandlung, das Objekt verbalsprachlich identifiziert hatte. Das in die Diskussion eingebrachte Objekt ist somit eindeutig einem „Minigolfplatz“ zugeordnet, auch wenn A an dieser Stelle nur seinen materiellen Träger erkennen kann.

```
02 B-ver: !INS! naturschutz |gebiet,
A-ver: |ne=ne;
B-gaz: >>FS>>>>>BS.....
        *6          *7
```



FishSign (FS) & BirdSign (BS) wurden bereits als "Naturschutzgebiet" ausgehandelt

B’s Reaktion auf A’s Orientierungsangebot verdeutlicht die Divergenz zwischen projizierter und interpretierter Orientierungsgestalt. Während sich A in seiner Vorschlagshandlung auf eine Objektlokalisierung bezieht, die von seinem perspektivischen Standpunkt ausgehend auf der gegenüberliegenden Seite des Sees (die grau markierte „Wohnfläche“) – also direkt vor seinem Ko-Partizipanten B – liegt, interpretiert dieser die vorgeschlagene Objektverortung diametral. Infolgedessen geht B davon aus, dass A das Objekt direkt in das unmittelbar vor dem Verweisgeber befindliche Naturschutzgebiet – in opponierter Position zur grau markierten „Wohnfläche“ – platzieren möchte. B’s spezifisches Verständnis der Vorschlagshandlung A’s wird dabei durch die prosodisch markierte Nachfrage „!INS! naturschutzgebiet,“ projiziert (02). Zeitgleich blickt er zuerst zum “FishSign” (FS) und dann zum “BirdSign” (BS)(*6+*7) – zwei Objekte, die bereits zuvor als ‚Platzhalter‘ und räumliche Begrenzung des Naturschutzgebietes ausgehandelt wurden.

B’s divergente Interpretation der referentiellen Orientierungspraxis A’s liegt hierbei darin begründet, dass beide Parteien unterschiedliche Relata der Raumreferenz und -orientierung wählen. Während A aus der subjektiven Orientierung ein *Origo*-zentriertes Verfahren anwendet, nutzt B in der Interpretation dieser Handlung die im Zuge der Objektaushandlung “Hot-Air-Balloon” (HAB) markierte und vermeintlich ausgehandelte Freifläche als Bezugspunkt der Raumorientierung (Relatum) (vgl. *2a). Die „fälschliche Wahl“ des Relatums resultiert aus der wahrnehmungsspezifischen Rezeptionsproblematik, dass in diesem Beispiel exakt diejenigen Orientierungshinweise der multimodalen Formgestalt des Verweisgebers nicht wahrgenommen werden konnten (Blick & Objekt-*pointing*), die den lokaldeiktischen Ausdruck situativ desambiguieren und in seiner referentiellen Spezifik nachvollziehbar figurieren (vgl. Abbildung 17).

Die Anschlusshandlung B's bestätigt retrospektiv die analytisch rekonstruierte Annahme, dass er A's Verweishandlung als ein in Relation zur initial markierten Freifläche (*Relatum*) stehendes Orientierungsangebot wahrgenommen und gedeutet hat. Äquivalent zur anfänglichen Aushandlung der Objektlokalisierung „HAB“ (vgl. *1) schlägt er vor, das Objekt „MG“ eher am Kartenrand zu platzieren (*9a+*9b) und ratifiziert auf diese Weise sein Bestreben, die initial markierte Region der vor ihm situierten Uferseite als Freifläche zu belassen.

Divergente Interpretation lokaldeiktischer Ausdrücke in multimodalen Orientierungsgestalten dieser Art zeigen, dass die Teilnehmer Schwierigkeiten haben, die zugrundeliegende Perspektivierung ihrer Handlungen mit der gegenperspektivischen Orientierung des jeweiligen Ko-Partizipanten zu koordinieren (s.a. Schober 1993), um die Wahrnehmbarkeit relevanter Orientierungshinweise zu taxieren. Umgekehrt kann der Rezeptionsprozess der Orientierungsgestalt nicht hinlänglich an die perspektivische Darstellungsform des Verweisebers adaptiert werden. Der Interpretationsmodus der Orientierungsgestalt ist dementsprechend auf die Beobachterperspektive beschränkt. Ein grundlegendes Prinzip des kommunikativen Handelns – die Reziprozität der Perspektiven (Schütz 1960) – ist hier verletzt.

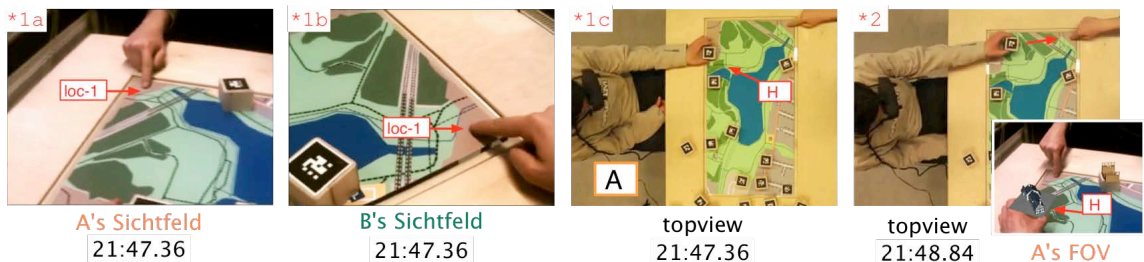
3.2.4.2 Verbalsprachliche Ausdrücke in multiplen Orientierungsangeboten

Während in den vorigen Beispielen (Fragment 19 - Fragment 21) Fälle behandelt wurden, in denen lokaldeiktische Verfahren zur Herstellung von Ko-Orientierung entweder direkt erfolgreich waren oder in der reziproken Teilnehmererkenntnis fehlgedeuteter Orientierungsangebote die Möglichkeit zur Reparatur boten, verweist das nachfolgende Fragment 22 auf eine Situation, in der die Divergenz zwischen dem intendierten Orientierungsangebot des Sprechers und der rezipierten Interpretation des Adressaten un bemerkt bleibt (es handelt sich hierbei um das Eingangsbeispiel der Gegenstandskonstitution, das hier jedoch unter einer anderen analytischen Zielsetzung aufgegriffen wird). Die Fehlinterpretation des Adressaten ist im folgenden Fall auf eine Simultanaktivität des Verweisebers zurückzuführen, in der eine rechts händige deiktische Verweigeste ‚akzidentell‘ mit einer linkhändigen Objektmanipulation zeitlich zusammenfällt. Äquivalent zu Fragment 21 manifestiert sich in Fragment 22 die empirische Beobachtung einer wahrnehmungsspezifischen Rezeptionsproblematik, in der die restringierten Bedingungen im *mutual monitoring* das Erkennen desambiguierender und spezifizierender Orientierungshinweise der offerierten Gestaltform erschweren. Darüber hinaus exemplifiziert das Fragment einen interaktionalen Mechanismus, in dem selbst subtile, feingraduelle Aktivitäten der orientierenden Partei vom Ko-Partizipanten sinnhaft gedeutet und in die emergente Orientierungshandlung integriert werden. Insofern reagiert der Adressat äußerst sensitiv auf semiotische Ko- und Kontextinformationen einer Orientierungsgestalt, denen eine potentielle Orientierungsrelevanz zuteilwird.

FRAGMENT 22:

AR-Baseline:	VP-Gruppe 10 (VP-A: Investor / VP-B: Naturschützer)
Zeit:	21:46 - 21:54

```
01 B-ver: |kann man DAS |nich HIER platzieren, so nen bisschen außerhalb;
B-act:   |dG>>@loc-1.....
A-lH:   |grasp>>>>>>>|@H.....lift.....frwd>>>>>>>>>>>>>>>>loc-1.....
                                         *1a-c                                     *2
```



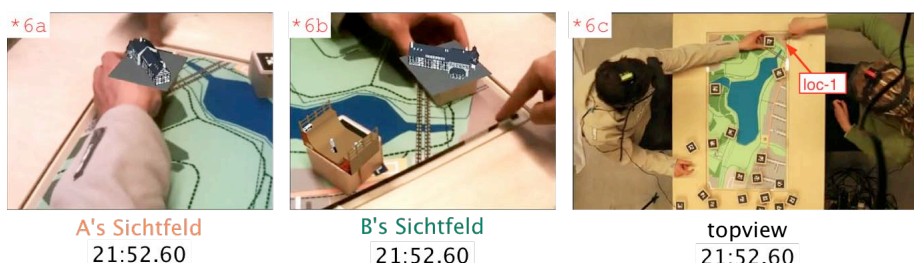
Wir steigen in die Analyse der Interaktion ein, nachdem A das Objekt „Hotel“ (H) als neue Interaktionsaufgabe ausgewählt, etabliert und einen initialen Vorschlag zur Platzierung verbalisiert hat, der von B abgelehnt wurde. Der Beginn des Transkripts markiert nun B's Gegenvorschlag, das Objekt „nen bisschen außerhalb“ zu platzieren (01). In Ko-Orientierung auf den vorgeschlagenen Lokalisationsort „loc-1“ (*1a-c) reagiert A auf B's Orientierungsangebot, bestehend aus deiktischer Geste und prosodisch markiertem Lokaladverb „HIER“, indem er das linkshändig manipulierte Objekt „H“ in Richtung des von B indizierten Platzierungsortes führt (*2).

```
02 B-ver: |HIER, (.)|!HIER! wär ich doch dafür;
  B-act: .....|circle..|tap.....@loc-1.....
  A-ver: |ja: genau |entweder HIER |außerhalb |hier is SCHWIERig
  A-lH:  |bwrD>>>>>>|place (loudly)
  A-rH:  |dG>>>>>>>@loc-2.....|tap
           *3           *4           *5
```



Kurz darauf bricht A die Aktivität ab und verbalisiert selbst einen Vorschlag zur Platzierung. Dabei zieht er das linkshändig geführte Objekt „H“ leicht zurück (*3), stellt es akustisch wahrnehmbar auf der Karte ab (*4) und weist deiktisch mit seiner rechten Hand einen alternativen Platzierungsort aus („loc-2“), der auf der gegenüberliegenden Seite der Karte liegt (*5). Betrachten wir B's Sichtfeld während dieser Vorschlagsepisode, wird deutlich, dass er in Fokussierung auf die kontemporär relevante Interaktionsaufgabe „Hotel“ den rechtshändig indizierten Alternativvorschlag „loc-2“ seines Ko-Partizipienten A nicht wahrnehmen kann. Stattdessen interpretiert er A's „provisorische Platzierung“ als intendiertes Angebot und orientiert sich zum vermeintlich vorgeschlagenen Lokalisationsort (*3+*4). In der sequentiellen Folgehandlung B's findet diese Teilnehmerinterpretation eine retrospektive Bestätigung: Auf den irrtümlich inferierten Platzierungsvorschlag reagierend enactiert B eine Relevanzsteigerung seines eigenen Vorschlags durch ein in prosodischer Nachdrücklichkeit markiertes „HIER, (.) !HIER! wär ich doch dafür;“ in Kombination mit einem ostentativen Fingertippen auf jenen vorgeschlagenen Lokalisationsort „loc-1“, den er als Standort des „Hotels“ präferiert. Analog zur Fehldeutung B's ist für A nicht erkennbar, dass sich B's Reaktion auf den falschen Referenten – die „provisorische Platzierung“ des Objektes bezieht. Infolgedessen muss A davon ausgehen, dass sein alternativer Platzierungsvorschlag von B wahrgenommen und geprüft wurde.

```
03 A-ver: |HIER is es gut (.) genau;|
  A-lH:  |place>>>>>>>>>>>>>>>>>>@loc-1|
           *6a-c
```



In Reaktion auf B's Insistieren, den von ihm vorgeschlagenen Standort zu ratifizieren, lenkt A ein und platziert das „Hotel“ auf den fortwährend ostentativ ausgewiesenen Lokalisationsort „loc-1“ (*6a-c).

Die analytische Rekonstruktion des Falles zeigt, dass Desambiguierungshinweise zur Etablierung von Ko-Orientierung (Blick-, Kopf-, Posturorientierung) in akzidentell-emergenten, multiplen Orientierungsangeboten unter der Einschränkung des *mutual monitoring* in ihrer semiotischen Relevanz nicht wirksam werden. Dabei konnte jedoch auch demonstriert werden, dass anschlussfähige nächste Handlungen Divergenzen zwischen der Orientierungskonstitution und ihrer Rezeption ‚überlagern‘. In diesem Sinne stellen interaktionale Fortsetzungsangebote orientierungsrelevante Strukturprinzipien dar, um die Bearbeitung interaktionaler Aufgaben zielgerichtet – im Fokus der Bearbeitung einer globalen Aufgabe – voranzutreiben.

3.3 Die Herstellung von Normalität: “wait & see” als Handlungsprinzip

Im vorigen Kapitel haben wir analytisch dargelegt, dass Fälle, in denen die fragmentarische Wahrnehmung der Interaktanten unterschiedliche Interpretationen zwischen der Bedeutungskonstitution und -rezeption eines Orientierungsverfahrens erzeugt, im Oberseeszenario der AR-Bedingung nur selten auftreten (“deviant cases”). In Bezug auf diese empirische Beobachtung stellt sich die Frage nach der phänomenologischen Ursache: Warum bilden divergente Interpretationen einer Orientierungsgestalt, trotz der gesteigerten Koordinierungsanforderungen und visueller Dispersionseffekte im *mutual monitoring*, nur vereinzelt auftretende Phänomene?

Unsere Analysen verweisen diesbezüglich auf ein gesteigertes Toleranzverhalten der Interaktanten, in der die Präferenzorganisation weitaus offener gehandhabt wird als in face-to-face Situationen: Dispräferierte Reaktionen oder ein Ausbleiben konditionell relevanter Folgehandlungen bilden keine grundlegenden Verletzungen normativer Erwartungen, sondern werden von den Interaktanten akzeptiert. Die Teilnehmer operieren dabei mit einer “wait & see” Strategie (s.a. Fox et al. 1996; Hayashi 1999), in der abgewartet wird, ob die dispräferierte Reaktion Anschlusshandlungen ermöglicht. Dieses organisatorische Prinzip wird im nachfolgenden Fragment am Beispiel eines Objektidentifikationsproblems deutlich.

FRAGMENT 23:

AR-Baseline:	VP-Gruppe 10 (VP-A: Investor / VP-B: Naturschützer)
Zeit:	19:56 - 20:12

```
01 B-ver: |ne ich weiß GAR nich was das is HIER|
    B-gaz: |@NT.....
    A-gaz: |@NT.....
    A-act: |ready-to-grasp-hand>>>NT
                                         *1a-c
```



Zu Beginn des Fragments verbalisiert B Probleme in der Identifikation eines Objekts (“NatureTrail (NT)”) (01), das er beidhändig greifend in den ‚intimen Anschauungsraum‘ seiner visuellen Wahrnehmung geführt hat (*1a). Während beide Interaktanten auf das unspezifische Objekt ko-orientiert sind (*1a+b), signalisiert A in einem ersten Schritt ein Hilfsangebot, indem er seine linke Hand in palmar geöffneter Greifbereitschaft zum Objekt führt (*1b+c). B’s Sichtfeld zu diesem Zeitpunkt lässt erkennen (*1a), dass er in Fokussierung auf das unbekannte Objekt A’s ausgestreckte Greifhand nicht wahrnehmen kann.

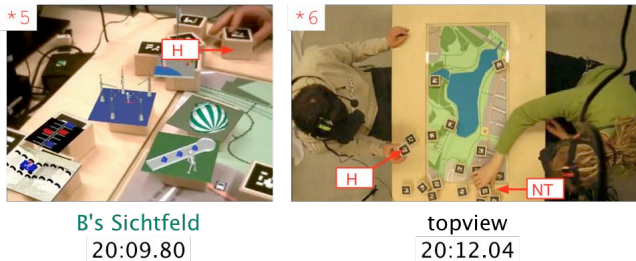
jedoch A, der die Anfrage nicht bearbeiten kann, da er zu diesem Zeitpunkt in der durch die Korrektur initiierten Aufmerksamkeitsverlagerung auf das am Kartenrand situierte Objekt „KT“ orientiert ist (*4a+*4c).

```

05 A-ver: |was ham wa denn HIER noch,|
A-act: |grasp>>>@H...|lift>>>hold.....|
B-gaz: |((A-act in B's fov))
B-rH: |place(NT) >>>>>>>>>@stack|

```

*5 *6



In Umkehrung zum eingangs geschilderten Ausfall konditioneller Relevanzeinlösung (A's Angebot zur Objektidentifikation) ist es nun Interaktant B, auf dessen salientes Orientierungsangebot A in dispräferierter Form reagiert: Anstatt sich erneut der Identifikation des Objekts "NT" zuzuwenden, greift A das Objekt "Hotel" und bereitet eine neue Interaktionsaufgabe vor („was ham wa denn HIER noch,“ (05)) (*5). Der "wait & see" Strategie als handlungsleitendes Prinzip folgend, eröffnet sich für B an dieser Stelle erst nach 4,5 Sekunden (vgl. 04) zwar eine dispräferierte, jedoch anschlussfähige Reaktion, auf die er mit der Aufgabe der Objektidentifikation „NT“ reagiert und das Objekt ‚namenlos‘ zurück an den Kartenrand stellt (*6).

3.4 Weitere Koordinierungsressourcen zur Herstellung von Ko-Orientierung

Die empirisch beobachtbare Robustheit interaktionaler Verfahren, mit der die Interaktanten – trotz Einschränkungen im *mutual monitoring* – die Bearbeitung lokal anstehender Aufgaben koordinieren und das Handlungsgeschehen mit Blick auf die globale Aufgabenanforderung sequentiell vorantreiben, ist letztendlich auch durch ein vielfältiges, semiotisch differentes und multimodales Repertoire an Orientierungsressourcen bedingt, derer sich die Interaktanten in situativ-adaptiver Kombinatorik bedienen. In der Analysesektion 3.2 haben wir bereits auf die Orientierungsrelevanz lokaldeiktischer Ausdrücke und der interaktionalen Fähigkeit zur Rekonstruktion deiktischer Gesten verwiesen:

- Lokaldeiktische Ausdrücke und deiktische Gesten
 - a) Lokaldeiktische Ausdrücke als "attention getter" (Mondada 2012) und "shifter" (Jakobson 1971) mit ‚zeitlicher‘ und ‚verweisender Orientierungsfunktion‘
 - b) Rekonstruktion deiktischer Gesten (Inferenz des Zeigeziels aus der Retraktionsphase)

Auch wenn die analytisch dargelegten Verfahren in ihrer Relevanz zur Aufmerksamkeitsmanipulation und Herstellung von Ko-Orientierung in der AR-Bedingung grundlegend sind, bilden sie die empirisch beobachtbare Robustheit interaktionaler Verfahren nur unzureichend ab. Im Sinne einer möglichst umfassenden Topographie der Teilnehmerorientierung ist die Liste um die folgenden Verfahren zu ergänzen, die in unterschiedlich ausgeprägter situativer Relevanz und semiotischer Assemblierung auftreten:

- Diskursstrukturierende Partikeln, Aufgabenratifikationen & -eröffnungen (vgl. 3.4.1)
 - a) Prospektive Orientierungsfunktion (3.4.1.1)
 - b) Retrospektive Orientierungsfunktion (3.4.2.2)

- Kohärenz- & Kohäsionsprinzipien, anhand derer Interaktanten (Re-)Orientierungsrelevanzen aus dem situativen Handlungsgeschehen ‚hypothetisieren‘ (3.4.1.3)
- Kopf- und Körperausrichtung als Orientierungsressource (in der AR-Bedingung eher selten)
- Signalisierung von Handlungsbereitschaft durch Hand- Armausrichtung (Bsp.: Fragment 28, *1)
- Prestrukturierung der materiellen Interaktionsumgebung und Vorbereitungshandlungen des Referenzobjekts (vgl. 3.4.2)
 - a) Präparation des Referenzobjekts im Vorfeld der Aufgabenankündigung (3.4.2.1)
 - b) Objektpräparation als Hinweis zur Ko-Konstruktion des Referenten (3.4.2.2)
- Steigerung der Objektsalienz (vgl. 3.4.3)
 - a) Anheben und Vorzeigen (bei initialer Objekteinführung) (3.4.3.1)
 - b) Präsentierendes Abstellen (“placing-for-procedure” i.S.v. Clark 2005; Bsp.: Fragment 28, *6)
 - c) Ostentativ-akustische Marker (z.B. Lautstarkes Abstellen der Objekte; Bsp.: Fragment 22, *4)
 - d) Trajektorische Modifikationen (zur Relevanzmarkierung bereits gegriffener Objekte) (3.4.3.2)
 - e) objektmanipulative Relevanzmarkierung des Referenten durch Kippen, Neigen, Drehen etc. (bei konkurrierenden Orientierungsangeboten, z.B. zwei gegriffene Objekte) (3.4.3.3)

Die aufgelisteten Koordinierungsressourcen zur Aufmerksamkeitsmanipulation und Herstellung von Ko-Orientierung bilden den thematischen Kern der nachfolgenden Analysesektion. Eine analytisch umfassende und ausdifferenzierende Rekonstruktion des mannigfaltigen semiotischen Repertoires ist dabei im Rahmen dieser Arbeit nicht zu leisten, so dass wir vor allem diejenigen Verfahren analytisch fokussieren, denen in der AR-Bedingung eine übergeordnete Koordinierungsrelevanz der Partizipientenorientierung zuteil wird (in der Liste durch die entsprechenden Absatznummern gekennzeichnet).

3.4.1 Diskursstrukturierende Partikeln, Aufgabenratifikationen und -eröffnungen

Diskursstrukturierende Partikeln haben nicht nur als gefüllte Pausen oder Hesitationssignale (mh::/ jahier also::/ so::/ okay:: etc.) eine überbrückende Qualität, sondern sie strukturieren auch komplexe Koordinierungsanforderungen, deren Orientierungsrelevanz sich vor allem im Übergang interaktionaler Aufgaben manifestiert. In diesem Zusammenhang merken Dausendschön-Gay & Krafft (2009) am Beispiel der Verkaufsinteraktion eines mobilen Markstandes für die – ausschließlich bei den Verkäufern und nicht bei den Kunden – frequente Verwendung des Markers „so“ an:

“So is used not to structure discourse (so the term “discourse marker” does not really seem appropriate) but to structure the interaction; it marks the end of an activity (or part of an activity) and indicates that the next one can now be tackled. So is an articulated prosodic gesture used to structure the interaction.” (Dausendschön-Gay & Krafft 2009: 255)

Die interaktionsstrukturierende Funktion steht im Mittelpunkt dieser Analysesektion, auch wenn wir nachfolgend aufgrund der terminologischen Übersättigung in der Literatur (Brinton (1996: 29) führt 22 Termini auf; s.a. Jucker & Ziv 1998: 1f.) stellvertretend die usuelle Bezeichnung ‚Diskursmarker‘ oder ‚diskursstrukturierende Marker/Partikeln‘ verwenden. Als interaktionaler Mechanismus zur wahrnehmbaren Gestaltung struktureller Einheiten sind Diskursmarker grundlegend: „sie machen das sprachliche Handeln interpretierbar, d.h. sie kommentieren es, steuern es, sichern seine formalen Voraussetzungen oder verankern es im Kontext“ (Auer & Günthner 2003: 1). In den nachfolgenden Analysen wird einerseits erkennbar, dass diese Steuerung und kontextuelle Verankerung bidirektional, in prospektiver und retrospektiver Orientierungsfunktion konzipiert werden kann. Andererseits kann gezeigt werden, dass diese Orientierungsrichtung auf zurückliegende, gegenwärtige oder nachfolgende Interaktionsepisoden in grammatikalischen Kategorisierungen einzelner Diskursmarker nur abstrakt erfasst werden kann, da die sequentielle Einbettung des Diskursmarkers und die präzise zeitliche Relationierung zum kontextuellen Geschehen seine strukturierende Qualität bestimmt.

3.4.1.1 Prospektive Orientierungsfunktion

In Kapitel 2.2 haben wir zwei phänomenologische Gestaltformen – die attrahierende und mediatisierende Aufmerksamkeitsbindung – des *mutual monitoring* in Transitionsphasen interaktionaler Aufgaben der face-to-face Bedingung rekonstruiert. Die analytischen Ergebnisse haben unterschiedliche Orientierungs- und Situierungsqualitäten offen gelegt, die in Abhängigkeit ihres Potentials stehen, Unklarheitsbedingungen in Transitionsphasen interaktionaler Aufgaben zu desambiguieren und eine für beide Parteien eindeutige Fortsetzungsoption in ‚sequentieller Lesart‘ aufzuzeigen (s.a. Kapitel 2.3). Vor diesem Hintergrund bilden instantane situative Desambiguierungsgestalten (z.B. Fragment 2, Fragment 4), in denen Interaktanten in zeitlich und multimodal verdichteter Ausdrucksform eine akute Relevanzminderung der gegenwärtigen und -steigerung der nächsten Aufgabe enactieren, ein besonders ergiebiges Orientierungsformat, das eine Verkürzung der Transitionsphase bedingt. Dabei geht es nicht darum, dass Übergangsphasen einen interaktional unhaltbaren, in jedem Falle zu minimierenden Zustand markieren (dies ist nicht der Fall, z.B. Planungsphasen etc.), sondern darum, dass sie eine Phase der Unklarheit konstituieren, in der die Interaktanten zwar individuell orientiert, aber interaktiv desorientiert sind – folglich das „joint project“ (i.S.v. Clark 1996) nicht vorantreiben können.

Wenn wir diese empirischen Beobachtungen mit denen der AR-Bedingung kontrastieren ist festzustellen, dass die Koordinierung des „What to do next?“ (Garfinkel 1967) von den Teilnehmern anders organisiert wird. Eine instantane situative Desambiguierung, die den Übergang interaktionaler Aufgaben – im Sinne eines Minimierungsanspruches transistionaler Phasen – in situ erkennbar werden lässt, erscheint dabei jedoch kaum möglich, da dieses Orientierungsformat in exzeptioneller Weise auf Projektionen kinesischer Displays und ein intaktes *mutual monitoring* angewiesen ist. Vielmehr wird diese Koordinierungsanforderung in der AR-Bedingung größtenteils verbalsprachlich gelöst. Dabei weisen die Transitionsphasen zum Teil deutliche Expansionen auf, da die verbalsprachliche Koordinierung einem weitaus stärkeren Sukzessionsdrang unterworfen ist als das kinesische Ausdrucksarsenal, dessen simultan-semiotisches Potential unmediatisiert und in vollem Umfang in der face-to-face Bedingung genutzt werden kann. Eine erfolgreiche Koordinierung zwischen interaktionalen Aufgaben, in der die Herstellung und Etablierung von Ko-Orientierung auf ein nächstes „joint project“ das Kooperationsziel darstellt, ist dabei an unterschiedliche Teilaufgaben gekoppelt, deren sukzessive Bearbeitung in der AR-Bedingung primär verbalsprachlich zwischen den Teilnehmern ausgehandelt und wechselseitig angezeigt wird. Die sequentielle Bearbeitungslogik dieser Teilaufgaben kann wie folgt skizziert werden:

- 1) Projektion einer bevorstehenden Aufgabeterminierung (Aufgabe-1)
- 2) Signalisierung der faktischen Aufgabeterminierung durch Ratifizierung (Aufgabe-1)
- 3) Gegenseitig wahrnehmbare Konstitution einer Transitionsphase
- 4) Gegenseitig wahrnehmbare Terminierung der Transitionsphase
- 5) Projektion einer bevorstehenden Aufgabenkonstitution („Aufgabe-2“)
- 6) Koordinierter Aufgabenbeginn (Aufgabe-2)

Die Bearbeitungslogik dieser Teilaufgaben ist hierbei nicht als striktes Ablaufschema zu verstehen, sondern als elidier- und expandierbare Grundordnung der Orientierung im Übergang interaktionaler Aufgaben. Während beispielsweise ein koordinierter Aufgabenabschluss für den Beginn einer neuen Aufgabe einen zentralen Orientierungsmechanismus konstituiert, ist die Existenz und zeitliche Ausdehnung der Transitionsphase variant und nicht obligat. Zudem können einzelne Phasen zeitlich zusammenfallen und wechselseitig substituiert werden. So markiert beispielsweise die Projektion einer nächsten interaktionalen Aufgabe gleichermaßen die Beendigung einer Übergangsphase usw. Das organisatorische Prinzip der verbalsprachlichen Koordinierung auf nächste Interaktionsaufgaben wird im nachfolgenden Fragment in der analytischen Rekonstruktion der Teilnehmerverfahren verdeutlicht.

Wir steigen in die Analyse der Interaktion nach einer längeren Aushandlungsepisode ein, in der sich das von Teilnehmer B vorgeschlagene Objekt „Hotel“ (ab 19:48 min) in einem Lokalisationsproblem sedimentiert, welches vor allem darin begründet liegt, dass unterschiedliche Objekte – das „Hotel“ (H) und der „CarPark“ (CP) – zusammengehörig positioniert werden sollen. Das Transkript setzt zu einem

Zeitpunkt ein, kurz nachdem Teilnehmer A einen Lösungsvorschlag verbalisiert hat, den B nachfolgend als „BESTe variante-“ ausweisend ratifiziert.

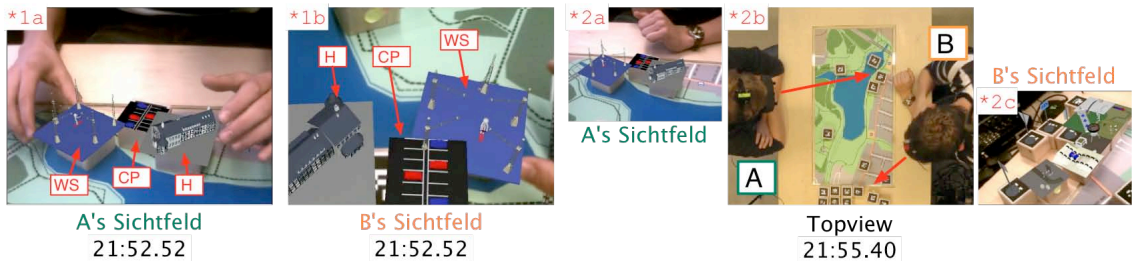
FRAGMENT 24:

AR-Baseline:	VP-Gruppe 06 (VP-A: Naturschützer / VP-B: Investor)
Zeit:	21:47- 22:13

01 B-ver: |gut; das machen wir sogar so; (0.8) is wahrscheinlich die
 B-gaz: |@H+CP+WS.....
 A-gaz: |@H+CP+WS.....

02 B-ver: BESTe variante- da sparn wir uns nochmal GELD für den
 B-gaz:
 A-gaz:
 *1a+b

03 B-ver: =für die neuen STRAßen die wir bauen müssen =OKAY;
 B-gaz:>>>>>>@stack.....>>>>>>
 A-gaz:
 *2a-c



04 B-ver: =<<all>ich glaub wir müssen uns nen=bisschen beeilen->
 B-gaz: ((blickt auf seine Uhr.....))>>>@H+CP+WS.....
 A-gaz:

Zu Beginn des Fragments ratifiziert B verbalsprachlich mit „gut; das machen wir sogar so;“ (01) den Lösungsvorschlag seines Ko-Partizipienten, das Verhandlungsobjekt „Hotel“ (H) zusammen mit dem „Car-Park“ (CP) in unmittelbarer Nachbarschaft zum Objekt „WasserSki“ (WS) an der Uferseite des Sees zu platzieren. Im Rahmen der präkontextuellen Konfiguration eines interaktional fossilisierten Lokalisationsproblems projiziert B's verbalsprachliche Ratifikation des Lösungsvorschlags handlungsschematisch die bevorstehende Aufgabeterminierung (Punkt 1). Während dieses Zeitintervalls sind die Teilnehmer auf die für die gegenwärtige Interaktionsaufgabe konstitutive Objektgruppe ko-orientiert (*1a+b). In B's verbalsprachlicher Expansion der Ratifikation „is wahrscheinlich die BESTe variante- [...]“ (01-03), in der A's Lösungsvorschlag nicht nur einfach konzidiert, sondern finalisierend argumentativ gestützt wird, fällt B's Aufmerksamkeitsverlagerung zu den übrigen Objekten am Kartenrand gegen Turnende auf (03, *2b+c). Gewissermaßen bereitet er hier durch alltäglich sedimentierte face-to-face Verfahren den Übergang zu einer neuen interaktionalen Aufgabe vor, der jedoch unter den limitierten Möglichkeiten im *mutual monitoring* von seinem Ko-Partizipienten – äquivalent zu den empirischen Beobachtungen aus Fragment 15 – nicht mitvollzogen werden kann (*2a). Im Gegensatz zu Fragment 15 entsteht hier jedoch keine interpersonelle Handlungsparallelität, da B sequentiell nachfolgend die interaktionale Aufgabe diskursstrukturrell mit einem prosodisch akzentuiertem „OKAY;“ faktisch terminiert (03) und die interaktionale Fortsetzung verbalsprachlich mit „<<all> ich glaub wir müssen uns nen=bisschen beeilen->“ einfordert (04) (Punkt 2 & z.T. Punkt 3). Auch wenn sein Insistieren durch einen Blick auf die Armbanduhr indexikal gestützt wird, kann A diesen Dringlichkeitsverweis aufgrund der Sichtfeldeinschränkung nicht wahrnehmen. Infolgedessen ist hier die Verbalsprache in isolierter Signalisierungsfunktion koordinierungsrelevant.

A reagiert auf B's sukzessiv hergestellte Fokussierung auf den Bearbeitungsbeginn einer neuen interaktionalen Aufgabe, indem er die diskursstrukturierende Startpartikel „also-“ in der Bestätigung des Aufgabengesuchs seines Ko-Partizipanten aufgreift („ja hier also-“, 07) und somit eine konkrete Objektauswahl zur bevorstehenden Aufgabenbearbeitung anzeigt. Erst im Greifen des ausgewählten Objekts „PettingZoo“ (PZ) (*5a) und der zeitgleichen Objektmanipulation B's (*5b) bestätigt sich für die Teilnehmer auch visuell-semiotisch die gemeinsame Orientierung auf die am Kartenrand situierten Objekte. In gegenseitig ratifizierter Ko-Orientierung schlägt A das Objekt (*6a+b) als nächste Interaktionsaufgabe vor und identifiziert es verbalsprachlich mit „so=ne HUNDEspielwiese; (08).

3.4.1.2 Retrospektive Orientierungsfunktion

Neben der prospektiven Orientierungsfunktion konstituieren diskursstrukturierende Partikeln mit retrospektiver Orientierungsfunktion (z.B. „ja wobei:“, „zumal“ oder „na=ja:“) eine zentrale Koordinierungsressource in der AR-Bedingung, die sich in einer Re-Orientierung des Ko-Partizipanten auf die kontemporäre Aufgabenbearbeitung auswirkt.

Darüber hinaus dient das nachfolgende Beispiel jedoch auch im Sinne der „deviant case analysis“ als Gegenbeispiel der prospektiven Orientierungsfunktion, in der sich *ex negativo* die Bedeutung der diskursstrukturierenden Orientierungsphasen im Übergang interaktionaler Aufgaben bei der verbalsprachlichen Koordinierung manifestiert (vgl. Punkt 1-6, 3.4.1.1). Hierbei wird deutlich, dass die alleinige ‚Abschlussratifikation‘ einer gegenwärtigen Aufgabe kein hinreichendes Kriterium der Aufgabenbeendigung darstellt, die eine Zuwendungsoption zu einer neuen interaktionalen Aufgabe projiziert. Grund dafür ist die prinzipielle Expandierbarkeit der gegenwärtigen Aufgabenbearbeitung, auch wenn diese von beiden Parteien ‚abschließend‘ ratifiziert wurde. Für eine interpersonell enge Koordinierung, in der beide Parteien ihre Handlungsaktivitäten auf fundierte Hypothesen zur gegenwärtigen Handlungsaktivität ihres Gegenüber stellen, erscheint es somit relevant, nicht nur den Aufgabenabschluss zu ratifizieren, sondern den strukturellen Übergang zur nächsten Aufgabe zu projizieren (Projektion der Transitionsphase) oder eine nächste Aufgabenbearbeitung direkt anzuzeigen, die sich wahrnehmbar von der vorigen Aufgabe unterscheidet.

An dieser Stelle ist jedoch darauf hinzuweisen, dass eben jene „interpersonell enge Koordinierung“ innerhalb des Oberseeszenarios, in der sich die interaktionale Relevanz vornehmlich auf die Objektauswahl und -platzierung erstreckt, keine notwendige Voraussetzung zur Bearbeitung der globalen Interaktionsaufgabe konstituiert. So ist auch im vorliegenden Fall die Verlagerung der Aufmerksamkeit direkt nach einer vermeintlich abgeschlossenen Aufgabenbearbeitung nicht von interaktionaler Bedeutung, da die Fortsetzung der gegenwärtigen Aufgabe, eingeleitet durch die retrospektive Orientierungsfunktion eines „ja wobei“, unmittelbar zur Re-Orientierung der abgewendeten Ko-Partizipantin führt. Die Konzeption als „abweichender Fall“ erscheint uns dennoch im Hinblick auf mögliche Anwendungsszenarien in der Medizin, technischen Wartung etc. (vgl. Azuma 1997) relevant, in der eben genau jene „enge interpersonelle Koordinierung“ für die interaktionale Aufgabenbearbeitung relevant und im *ex negativo*-Fall‘ folgenreich sein kann.

Das analytische Verständnis der teilnehmerrekonstruktiven Verfahren und Aufmerksamkeitspräferenzen im folgenden Fragment hängt entscheidend von einer präkontextuellen Handlung ab, in der VP-A siebzehn Sekunden vor dem dargestellten Transkriptausschnitt (ab Minute 18:41) ihrer Ko-Partizipantin den Vorschlag unterbreitet „dann lass und doch SA:gen =<<all>weiß nich> =DU kriegst dein WASSERski und ich mach dann so=nen (-) TOLLEN WANDERweg der alles erklärt; [...]“. Zusammen mit B's sequentiell anschließender Vorschlagsratifikation „joa: das find ich GUT;“ wird hier interaktiv ein duales Bearbeitungsschema der nächsten interaktionalen Aufgaben etabliert, das der nachfolgenden Interaktionsepisode zugrunde liegt.

die wiederum von A mit „genau;“ bestätigt wird. Auf diese Abfolge von dreizügigem Ratifikationsformat reagiert A, indem sie sich von der kontemporären Interaktionsaufgabe “NT” abwendet und das zweite Element “WasserSki” (WS) des präkontextuell etablierten Bearbeitungsschemas fokussiert (*2a+b (das Sichtfeld der Akteure ist in der Draufsicht durch gestrichelte und farbig den Akteuren zugeordnete Linien markiert)). Die teilnehmerrekonstruktive Wahrnehmung A’s lässt hier folglich relativ zuverlässig erkennen, dass für sie der erste Teil des Bearbeitungsschemas ‚abgearbeitet‘ ist. Für B hingegen ist dies nicht der Fall: In konstanter Orientierung auf den das Objekt “NT” inkludierenden Raumausschnitt (*2c) leitet sie nach 1.4 Sekunden durch den diskursorganisatorischen Expansionsmarker „ja wobei“ eine Handlungserweiterung ein, auf die A mit der Re-Orientierung auf das diskursrelevante Objekt “NT” reagiert (04, *3a-c).

In Fällen dieser Art manifestieren sich neben der situativ vorhandenen Expansionsmöglichkeit ratifizierter Aufgabenabschlüsse auch interaktionale Zuordnungsprobleme ratifizierender Marker. Ohne die Möglichkeit der visuellen Einsichtnahme desambiguierender Orientierungshinweise innerhalb multimodaler Ausdrucksgestalten ist für die Interaktanten kaum ermittelbar, worauf sich der Gültigkeitsanspruch einzelner oder sequentiell geordneter Ratifikationsmarkierungen de facto bezieht. Dennoch wirken sich punktuelle Desorientierungen im Übergang interaktionaler Aufgaben durch das Potential re-orientierender Verfahren nicht ‚folgeschwer‘ aus, da die Interaktanten flexibel auf situative Präferenzstrukturen reagieren und die interaktionale Fortsetzung eine ‚relevantere‘ Ordnungsstruktur konstituiert als der individuelle, intrinsisch motivierte Progressionsanspruch (vgl. 3.3).

3.4.1.3 Abweichender Fall zur retrospektiven Orientierungsfunktion

Eine (re-)orientierende Funktion liegt nicht allein in der ‚Semantik‘ des verwendeten Diskursmarkers als prospektiv- hinweisendes oder retrospektiv- rückweisendes Sprachelement begründet, sondern hängt entscheidend von seiner sequentiellen Einbettung im Interaktionsgeschehen ab. Ein „wobei:“, „zumal“ oder „na=ja:“ kann sowohl in Ko-Orientierung auf einen Sachverhalt/eine Entität konzessiv eine Aussageneinschränkung oder -präzisierung ankündigen als auch im Sinne eines Expansionsmarkers der gegenargumentativen Handlungserweiterung einer vermeintlich abgeschlossenen und ko-ratifizierten Interaktionsaufgabe dienen (zur funktionalen Vielfalt des Pronominal-/Relativadverbs *wobei* vgl. Günthner 2000; Auer & Günthner 2003). Unter einer praxeologischen Perspektive auf Interaktion erscheinen hier weniger grammatikalische Funktionen oder Grammatikalisierungsprozesse der Bedeutungsverschiebung relevant als die konkrete situative Nutzung, in der Adverbformen wie „wobei“ unterschiedliche diskursstrukturelle Funktionen in einem spezifischen interaktionalen Kontext übernehmen können. Die situative Orientierungsfunktion markiert somit die relevante Bezugsgröße der analytischen Rekonstruktion unterschiedlicher Verwendungsoptionen.

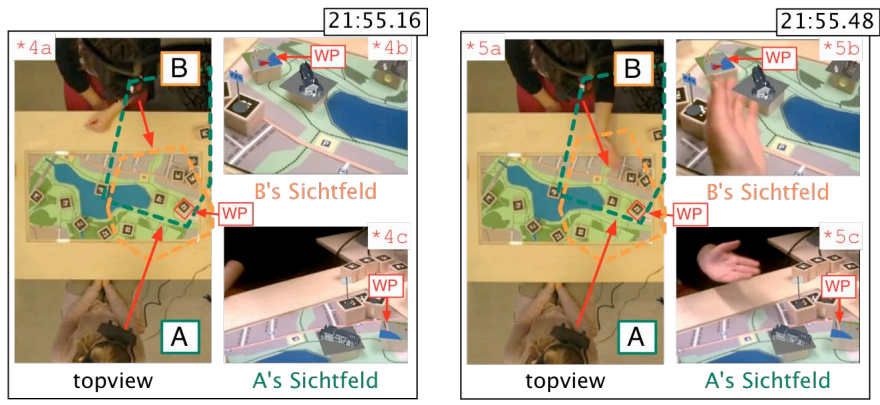
Vor diesem Hintergrund verdeutlicht das nachfolgende Beispiel, dass sich die interaktiv-pragmatische Funktion jener Adverbform „wobei“ nicht a priori durch eine Bedeutungs- und Funktionsbestimmung festschreiben lässt, sondern in fallspezifischer Varianz aus dem Handlungskontext emergiert. Zur Initiierung einer Selbstkorrektur wird die Adverbform in eine bereits etablierte Fortsetzungsaktivität eingebunden, in der die bestehende Ko-Orientierung auf das gegenwärtige Handlungsgeschehen eine etwaige Re-Orientierungsfunktion gegenstandslos werden lässt.

Der Beginn unserer Analyse zeigt den interaktionalen Übergang zwischen der gerade abgeschlossenen Aushandlungsepisode “Hotel” (H) und der interaktiven Herstellung einer neuen lokalen Interaktionsaufgabe, in der VP-A das Objekt “WaterPark” (WP) in den Diskurs einbringt. Die verbalsprachliche Einführung des Diskursobjekts erfolgt durch eine semantische Bezugnahme zum Präkontext, in der das Objekt “WP” (hier als „Spielplatz“ bezeichnet) – äquivalent zum vorig ausgehandelten Objekt “H” – als integraler Bestandteil der topographischen Oberseekarte hervorgehoben wird.

erkennen, wie B ihre Ko-Partizipantin A ‚strategisch vorteilhaft‘ zuerst auf einen Spielplatz des bereits enaktierten Raumausschnitts um das ‐Hotel‐ orientiert (*2a+b) und dann ihre Aufmerksamkeit aus der gemeinsamen Orientierung auf einen situativ ‚unbekannten‘ Raumausschnitt lenkt, in welchem sie auf den zweiten Spielplatz der Karte verweist (*3a+b).

07 B-ver: |ja is die frage ob da nen zusätzlicher jetzt noch nötig wäre
 B-gaz: |@WP+H+BS.....
 A-gaz: |@loc-2... >>>>>WP+H+BS.....
*4a-c

08 B-ver: |°h WOBEI: das| is natürlich ähm: in gewisserweise auch
 B-act: |dG>>>WP>>>HP|
 B-gaz: |@WP+H+BS.....
 A-gaz: |@WP+H+BS.....
*5a-c



09 B-ver: nen ANderer =sieht nach nem ANderen spielplatz aus
 10 B-ver: |=wenn es DA: halt anlage zum WASSER gibt |(xxx)
 B-act: |dG>>>>>WP.....>>>HP.....|
 A-ver: |^mhm ja;|

Unmittelbar im Anschluss der Zweifach-Orientierung auf die beiden bestehenden Spielplätze etabliert B eine logisch implikative Fortsetzungsaktivität, die bereits im initialen Hinweis vorhandener Spielplätze (01+02) angelegt war: In expliziter Frageform ausgewiesen („ja is die Frage“) konstituiert Sprecherin B im *first-pair part* die Präferenzstruktur einer diskursiven Bearbeitungsanforderung („ob da nen zusätzlicher jetzt noch nötig wäre“, 07). Bemerkenswert ist hierbei, dass sich A auf Grundlage der logischen Emergenz einer der Verweishandlung sequentiell subordinierten Bearbeitungsanforderung bereits mit Beginn der Fragekonstitution auf den relevanten Raumausschnitt re-orientiert, in dem das Aushandlungsobjekt ‐WP‐ situiert ist (07). Hier manifestieren sich Kohärenz- und Kohäsionsprinzipien der Orientierung, anhand derer Interaktanten (Re-)Orientierungsrelevanzen aus dem situativen Handlungsgeschehen ‚hypothetisieren‘. In Ko-Orientierung auf den relevanten Raumausschnitt und das in ihm befindliche Objekt ‐WP‐ (*4a-c) initiiert B – sequentiell an die konsistente Fortsetzungsaktivität gekoppelt – durch die Konzessivform „WOBEI:“ eine Selbstkorrektur der etablierten Präferenzstruktur im *second-pair part* (08), in der auf die Unterschiedlichkeit der bestehenden und der gegenwärtig ausgehandelten Objektspezifikationen hingewiesen wird (08-10). Die sequentielle Einbettung in eine etablierte Ko-Orientierungssituation richtet die Funktionalität des Diskursmarkers „WOBEI:“ auf eine konzessive Lesart aus, in der eine etwaige Re-Orientierungsfunktion (vgl. Fragment 25) obsolet ist.

3.4.2 Prestrukturierung der Interaktionsumgebung – Präparation des Referenzobjekts

Zum Zwecke der gemeinsamen Orientierung auf kooperativ zu bearbeitende Aufgaben (“joint projects”) ,drängen‘ Interaktanten der materiellen Interaktionsumgebung Strukturen auf. Dies geschieht sowohl durch verbalsprachliche Raumausdrücke, mit denen die Akteure Relationen zwischen den Raumentitäten und der Beteiligungsstruktur (i.S.d. “participation framework”, Goodwin & Goodwin 2004) abbilden (vgl. 3.1 (Eingangsbeispiel ‚Wahrnehmungsraum‘; 3.2 (Lokaldeiktika)), als auch durch Manipulationshandlungen, wie wir im Folgenden analytisch darlegen. Der Strukturierung des Interaktionsraums liegt dabei eine Orientierungslogik zugrunde, in der die partikuläre Veränderung der Raumstruktur die Wahrnehmbarkeit ihrer subjektiven Aneignung ermöglicht. Der Raum ist damit nicht mehr ein ,objektiv‘ zugänglicher, sondern ein ,subjektiv‘ strukturierter, der als solcher entweder in situ wahrnehmbar wird oder ,wahrnehmbare Spuren‘ hinterlässt.

Die beiden folgenden Beispiele deuten darauf hin, dass die intrapersonelle Strukturierung des Interaktionsraums weniger ein ,aktives‘ Verfahren zum Zwecke der Ko-Orientierung als eine subjektive Raumaneignung repräsentiert, der vom Ko-Partizipanten eine Orientierungsrelevanz zugeschrieben wird. Dabei zeigt sich im Besonderen, dass die Herstellung von Ko-Orientierung deutlich vor der eigentlichen Referenzherstellung beginnt.

3.4.2.1 Präparation des Referenzobjekts im Vorfeld der Aufgabenankündigung

Das ,Aufdrängen‘ subjektiver Strukturen der Raumaneignung manifestiert sich im folgenden Fall durch eine vorbereitende Manipulationshandlung eines Objektes, welches erst im weiteren Interaktionsverlauf als Referenzobjekt konstituiert wird (vgl. Neumann, Schnier, Hermann & Pitsch 2013). Nach Abschluss einer terminierten Aufgabebearbeitung isoliert Teilnehmer B ein Objekt aus der Gesamtheit der übrigen Verhandlungsobjekte am Kartenrand (“stack”). Auffällig ist hierbei, dass der objektmanipulativen Isolationshandlung von B selbst zunächst keine interaktive Relevanz im Sinne einer präferierten Anschluss-handlung seines Ko-Partizipanten beigemessen wird. Vielmehr lenkt B die Aufmerksamkeit seines Ko-Partizipanten vorerst systematisch von der eigenen Objektmanipulation auf die Möglichkeit der Unterbreitung einer neuen, den Handlungsbedürfnissen seines Ko-Partizipanten entsprechenden Aufgabenkonstitution, bevor er sein präpariertes Objekt an geeigneter Stelle als faktisches Referenzobjekt etabliert und ihm die Bedeutung einer nächsten relevanten Interaktionsaufgabe zuweist.

FRAGMENT 27:

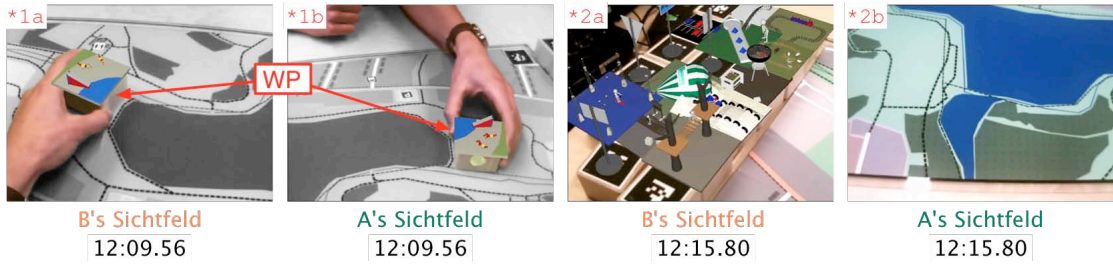
AR-Baseline:	VP-Gruppe 06 (VP-A: Naturschützer / VP-B: Investor)
Zeit:	12:09 - 12:53

```

01 A-ver: |DA:: wollen=sie=es HIN haben,|
    A-gaz: |@WP.....|
    B-ver: |JA;|
    B-act: |place (WP) >>>@loc...|grasp@WP.....>>>HP|
    B-gaz: |@WP.....|
           *1a+b

02 A-ver: |ähm:: °h JA ich denke AUCH
    A-gaz: .....>>>map~~~~~>>>WP.....
    B-ver: |ich denke mal (da sind=wa)|
    B-gaz: |>>>>@stack.....|
           *2a+b

```



03 A-ver: das: es DA ganz gut platziert is; =vielleicht ja;|
 A-gaz:|
 B-gaz:|

Den Ausgangspunkt der Analyse bildet eine Situation, in der die Interaktanten in Ko-Orientierung die Platzierung des Objekts “WaterPark” (WP) aushandeln (*1a+b). Als die das Objekt manipulierende Partei ,testet‘ B geeignete Orte an der Uferseite des Sees aus und unterbricht diese Testphase, indem er das Objekt kurzweilig auf einem partikulären Lokalisationsort (@loc, 01) abstellt. Darauf reagiert A durch die Nachfrage „DA:: wollen=sie=es HIN haben,“ (01) und der anschließenden visuellen Überprüfung der Karte nach potentiellen Alternativen (02, *2b). In Erwiderung der Nachfrage ratifiziert B den ,ausgetesteten‘ Lokalisationsort mit einem „JA,“, retrahiert seine objektführende Hand in ihre *home position* und verlagert seine Orientierung auf die verbleibenden Objekte am Kartenrand (01+02, *2a). Infolgedessen sind beide Interaktanten gegen Ende der Aufgabenbearbeitung “WP” auf unterschiedliche Orte der Karte orientiert (*2a+b).

12:20
 |
 04 B-ver: |SO:; was is IHNEN denn noch besonders !WICH!tig,
 B-act: |grasp>>>>>>>@H.....((zieht H aus stack))>>>>>>>HP
 A-gaz: |@stack((+B-act)).....((-B-act))
 *3 *4 *5
 05 A-ver: |°h (2.4) JA: also wie gesagt wir dürfen hier die GRÜNflächen
 A-gaz: |@stack.....(foc@H)....
 *6



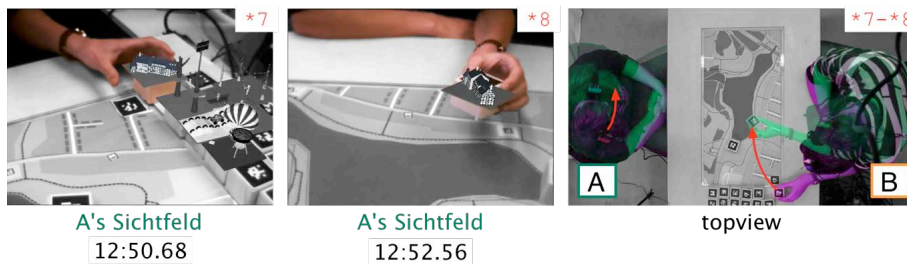
06 A-ver: natürlich nich beRÜHREN- und |ähm
 B-ver: |JA:: =aber das is ja kein
 07 B-ver: natURschutzgebiet =beSTIMMTE grünflächen können=wa ja ruhig-
 (...) Fortführung des Disputs

B reagiert auf A’s abschließenden Kommentar zur Interaktionsaufgabe “WP” (03), indem er den Übergang zu einer neuen interaktionalen Aufgabe markiert (SO:;), zeitgleich seine rechte Hand in Handlungsbereitschaft zur Objektmanipulation führt (*3) und seinen Ko-Partizipienten A mit „was is IHNEN denn noch besonders !WICH!tig,“ auffordert, ein nächstes Verhandlungsobjekt vorzuschlagen (04). Die Kombination aus verbalsprachlichem Startmarker „SO:;“ und B’s in Handlungsbereitschaft zur Objektmanipulation geführte rechte Hand bilden hierbei eine Orientierungsgestalt, anhand derer A seine Aufmerksam-

keit von der abgeschlossenen Interaktionsaufgabe “WP” auf die übrigen Verhandlungsobjekte am Kartenrand verlagert (*3). Zu diesem Zeitpunkt sind beide Interaktanten auf die noch zu verhandelnden Objekte am Kartenrand ko-orientiert, so dass A der objektmanipulativen Handlung B’s folgen kann: Simultan zur Anfrage einer neuen Aufgabenkonstitution, zieht B das Objekt “Hotel” einige Zentimeter aus der Anordnung der verbleibenden Objekte am Kartenrand (*4+*5). Dieser Salienzkonfiguration wird jedoch zunächst durch B’s explizite Anfrage „was is IHNEN denn noch besonders !WICH!tig.“ eine interaktive Relevanz abgesprochen.

Die Suffizienz dieser systematischen Aufmerksamkeitsmanipulation spiegelt sich im Sichtfeld seines Ko-Partizipanten wider: A kann zwar das initiale Stadium der Objektmanipulation erkennen (also die Aktivität als eine objektmanipulative Handlung wahrnehmen), bearbeitet dann jedoch B’s Anfrage durch eine visuelle Überprüfung der übrigen Verhandlungsobjekte (05, @stack) und verbalsprachlichen Bearbeitung der durch die Anfrage erzeugten Relevanzstruktur: Anstelle eines konkreten Objektvorschlags antwortet A mit einer Darstellung grundsätzlicher Prinzipien, die es in der weiteren Diskussion zu berücksichtigen gilt (05-07). Eher beiläufig exploriert er im Zuge dieser Grundsatzdarstellung das von B manipulierte Objekt “Hotel” (*6).

```
08 B-ver:  ALSO:: was natürlich noch  |!GANZ! WICHtig is
      B-act:  |grasp>>>>>>>>>>>>@H
                                     *7
09 B-ver:  is nen DICKES hotel;
      B-act:  lift>>>>>>>>>>hold...
                                     *8
```



Erst nachdem der aus A’s Grundsatzdarstellung entwickelte Disput über die Bebauung und den Erhalt von Grünflächen zum Erliegen kommt, präsentiert B (27 Sekunden nach der initialen Objektpräparation) das Objekt “Hotel” als nächste interaktionale Aufgabe. Die faktische Konstitution als Referenzobjekt erfolgt dabei in präziser Koordinierung zu einer turnübergaberelevanten Stelle (TRP), in der die Beendigung des Disputs in einer 2-sekündigen Pause mündet, die eine Neuverhandlung des Rede- und Handlungsrecht ermöglicht. Bemerkenswert ist hier vor allem die aktive Orientierung beider Interaktanten gegen Ende des Disputs auf das Objekt “Hotel”, aus der sich schlussfolgern lässt, dass A die sich sequentiell anschließende Objekteinführung seines Ko-Partizipanten B antizipiert. – Die initiale Objektvorbereitung macht eine interaktionale Relevanz erwartbar.

An dieser Stelle wird deutlich, dass beide Interaktanten interaktiv, jedoch ohne sich verbalsprachlich darüber zu verständigen, einen Handlungsplan der nächsten Aktivität entwickelt haben – B durch die manipulative Objektpräparation und A durch die inferentielle Erkenntnis, dass diese Aktivität interaktional relevant werden wird. Tatsächlich nutzt B die ungefüllte Pause und referenziert das initial vorbereitete Objekt “Hotel” mit dem *account* „ALSO:: was natürlich noch !GANZ! wichtig ist, gefolgt von der Objektidentifikation „is nen DICKES hotel;“ (08+09). Simultan zur verbalen Ankündigung greift er das isolierte Objekt aus dem “stack” (*7) und führt es ins Zentrum des gemeinsamen Wahrnehmungs- und Handlungsraums (*7-*8).

3.4.2.2 Objektpräparation als Hinweis zur Ko-Konstruktion des Referenten

Im zweiten Beispiel der Fallgruppe lässt sich das Verfahren der objektpräparativen Referenzherstellung in einer phänomenologisch modifizierten Form beobachten. Das Beispiel bildet die interaktionale Folgeaufgabe der Objektaushandlung “Hotel” (vgl. Fragment 27) und setzt unmittelbar nach einer Einigung der

objektzentrierten Bearbeitung spiegelt sich in dem salienzkonfigurierenden Verfahren des ‚Anhebens und Vorzeigens‘ zur Relevanzkonstitution singularer Objekte wider. Handlungsschematisch wird das Verfahren heterogen in die Bearbeitung unterschiedlicher interaktionaler Teilaufgaben eingebunden. Neben dem spezifischen Verfahren, das Referenzobjekt in salient angehobener Position als verweisende Instanz auf einen möglichen Platzierungsort zu nutzen (vgl. Abbildung 17, vgl. „poising“ bei Clark 2005: 512f.), verweist Abbildung 18 auf weitere handlungsschematische Kontextualisierungen, die im vorliegenden Beispiel konglomeratisch auftreten. Zu diesen zählen vor allem i) die Vorschlagskonstitution, ii) intra- und interpersonelle Objektidentifikationsprobleme und iii) die Bearbeitung von Objektidentifikationsproblemen.



Abbildung 18: Anheben und Vorzeigen als Salienzkonfigurationen der Vorschlagskonstitution (Teil A) und Verfahren zur Bearbeitung von Objektidentifikationsproblemen (Teil B)

Epistemisch interessant ist in diesem Zusammenhang die numerische Distribution objektmanipulativer Handlungen zwischen der AR- und F2F-Bedingung (vgl. Abbildung 19). Im kontrastiven Vergleich der Bedingungen manifestiert sich die Bedeutung salienzkonfigurierender Verfahren zur Relevanzmarkierung von Objekten in den AR-Dyaden. Besonders auffällig ist hierbei die Verteilung jener Annotationskategorien, die objektmanipulative Handlungen ‚in der Luft‘ beschreiben – also Handlungen, in denen Objekte in äquivalenter oder ähnlicher Formgestalt zum dargestellten Verfahren ‚Anheben und Vorzeigen‘ manipuliert werden. Hier sticht vor allem die Verteilung der Kategorie ‚Hand am Objekt in der Luft in starrer Haltung‘ hervor (rot markiert, Annotationskategorie: ^@), in der die Interaktanten der AR-Bedingung faktisch 4,5 mal so viel Zeit auf objektmanipulative Handlungen dieser Art aufwenden wie die Interaktanten der F2F-Bedingung. In einem normalisierten Vergleich entspricht dies einer Steigerung um den Faktor 5,5. In den numerischen Häufigkeiten bestätigt sich einerseits die empirische Auffälligkeit der unterschiedlichen Verteilung (AR=141 Annotationen; F2F=42 Annotationen), andererseits ist abzuleiten, dass dieses Objektpräsentations- und -identifikationsverfahren im kontrastiven Vergleich der Bedingungen in den AR-Dyaden in einer zeitlich expandierten Form auftritt (durchschnittliche Ausführungszeit: AR=2,97s; F2F=2,19s).

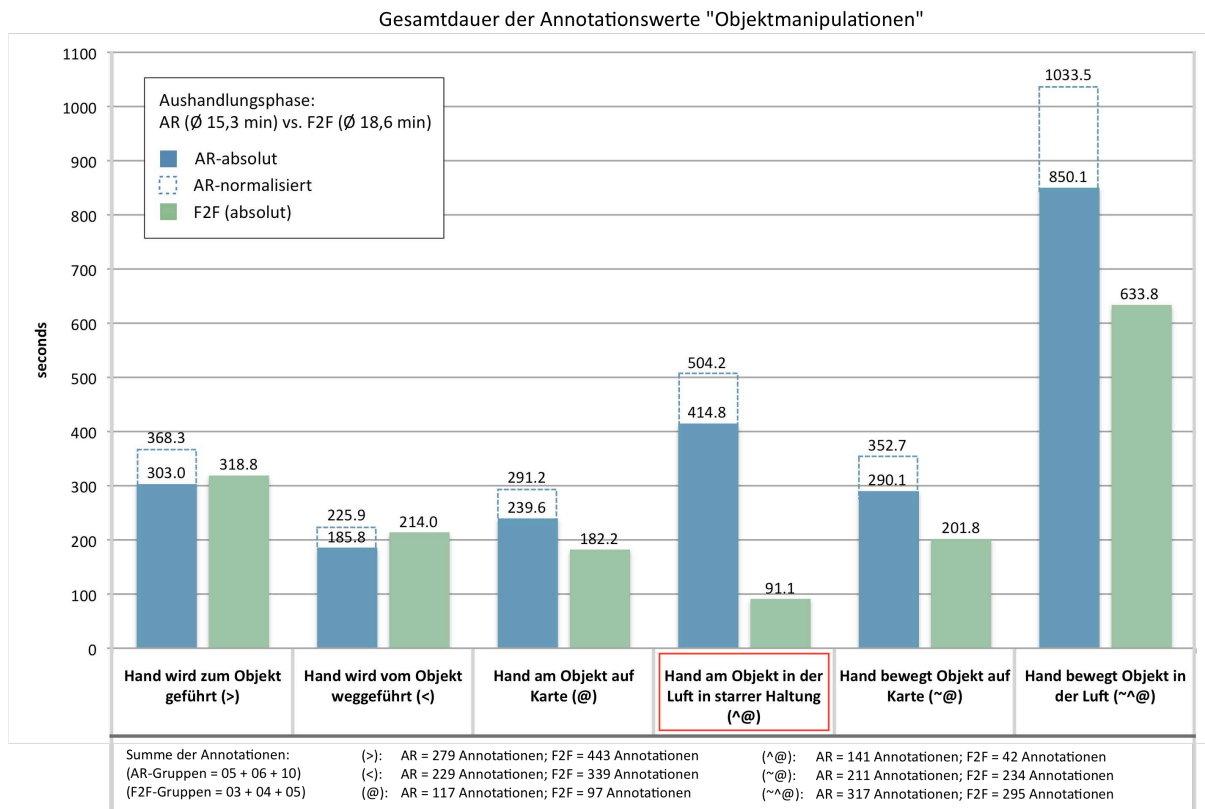


Abbildung 19: Objektmanipulative Handlungen im kontrastiven Vergleich der AR- und F2F- Bedingung. Dargestellt sind die summierten Annotationszeitwerte der jeweils drei vollannotierten Versuchspersonengruppen (VP-Gruppen AR: 05+06+10; VP-Gruppen F2F: 03+04+05).

Betrachten wir die Auffälligkeiten in der Distribution objektmanipulativer Handlungen im Verhältnis zur deiktischen Gestenverteilung (vgl. 3.2.3, Abbildung 16) lassen sich unterschiedliche Relevanzsysteme und interaktionale Bearbeitungspräferenzen erkennen: Während in der F2F-Bedingung Objektvorschläge multireferentiell in argumentative Strukturen eingebunden und in komplexen Relationen zwischen Objekten und Entitäten der Karte entfaltet werden, lässt sich unter dem eingeschränkten *mutual monitoring* in der AR-Bedingung eine Präferenz zur objektzentrierten Bearbeitung durch objektmanipulative Handlungen feststellen, in der die interaktive Herstellung eines Referenzobjekts den koordinierungsintensiven Kern der Aufgabenbearbeitung darstellt. Insofern führt die Restriktion des *mutual monitoring* zu einer Verschiebung der Relevanzsysteme: Referentielle Verfahren der Objektmanipulation sind in der AR-Bedingung nicht nur Mittel zum Zweck, sondern konstituieren einen Selbstzweck, der sich in der gemeinsamen Koordinierung über Objekte als “centers of coordination” (vgl. Suchman 1997) manifestiert.

3.4.3.2 Trajektorielle Modifikationen

Die feingraduelle Koordinierung über Objekte als “centers of coordination” (vgl. Suchman 1997) wird im folgenden Verfahren besonders deutlich, das wir terminologisch mit dem Begriff der „trajektoriellen Modifikation“ fassen. Das Beispiel zeigt, wie ‚unbelebte‘ Objekte durch manipulative Veränderungen bedeutungskonstitutiv ‚belebt‘ und mit bedeutungsdistinktiven Funktionen wie der Darstellung akuter und suspensierter Relevanz belegt werden. Objekte können so als Mediatoren eingesetzt werden, die nicht wahrnehmbare Relevanzmarkierungen des Blickes und der Körperpositur durch objektmanipulative Handlungen in etablierten Raumausschnitten substituieren.

Das Fragment zeigt den ersten Objektvorschlag dieser Versuchspersonengruppe. Wir steigen in die Analyse der Interaktion ein, nachdem VP-B (Investor) im “pre-proposal account” (Houtkoop-Steenstra 1990) seine Position erörtert, man brauche insbesondere Attraktionen, die Spaß bringen und es ihm nicht so wichtig sei, die Grünflächen zu erhalten. Am Ende dieser einleitenden Positionierung wählt B mit den

Worten „ich hab HIER zum beispiel schon was SCHÖnes“ das Objekt “WaterPark” (WP) aus und führt es vom Kartenrand in den unmittelbaren Wahrnehmungs- und Handlungsraum zur Kartenmitte. Währenddessen startet VP-A eine Gegenargumentation, in der er darauf hinweist, dass auch die „älteren und naturbegeisterten Leute“ bedacht werden müssen. Das Ende dieser Gegenargumentation markiert den Anfang unserer Analyse im nachfolgenden Fragment. Beide Interaktanten sind zu diesem Zeitpunkt bereits auf das potentiell relevante Objekt “WP” ko-orientiert. Es ist sozusagen zeitlich und lokal in einer ‚Karenzposition‘ situiert, in der die vorschlaggebende Partei B einen geeigneten Zeitpunkt zur sequentiellen Einbettung abwartet.

FRAGMENT 29:

AR-Baseline:	VP-Gruppe 06 (VP-A: Naturschützer / VP-B: Investor)
Zeit:	11:32 - 11:43

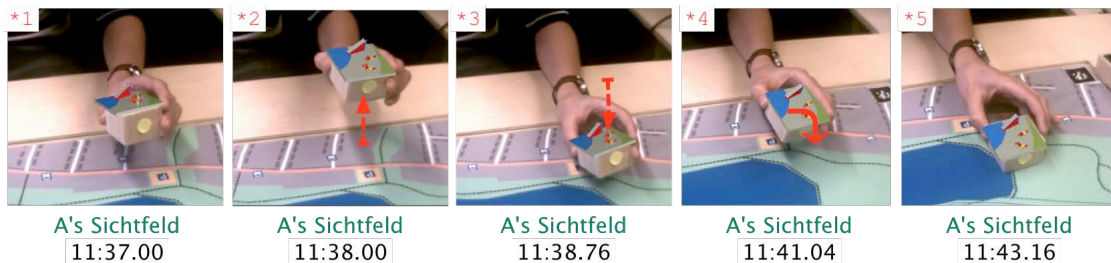
```

01 A-ver: |und ich denke mal es gibt geNU:G grünflächen die wir SCHÜTZEN
    B-act: |<<lift(WP)...<hold.....><up.....

02 A-ver: können und trotzdem noch genug platz um: (-)|°hh |bestimmte
    B-ver: |also was |ha=halten
    B-act: .....>|<hold.....>|<up.....>
    B-gaz: |((schaut sich WP genauer an.....))|
                                           *1 *2

03 A-ver: unterhaltungs|möglichkeiten zu bieten; |
    B-ver: |<<all>wir siezen uns aber ne,> |((lacht))
    B-act: <down.....>|<hold.....>
                                           *3

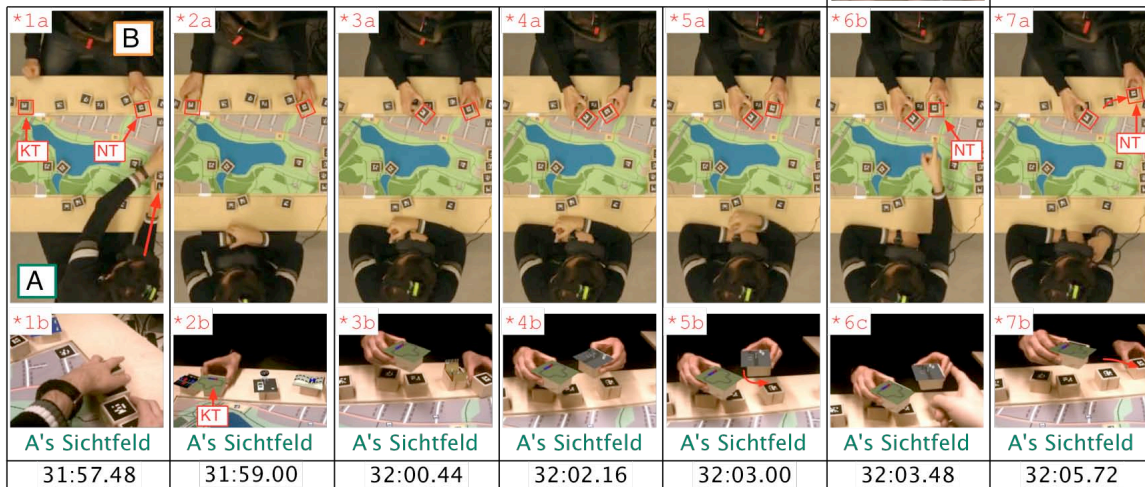
04 B-ver: WAS halten sie=denn von DIEser:=äh möglichkeit |
    B-act: ((leichtes Bewegen/Kippen von WP.....))>|place(WP)@map
                                           *4 *5
  
```



Der Beginn des Fragments zeigt die konnektierte zweite Proposition der gegenargumentativen Darstellung A's (01+02), die B's Fortführung der initial vorbereiteten Vorschlagskonstitution („ich hab HIER zum beispiel schon was SCHÖnes“) zeitweilig unterbindet. Gleichzeitig entbindet A's Gegenargumentation seinen Ko-Partizipanten B von akuten Zugzwängen, so dass dieser die Zeit ‚nutzt‘, um das Objekt “WP” genauer zu inspizieren (02). Erst nachdem A eine Gesprächs- und Atempause in der Finalsatzkonstruktion „noch genug platz um: (-) °hh [...]“ zulässt (02), reagiert B mit einem erneuten Versuch, das Objekt “WP” als Aushandlungsobjekt zu etablieren: In einer multimodalen Gestalt aus [abrupter Salienzmarkierung durch Hochziehen des Objekts] + [verbalsprachlicher Instanzierung] setzt er das Objekt in zeitlich präziser Koordinierung zur emergenten Gesprächs- und Atempause relevant (vgl. *1 und *2).

Seine verbalsprachliche Initiierung „also was ha=halten“ wird hierbei jedoch durch A's verzögerte Fortsetzung der „um...zu“-Konstruktion im verbalsprachlichen overlap unterbunden. Darauf reagierend nimmt B seinen Redrechtsanspruch zurück und suspendiert die Relevanzmarkierung des Objektes durch ein abruptes Absenken zurück in die ‚Karenzposition‘ (*3). Erst nachdem A seine konzessiv ausgelegte Gegenargumentation „[...] genug platz um: (-) °hh bestimmte unterhaltungsmöglichkeiten zu bieten;“ beendet, greift B seinen Handlungsanspruch mit „WAS halten sie=denn von DIEser:=äh möglichkeit“ (04)

04 B-ver: |<<p>weiß nich;> |
 B-act: |((stellt NT weg))|
 *7a+b



Den Ausgangspunkt der Analyse bildet eine Situation, in der VP-B mit der linken Hand das Objekt "NatureTrail" (NT) und mit der rechten das Objekt "KidsTrain" (KT) greift (*1a), während er ein Problem in der Differenzierung zwischen den Objekten verbalisiert („<<p> mir is grad nich so klar wo der unterschied HIER is;>“ (01)). VP-A ist zu diesem Zeitpunkt auf einen anderen Kartenbereich orientiert, so dass er die Auswahl der Objekte nicht wahrnehmen kann (*1b). Infolgedessen reagiert er auf das verbalisierte Differenzierungsproblem mit einer Nachfrage zur Verständnissicherung („zwischen was,“ (02)). In der Bearbeitung der Nachfrage rekonfiguriert VP-B die disperse Anordnung der Objekte (*2a+b), indem er zuerst das rechthändig manipulierte Objekt "KT" (*3a+b) gefolgt vom linkshändig manipultierten Objekt "NT" in einem zentrierten Wahrnehmungsraum zusammenführt (*4a+b). VP-A ratifiziert diese Handlung mit „achso;“ (02) und zeigt damit an, dass er nun ein fundiertes Verständnis des initial verbalisierten Differenzierungsproblems erlangt hat.

Bemerkenswert ist im Folgenden die interaktive Zuschreibung eines spezifischen Objektidentifikationsproblems, aus dem sich das Differenzierungsproblem zwischen den beiden Objekten begründet: Während B das rechthändig manipulierte Objekt "KT" in einer fixierten Positionierung hält, beginnt er das in der linken Hand geführte Objekt "NT" salienzkonfigurierend zu drehen (*5a+b). Darauf reagiert A, indem er auf das salient gemachte Objekt zeigt und in der Nachfrage „was soll DAS sein,“ (03) Probleme in der Objektidentifikation zum Ausdruck bringt. Dabei ist die Verweisgeste auf das spezifische Objekt "NT" in B's Sichtfeld nicht zu erkennen (*6a-c). Trotzdem erscheint ihr referentieller Bezug auf das Objekt "NT" für B eindeutig zu sein. In Reaktion auf A's Nachfrage antwortet B mit „<<p>weiß nich;>“ und stellt das nicht identifizierte Objekt "NT" zur Seite (*7a+b).

An diesem Beispiel wird erkennbar, wie kohärente Abfolgen sequentieller Aktivitäten bedeutungsstiftend einen interpretativen Rahmen erzeugen, der die jeweiligen Anschlussaktivitäten in ihrem ‚intendierten‘ Ausdruck inferentiell erschließbar macht.

3.5 Analytische Kontrastierung zum Exhibition-Design-Szenario

(Schnier et al. 2011a,b)

Im kontrastiven Vergleich der F2F- und AR-Bedingung im Obersee-Szenario konnte gezeigt werden, dass die Einschränkung des *mutual monitoring* eine Verschiebung der Relevanzsysteme bedingt: Referentielle Verfahren der Objektmanipulation sind in der AR-Bedingung nicht nur Mittel zum Zweck, sondern konstituieren einen Selbstzweck, der sich in der gemeinsamen Koordinierung über Objekte als "centers of coordination" (vgl. Suchman 1997) manifestiert. Die Koordinierungsrelevanz der materiellen Raumstruktur äußert sich in der empirischen Tendenz von einer multireferentiellen Bearbeitungsstruktur in der F2F-

Bedingung zu einer objektzentrierten Bearbeitung in der AR-Bedingung. Die Komplexität der materiellen Raumstruktur wird gewissermaßen durch die Präferenz der Eingrenzung auf situativ bearbeitungsrelevante Referenzentitäten reduziert. Als handlungsleitendes Prinzip konnte die Orientierung am Progressionsanspruch festgestellt werden, in dem die Interaktanten der “wait & see”-Strategie folgend darauf ‘vertrauen’, dass sich vielleicht nicht präferierte, jedoch anschlussfähige Reaktionen ergeben. Vor diesem Hintergrund stellt sich die Frage, wie Interaktionsteilnehmer verfahren, wenn die koordinierungsrelevante Raumstruktur nicht mehr stabil ist? Welche interaktionalen Effekte sind in einer Interaktionsumgebung beobachtbar, in der die kooperative Gestaltung einer interaktionalen Fortsetzung nicht ohne weiteres auf die Verfügbarkeit der materiellen Raumstruktur zurückgreifen kann?

In einer analytischen Kontrastierung zum Exhibition-Design-Szenario (vgl. Abbildung 20) – als ein anfänglicher Prototyp im Entwicklungsprozess des ARbInI Systems (vgl. Korpusstruktur, Abbildung 4) – manifestiert sich die Koordinierungsrelevanz einer stabilen Interaktionsumgebung. Im direkten Vergleich zum Obersee-Szenario weist diese Versuchsreihe eine technisch bedingte Instabilität in der Augmentierung der Objekte auf. Die Interaktionsumgebung konstituiert eine herausfordernde “ecology of sign systems” (e.g. Goodwin 2003a: 35ff.), in der semiotische Umgebungsressourcen nicht kontinuierlich verfügbar sind und somit keine stabilen Bezugspunkte der Raumorientierung bilden.

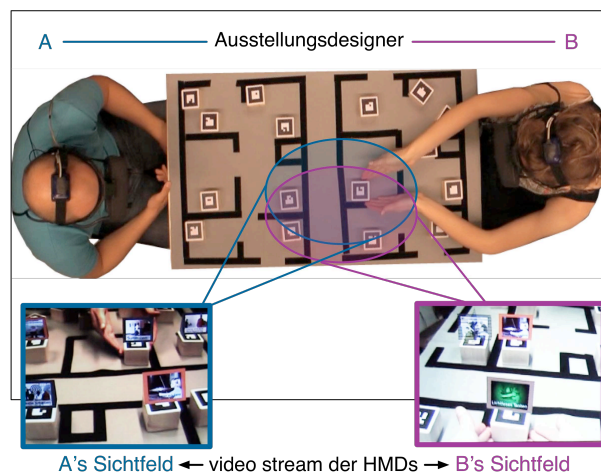


Abbildung 20: AR-System im Gebrauch mit dem Exhibition-Design-Szenario (2010)

Es ist darauf hinzuweisen, dass die Vergleichbarkeit der Szenarien in einigen Punkten leicht beeinträchtigt ist: Aufgrund unterschiedlicher Ausrichtungen im Forschungsvorhaben weist die Konzeption des Exhibition-Design-Szenarios (Dierker 2012) zwei typgleiche Grundrisspläne auf. In der hier analytisch relevanten dritten Phase des Experiments waren die Interaktanten angehalten, eine gemeinsame Lösung der 16 zur Verfügung gestellten Ausstellungsobjekte in einem der beiden Grundrisspläne eines Museums auszuhandeln (vgl. Dierker 2012: 87). Anders als im Obersee-Szenario liegt folglich für einen der beiden Interaktionsteilnehmer eine räumlich leichtere Zugänglichkeit vor. Ein weiterer Unterschied besteht in der Repräsentation der Mediationsobjekte: Im Exhibition-Design-Szenario stellen sie zweidimensionale Abbildungen dar, während im Rahmen der konzeptionellen und technischen Weiterentwicklung des Obersee-Szenario auf dreidimensionale Repräsentationen zurückgegriffen wurde.

Für die folgende analytische Kontrastierung sind diese konzeptionellen Abweichungen zum Obersee-Szenario vertretbar, da es uns nicht um quantitativ messbare Vergleichsparameter geht (z.B. Effizienz, zeitliche Dauer einzelner Objektaushandlungen, etc.), sondern um strukturelle Unterschiede, in denen partikuläre Verfahren der referentiellen Praxis zwischen einer stabilen und instabilen Interaktionsumgebung fokussiert werden.

3.5.1 Interaktionsbedingungen im Exhibition-Design-Szenario

Um den Kernunterschied der Interaktionsbedingungen zwischen dem Obersee- um dem Exhibition-Design-Szenario zu verdeutlichen, starten wir mit einem Beispiel, an dem sich die Vitalität der Referenz-

objekte im Exhibition-Design empirisch darlegen lässt (vgl. Abbildung 18). Die Interaktionsbedingung verweist auf eine „dual ecology“⁶⁴ aus Stabilität und Instabilität, in der die physisch manipulierbaren Objektträger stabile, die augmentierten Inhaltsobjekte jedoch instabile Referenzentitäten konstituieren.



Abbildung 21: Instabilität der Referenzobjekte im Exhibition-Design-Szenario

In den jeweiligen Sichtfeldern der Akteure ist die Instabilität der Referenzobjekte zu beobachten: Während B ein Objekt anhebt und in der visuellen Augmentierung das Exponat „Plasmascheibe“ erkennen kann, ist für A diese Augmentierung zuerst nicht sichtbar (#1), erscheint nachfolgend orthogonal versetzt (#2) und dann als Exponat „Dreieck im Haus“ (#3). Zu diesem Zeitpunkt werden den Interaktanten folglich unterschiedliche Exponate angezeigt, obgleich sie auf den identischen Objektträger der realen Umgebungsdimension ko-orientiert sind. Erst nachdem B das Objekt an A übergibt, wird diesem das Exponat „Plasmascheibe“ angezeigt, während es nun für B in orthogonal versetzter Augmentierung erscheint (#4).

Infolge der volatilen Interaktionsbedingungen im Exhibition-Design-Szenario können die Interaktanten nicht sicher sein, identische Repräsentationen ihrer interaktionalen Umwelt zu teilen. Neben interaktionalen Problemen in der Referenzherstellung hat dies vor allem gesprächsstrukturelle Konsequenzen, in denen sich der zeitliche Erkenntnisprozess der Teilnehmer widerspiegelt, nicht zwangsläufig über die gleichen Repräsentationen zu verfügen: So lässt sich für einzelne Versuchspersonengruppen beobachten, dass die initiale Verwendung usuelier face-to-face Verfahren (z.B. deiktische Gesten) aufgegeben wird, neue Verfahren erprobt und zu funktionalen Bearbeitungsrouninen weiterentwickelt werden. Im kontrastiven Vergleich zum Obersee-Szenario ist dabei ein massiver Relevanzanstieg objektzentrierter Orientierungs- und Bearbeitungsstrategien zu beobachten, in dem die koordinativen Anforderungen der Referenzherstellung die Bearbeitungskapazitäten der Interaktanten ‚auslastend und umfanglich‘ binden.

3.5.2 Aushandlung neuer Orientierungsroutinen

Abbildung 22 zeigt die chronologische Entwicklung referentieller Verfahren auf Objekte in der kooperativen dritten Experimentphase der Versuchspersonengruppe 08 über einen Zeitraum von 16 Minuten. Die interaktive Herstellung einer suffizienten Orientierungs- und Bearbeitungsrounne lässt sich in drei Phasen ausdifferenzieren. Als habituelles und funktionales Verfahren in haptisch zugänglichen Verweissräumen ist in der ersten Phase eine prädominante Verwendung deiktischer Gesten beobachtbar (rote Kennzeichnung „Phase-1“). Äquivalent zum Obersee-Szenario ist auch hier festzustellen, dass unter der *default*-Annahme von Ko-Orientierung agiert wird, da die feingraduelle Koordinierung der Teilnehmeraktivitäten unter den Einschränkungen im *mutual monitoring* inhibiert ist. Unterscheidend ist jedoch anzumerken, dass die Verwendung deiktischer Gesten im Exhibition-Design ein evident höheres Störpotential birgt, das sich vergleichsweise oft in deutlichen Problemen der Referenzherstellung ausprägt. Diese interaktionalen Probleme liegen dabei in einer Konstellation begründet, in der die Instabilität des Referenzobjektes nicht für beide Interaktionspartner gleichermaßen zu erkennen ist. Während die verweisgebende Partei aus einer Situation agiert, in der ihr das Referenzobjekt visuell zugänglich ist, schlägt die Markerdetektion aufgrund

⁶⁴ Der Terminus „dual ecology“ stammt von Kuzuoka et al. (2004) und wurde für unsere Zwecke redefiniert: Kuzuoka et al. nutzen die Bezeichnung für ein „remote collaboration system“, um auf die Differenz zwischen einer lokal-verfügbaren und mediatisiert-vermittelten Interaktionsumgebung hinzuweisen.

der räumlichen Distanz zum Referenzobjekt beim Rezipienten der Orientierungshandlung zumeist fehlt (vgl. Abbildung 21). Deiktische Verweisgesten gründen somit auf der Annahme, dass das Referenzobjekt dem Ko-Partizipanten gleichermaßen zugänglich ist wie dem Verweisenden selbst.

Erst in der beiderseitigen Erfahrung des Produzenten- und Rezipientenstatus deiktischer Verweisgesten entwickelt sich sukzessiv eine Art ‚interaktionale Erkenntnis‘, in der die Partizipanten die Insuffizienz des Verfahrens aus den jeweiligen Anschlussreaktionen („next turn proof“) erschließen und mit einer Instabilität der materiellen Raumstruktur in Verbindung bringen. In dieser interaktionalen Phase, die den Übergang zur Konstitution einer neuen Orientierungs- und Bearbeitungsroutine markiert, ist zu beobachten, wie deiktische Gesten kompensatorisch mit verbalen Explizitheitshinweisen versehen und objekt-manipulative Verfahren erprobt werden (auf der Zeitachse durch den gestrichelten Übergang der Phasen markiert).

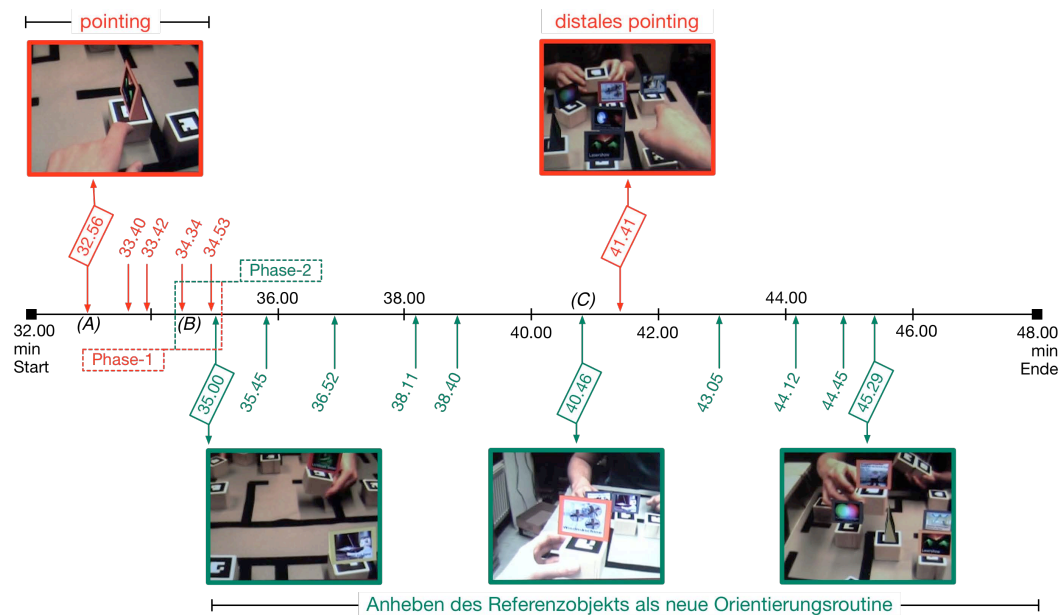


Abbildung 22: Chronologie referentieller Verfahren auf Objekte; 3. Experimentphase (dyadic phase), VP-Gr. 08

Im Rahmen der Erprobung geeignet erscheinender Referenzverfahren erhöht ein Interaktant erstmals nach drei Minuten (35:00min) den Figur-Grund-Kontrast durch Anheben, Drehen und Neigen eines Referenzobjektes. In der sequentiellen Nachfolge (markiert als Phase-2) wird dieses Verfahren leicht modifiziert und stabilisiert: Die Referenzobjekte werden in salienten trajektoriiellen Gestalten in das Sichtfeld des Ko-Partizipanten geführt, der mit einem Vorlehnen zum Referenzobjekt reagiert. Aus der beiderseitig wahrnehmbaren Suffizienz etablieren die Interaktanten das Verfahren als konstante Orientierungs- und Bearbeitungsroutine, die das Instabilitätsproblem der Referenzentitäten adressiert und überwindet.

In den nachfolgenden Fragmenten kann dieser Entwicklungsprozess empirisch veranschaulicht werden. Da wir im Rahmen dieser Arbeit nicht den gesamten interaktionalen Verlauf analytisch darlegen können, beschreiben die drei ausgewählten Fragmente exemplarisch und stellvertretend einzelne Phänomeninstanzen der drei relevanten Entwicklungsphasen. Ihre zeitliche Verortung ist in Abbildung 22 mit den Buchstaben (A)-(C) gekennzeichnet.

Ko-Orientierung als praktisches Problem (A)

Das nachfolgende Beispiel zeigt, mit welchen referentiellen Verfahren die Interaktanten im initialen Stadium der kooperativen Experimentphase operieren und welche praktischen Probleme sich dabei in der Koordinierung einer gemeinsamen Orientierung manifestieren. Dargestellt ist die allererste Referenzherstellung auf ein konkretes Objekt, die von VP-A im Kontext einer einleitenden Erklärung zur Anordnung seiner Exponate im Grundrissplan initiiert wird.

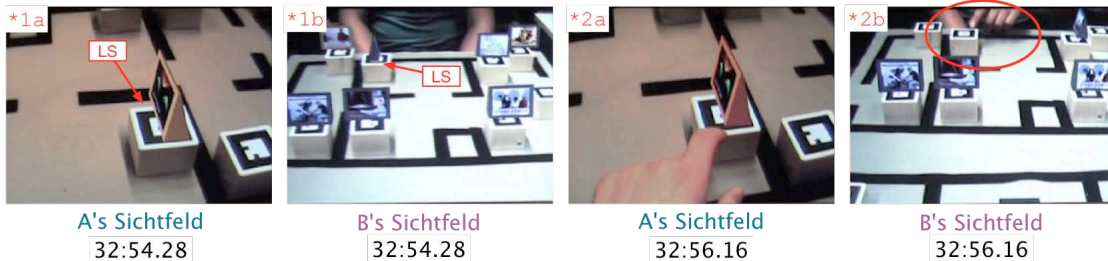
FRAGMENT 32:

AR-Disturbed:	VP-Gruppe 08 (VP-A: Ausstellungsdesigner / VP-B: Ausstellungsdesignerin)
Zeit:	32:53 - 33:08

01 A-ver: äh=also diese LASERshow hier (-)|und die |(-) siehste hier, (--)|
 A-act: |dG>>>>>>>>>@LS...>>>HP|
 B-ver: |hm:=hm:-|

*1a+b

*2a+b



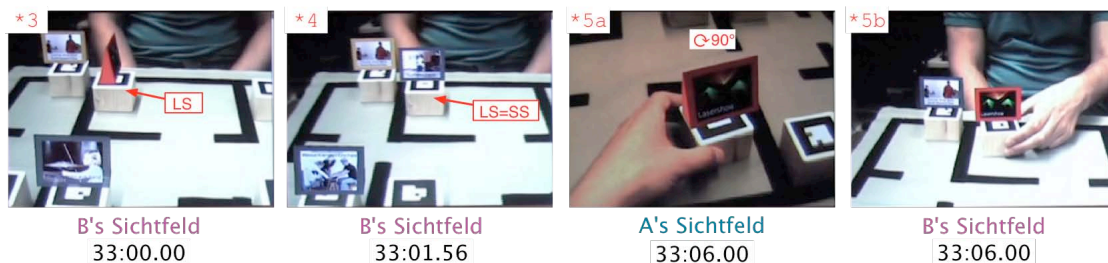
Zu Beginn des Fragments versucht A seine Ko-Partizipantin B durch die exklusiv verbalsprachliche Referenzierung „äh=also diese LASERshow hier“ (01) auf das Objekt „Lasershow“ (LS) zu orientieren (*1a+b). Der verbaldeiktische Verweis erscheint B jedoch keine präzise Orientierung auf das Objekt zu ermöglichen. In ihrem Sichtfeld ist zu beobachten, dass die Ausrichtung des Objekts eine erkennende Wahrnehmung verhindert und keine Orientierungsverlagerung einsetzt (*1b). Auf den verbalsprachlichen Orientierungsversuch reagierend antwortet sie mit einem prosodisch gedehnten „mh:=mh:-“, das B als eine Form des behaupteten, jedoch nicht fundierten Verständnisses interpretiert. Infolgedessen re-organisiert er seinen Orientierungsvorschlag: Im zweiten Anlauf ergänzt er den Orientierungsversuch um eine deiktische Geste (*2a) und adressiert die Verweishandlung in expliziter Frageform „siehste hier,“, die eine Ratifikation oder Problemschilderung konditionell relevant setzt. In B's Sichtfeld ist erkennbar, dass sie der Verweishandlung auf den materiellen Objektträger folgen kann, die Objektaugmentierung zu diesem Zeitpunkt jedoch aussetzt (*2b).

02 A-ver: |in diesem RAUM hab ich den extra reingestellt- |mh=weil
 B-ver: |warte- (0.5)
 B-act: |((lehnt sich in Richtung des Objekts))

*3

03 B-ver: irgendwie is das das UMge-(.)dreht ich seh da nen SUMMstein;
 *4

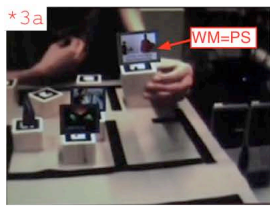
04 A-act: ((dreht Objekt um 90°))|
 B-ver: | OH- LASERshow; (.) ja=ja okay lasershow;
 *5a+b



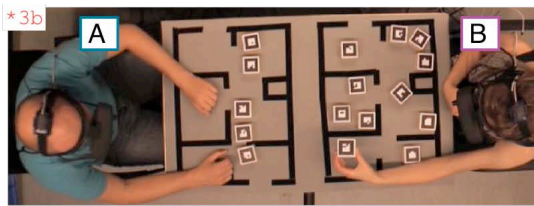
Sequentiell nachfolgend reagiert B auf das bestehende Orientierungsproblem und die akuten Zugzwänge des zweiten Orientierungsversuchs (01) durch ein Vorlehnen zum Objektträger, gefolgt von der Aufforderung „warte-“ zur einstweiligen Unterbrechung der Fortsetzungsaktivitäten ihres Ko-Partizipanten (02). In der distalen Verkürzung zum Referenzobjekt erscheint ihr die virtuelle Objektaugmentierung in

Zu Beginn des Fragments eröffnet VP-B in einer multimodalen Gestalt aus verbalsprachlicher Handlungsprojektion „und DANN ist noch die frage“ (01) und kinesischer Manipulation des Referenzobjekts „Windmaschine“ (WM) eine neue interaktionale Aufgabe. In den Sichtfeldern der Interaktanten ist erkennbar, dass B's zum Objekt „WM“ geführte Greifhand (*1b) im initialen Stadium dieses Orientierungsverfahrens für A nicht wahrnehmbar ist (*1a). In Reaktion auf B's verbalsprachliche Handlungsprojektion verlagert A jedoch umgehend – innerhalb von 600ms – seine Orientierung auf den relevanten Kartenabschnitt zu seiner Rechten (*2a), den er aus dem absenten Handlungsvollzug zu seiner Linken erschließen kann. Zu diesem Zeitpunkt erscheint ihm allerdings noch nicht das augmentierte Inhaltsobjekt „Windmaschine“ (WM), sondern das Exponat „Paukenschlag“ (PS) (vgl. *2a+b).

```
02 B-ver:   problematisch die WINDmaschine; (0.5) |
      B-act: >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>hold..... |
      A-ver:                                     |mh=mh; |
      A-act:                                     | frwd>>>WS..... |
                                     *3a+b           *4a+b
```



A's Sichtfeld
40:45.88



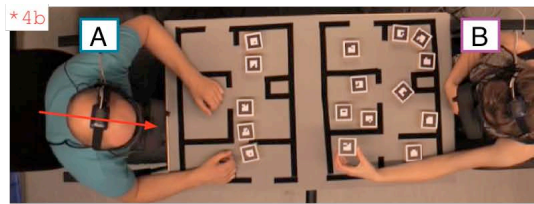
topview
40:45.88



B's Sichtfeld
40:45.88



A's Sichtfeld
40:47.40



topview
40:47.40



B's Sichtfeld
40:47.40

Im unmittelbaren Anschluss der Handlungsprojektion führt VP-B das Manipulationsobjekt mit leichten Drehbewegungen in einen räumlich verdichteten Wahrnehmungs- und Handlungsraum zu ihrem Ko-Partizipanten A (*3b+c) und referenziert es verbalsprachlich mit „also (.) ich fand (.) bei mir SUPER problematisch die WINDmaschine;“ (01+02). Zu diesem Zeitpunkt erscheint A weiterhin das falsche Inhaltsobjekt „Paukenschlag“ (PS) (*3a). Erst in der Reaktion auf B's sichtbare Darbietung, in der er sich zum Objekt vorlehnt und kooperativ den Wahrnehmungs- und Handlungsraum weiter proximal verkürzt, wird ihm das korrekte Referenzobjekt „Windmaschine“ (WM) angezeigt (*4a-c).

Die analytisch rekonstruierte Orientierungs- und Bearbeitungsroutine verweist in ihrer repetitiven Verwendung auf eine Konstitutionsform aus zwei zentralen Gestaltelementen, die in ihrem interpersonellen Zusammenwirken eine stabile Herstellung von Ko-Orientierung ermöglichen. Auf der Seite der verweisgebenden Partei wird das Referenzobjekt ostentativ präsentiert und aktiv der unmittelbaren Wahrnehmung des Rezipienten auf Sichthöhe zugeführt, wobei unterschiedliche Manipulationsbewegungen (Drehen, Neigen, etc.) die Instabilität der Augmentierung adressieren. Auf der Seite der rezipierenden Partei wird diese Orientierungsanforderung durch weitere proximale Verkürzungen des Wahrnehmungs- und Handlungsraumes sowie einem ‚strategisch vorteilhaften Blick‘ auf das Referenzobjekt kooperativ bearbeitet.

3.6 Zusammenfassung und Fazit: Robustheit von Interaktionssystemen

Analytisch motiviert wurde das vorliegende Kapitel 3 aus den empirischen Beobachtungen zur Emergenz und Auflösung konkurrierender Aktivitäten im kontrastiven Vergleich der F2F- und AR-Bedingung (vgl. 2.4). Die analytischen Kernbefunde einer *mutual monitoring* einschränkenden Interaktion auf die sequentielle Gesprächsorganisation der Interaktanten sollen nachfolgend kurz vergegenwärtigt werden:

Für die F2F-Bedingung konnte gezeigt werden, dass zwei am Tisch gegenüberliegende Interaktanten selbst bei starker Aufgabenfokussierung ihr jeweiliges Verhalten permanent direkt oder peripher („peripheral’ monitoring” (Heath et al. 1995)) beobachten und die Aktivitäten des Gegenüber in situ interpretieren. Im Sinne der Schegloff’schen Lesart „order *possible at any point*” (2005), als Korrektur zum „order at all points” Prinzip (Sacks 1984), haben wir analytisch dargelegt, dass ein kontinuierliches *mutual monitoring* die Voraussetzung bildet, emergierende Handlungsparallelität in sequentielle Ordnungsstrukturen zu überführen. Der gegenwärtige Turninhaber verfügt über ein differenziertes Repertoire an Verfahren, um Simultanaktivitäten des Ko-Partizipienten, die einen potentiellen Störfaktor zum eigenen Handlungsvollzug darstellen, instantan zu unterbinden.

Bei eingeschränktem *mutual monitoring* in der AR-Bedingung ist das Prinzip „order *possible at any point*” gestört. Ko-Orientierung stellt dabei eine *default*-Annahme der Interaktanten dar. Infolgedessen agieren sie unter der Prämisse, ihre Handlungsaktivitäten werden vom jeweiligen Gegenüber wahrgenommen. Dies hat gesprächsstrukturelle Konsequenzen: Emergierende Handlungsparallelität kann – anders als in der F2F-Bedingung – in ihrer initialen Form nicht wahrgenommen und bearbeitet werden. Den Interaktanten stehen somit keine Verfahren zur Verfügung, um Ko-Orientierung in diesem Interaktionsstadium zu reparieren und eine sich verlagernde Aufmerksamkeit des Ko-Partizipienten zurückzugewinnen.

Auch wenn diese empirischen Beobachtungen ein grundlegendes Differenzierungsparadigma der untersuchten Interaktionsbedingungen bilden, stehen sie im vermeintlichen Widerspruch zu den empirischen Beobachtungen eines funktionalen Progressionsanspruchs in der AR-Bedingung, aus dem die Interaktanten – trotz Einschränkungen im *mutual monitoring* – die Bearbeitung lokal anstehender Aufgaben koordinieren und das Handlungsgeschehen mit Blick auf die globale Aufgabenanforderung sequentiell vorantreiben. Insofern wirkt sich die episodisch-sequentielle Entkopplung in Situationen emergierender Handlungsparallelität nur marginal auf die interpersonelle Koordinierung zur Bearbeitung lokal anstehender Aufgaben aus. Störausprägungen, in denen sich teilnehmerrelevante Probleme in der Bearbeitung interaktionaler Aufgaben abzeichnen, manifestieren sich eher in singulären Randphänomenen, die jedoch nicht den konstitutiven Normalfall beschreiben. Aus diesem Grund sind wir im vorliegenden Kapitel der Frage nachgegangen, mit welchen multimodalen Verfahren die Interaktanten der AR-Bedingung den gesteigerten Koordinierungsanforderungen begegnen, wie diese dazu beitragen, Distorsions- und Störeffekte zu minimieren und welche Rückschlüsse sich daraus auf die Robustheit interaktiver Koordinierungsressourcen/-verfahren im Speziellen und interaktionaler Systeme im Allgemeinen ziehen lassen.

In einem ersten analytischen Schritt (Kapitel 3.2) ließen sich empirische Hinweise auf die zentrale Relevanz lokaldeiktischer Ausdrücke bei der Herstellung von Ko-Orientierung extrahieren. In Situationen, in denen die Teilnehmer nicht ko-orientiert sind, setzen lokaldeiktische Ausdrücke eine Orientierungsverlagerung des Ko-Partizipienten konditionell relevant. Während die lokaldeiktischen Abbildungseigenschaften für proximale („hier“) oder distale („dort“) Relationen innerhalb der definiten räumlichen Dimensionierung des Settings eher von untergeordneter Relevanz sind, bildet ihre Funktion als „attention getter” (Mondada 2012; Jakobson 1971: „shifter”) ein konstitutives Merkmal bei der Herstellung von Ko-Orientierung. Ihre Orientierungsprävalenz liegt dabei in einer doppelten Orientierungslogik begründet: Einerseits dienen sie als ‚zeitliche Orientierungsmarker’, um auf ‚Wahrnehmung als relevante Interaktionsmodalität gleichsam umzuschalten“ (Hausendorf 2010: 177). Andererseits fungieren sie als ‚verweisende Orientierungsmarker’, die nicht nur raumreferentiell auf die materielle Umgebungsstruktur, sondern auch selbstreferentiell auf weitere körperliche Ressourcen (z.B. deiktische Gesten) der holistischen Gestaltform verweisen. Im Hinblick auf die Robustheit interaktiver Koordinierungsressourcen ist dieser Aspekt zentral, da er verdeutlicht, dass in der Teilrezeption einer orientierenden Äußerungsgestalt eine holistische Konzeption unterstellt wird, die eine Aufmerksamkeitsverlagerung des Rezipienten motiviert. Die Herstellung von Ko-Orientierung mittels lokaldeiktischer Ausdrücke liegt somit nicht allein in der

initiiierenden Funktion zur Aufmerksamkeitsverlagerung begründet, sondern gleichermaßen in der Selbstwahrnehmung des Rezipienten, die Orientierungsgestalt nur fragmentarisch wahrgenommen zu haben.

Im kontrastiven Vergleich der untersuchten Bedingungen manifestieren sich sequenzstrukturell divergente Orientierungsprävalenzen (vgl. Abbildung 23): Während in der F2F-Bedingung bereits der *Onset* einer deiktischen Geste eine aufmerksamkeitsbindende und orientierungslenkende Funktion ausübt, ist in der AR-Bedingung eine Orientierungsdominanz lokaldeiktischer Ausdrücke festzustellen. Aufgrund der sequentiellen Positionierung lokaldeiktischer Ausdrücke (meist im *Stroke* der Verweisgeste oder kurz danach) wirkt sich diese Orientierungsdominanz in einer gegenüber der F2F-Bedingung zeitlich verzögerten Ko-Orientierung aus.

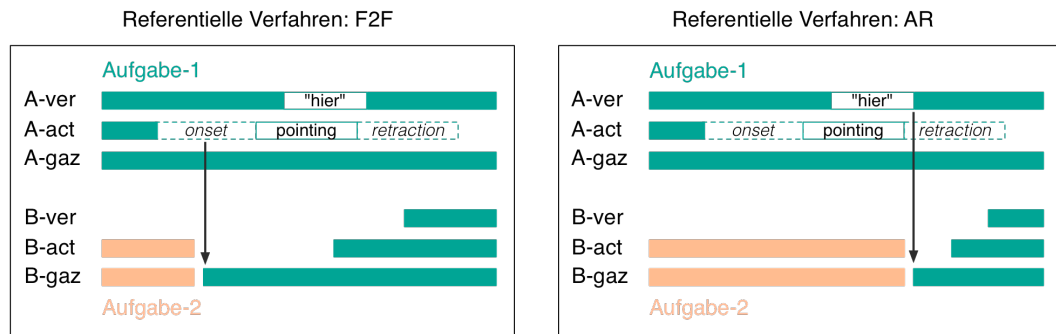


Abbildung 23: Referentielle Verfahren im kontrastiven Vergleich der F2F- und AR-Bedingung

Das (Re-)Orientierungspotential lokaldeiktischer Ausdrücke bildet jedoch nicht per se ein suffizientes Verfahren für die Herstellung von Ko-Orientierung, da die verweisgebende Partei das referentielle Verfahren aufgrund des restringierten *mutual monitoring* nicht mit der Blickorientierung des Ko-Partizipanten in situ koordinieren kann (Unterstellung von Ko-Orientierung, vgl. 2.4.2.1). Die Robustheit des Verfahrens ist vielmehr auf ein Ineinandergreifen zweier Mechanismen zurückzuführen: Während der lokaldeiktische Ausdruck eine Blickrichtungsänderung des Ko-Partizipanten konditionell relevant setzt, konstituiert die sich in der Retraktionsphase befindliche Zeigegeste des verweisenden Sprechers einen suffizienten Orientierungshinweis, anhand dessen der Ko-Partizipant unter spezifischen interaktionalen Bedingungen das Zeigeziel inferentiell ermitteln kann. Auf Grundlage dieser empirischen Beobachtungen ließen sich die folgenden Erkenntnisse gewinnen, die bestehende Untersuchungsergebnisse präzisieren und zum Teil erweitern:

- Nicht nur der *Onset* einer deiktischen Verweishandlung eröffnet einen vektoriellen Suchraum, in dem das Zeigeziel aufgefunden werden muss (vgl. Stukenbrock 2012; Goodwin 2000b), sondern auch die *Retraktion*, die es dem jeweiligen Ko-Partizipanten unter spezifischen Bedingungen (s.u.) erlaubt, den vektoriellen Suchraum retrospektiv zu erschließen und inferentiell das Zeigeziel aus dem situativen Handlungskontext zu ermitteln.
- Die methodologische, der Gestaltpsychologie entlehnte Annahme einer holistischen Produktion und Rezeption multimodaler Gestalten ("contextual configurations", Goodwin 2000a, 2003), die in ihrem Zusammenwirken deutliche Projektions- und Bedeutungsdivergenzen zu ihren Einzelkomponenten aufweisen (Goodwin 2011), kann dahingehend präzisiert werden, dass Interaktanten versuchen, die bedeutungskonstituierenden Elemente der holistischen Gestaltform zu rekonstruieren, wenn diese ihrer Wahrnehmung nicht zugänglich sind. Das zeigt, dass die holistische Rezeption nicht per se ein notwendiges Kriterium zum Verständnis einer sprecherseitigen Äußerungsgestalt darstellt, sondern durch inferentielle Verfahren kompensiert werden kann. In diesem Sinne manifestiert sich in der Rezeption eine ausgeprägte Störresistenz.
- Im Aufgabensplitting des Verfahrens aus initiiierendem Lokaldeiktikum und inferentieller Erschließung des Referenten wird analytisch transparent, wie unterschiedliche, individuelle Mechanismen ineinandergreifen können, sich aus ihrer singulären Insuffizienz gegenseitig stabilisieren und erst im interaktiven Zusammenwirken eine suffiziente Orientierung ermöglichen.

Die direkt erfolgreichen Ko-Orientierungsfälle durch lokaldeiktische Ausdrücke haben wir in 3.2.2 mit den indirekt erfolgreichen Phänomeninstanzen kontrastiert, in denen eine Notwendigkeit zur Reparatur beobachtbar wurde. Im analytischen Vergleich konnten die spezifischen Bedingungen des Verfahrens präzisiert werden (vgl. erster Punkt der Auflistung). Die interaktionale Fähigkeit, den Referenten aus einer fragmentarisch rezipierten Orientierungsgestalt zu ermitteln, ist an bestimmte ‚Salienzindikatoren‘ gekoppelt (die ausgehandelte Relevanzstruktur, der spezifische Angebotscharakter, die Beschäftigungsdauer, die räumliche Nähe zur gegenwärtigen Orientierungshandlung, etc.), die den Referenten entweder a priori (i.S.d. „Affordanz“) oder in situierten Interaktionsepisoden mit Bedeutung ‚aufladen‘ und ihm ggf. eine ‚Orientierungsrelevanz‘ zuweisen. Die vermeintlich problematischen Fälle, in denen das Zeigeziel in seiner figurativ-räumlichen Gestalt nicht eindeutig erscheint (z.B. unspezifische Raumentitäten der Karte) und eine Reparatur zur Ko-Orientierung notwendig wird, beschreiben aus der emischen Teilnehmerperspektive eher unproblematische Fälle. Eine interpersonell festgestellte Inkohärenz zwischen dem, was der Verweisgeber zu Orientierungszwecken sichtbar gemacht hat und dem, was der Rezipient davon wahrgenommen hat, kann direkt bearbeitet werden, so dass sie nicht Gefahr läuft sich in unterschiedlichen Wissensrepräsentationen eines Sachverhalts zu fossilieren.

Vor diesem Hintergrund erwiesen sich vor allem jene seltenen Fälle („deviant cases“, vgl. 3.2.4) als epistemisch besonders aufschlussreiche Phänomeninstanzen, in denen eine ausgeprägte Divergenz zwischen der projizierten und der interpretierten Orientierungsgestalt sichtbar wurde. Im methodischen Sezierungsparadigma des AR-Systems wurde empirisch beobachtbar, welche partikulären Orientierungsressourcen einer rezipierenden Partei ‚fehlen‘, um eine Orientierungsgestalt gemäß der sprecherintendierten Ausdruckskonstitution wahrzunehmen. Hierbei konnte analytisch gezeigt werden, dass divergente Interpretationen nicht allgemein darin begründet liegen, eine orientierende Gestalt nicht ganzheitlich wahrgenommen zu haben (vgl. Goodwin 2011), sondern im Speziellen darin, dass die relevanten Desambiguierungshinweise einer multimodalen Orientierungsgestalt als semiotische Interpretationsressourcen visuell nicht zur Verfügung stehen. Im Anwendungsfall der Technik dürften sich dabei vor allem solche Phänomeninstanzen problematisch auswirken, in denen die Divergenz zwischen der projizierten und rezipierten Orientierungsgestalt nicht erkannt wird, da die zur Verfügung stehenden Orientierungsressourcen – ohne profundes Wissen ihrer faktisch-funktionalen Einbettung – sinnhaft gedeutet werden (vgl. „multiple Orientierungsangebote, Fragment 22).

In einem zweiten analytischen Schritt (Kapitel 3.3) konnte die empirisch beobachtbare Robustheit auf ein zentrales Handlungsprinzip der Teilnehmerkoordinierung zurückgeführt werden. Das handlungsleitende Prinzip manifestiert sich in einem gesteigerten Toleranzverhalten der Interaktanten, in der die Präferenzorganisation innerhalb der AR-Bedingung weitaus offener gehandhabt wird als in der F2F-Bedingung: Dispräferierte Reaktionen oder ein Ausbleiben konditionell relevanter Folgehandlungen bilden keine grundlegenden Verletzungen normativer Erwartungen und werden von den Interaktanten akzeptiert. Die Teilnehmer operieren dabei mit einer „wait & see“ Strategie (s.a. Fox et al. 1996; Hayashi 1999), in der beobachtend abgewartet wird, ob eine dispräferierte Reaktion Anschlusshandlungen ermöglicht. Der Orientierung am Progressionsanspruch kommt hierbei eine zentrale Relevanz zu: Solange ein sinnhaftes Fortschreiten interaktionaler Aufgabenbearbeitungen beiderseitig erkennbar ist, besteht kein akuter Handlungsbedarf, etwaige Dispräferenzen in den Anschlussreaktionen zu bearbeiten. In diesem Sinne stellen die Teilnehmer interaktiv ‚Normalität‘ her, die sich zwar von der normativen Bezugsgröße der Präferenzorganisation in der F2F-Bedingung unterscheidet, jedoch gleichermaßen – in einer der Situation angepassten Form – der Gewährleistung einer Handlungskompetenz dient. Interaktionale ‚Normalität‘ kann somit als ausgehandeltes Resultat einer dualen Konzeption verstanden werden, in der normative Maßstäbe einer sedimentierten Handlungspraxis alltäglicher face-to-face Interaktion auf kontextsensitive Herstellungsmechanismen situativer Normerwartungen treffen.

In einem dritten analytischen Schritt (Kapitel 3.3) ließ sich das Anforderungsprofil einer interaktional robusten Aufgabenkoordinierung und -bearbeitungen weiter präzisieren. Im Zuge des Kapitels wurden einzelne Verfahren analytisch rekonstruiert, die in ihrer Gesamtheit auf ein vielfältiges, semiotisch differentes und multimodales Repertoire an Orientierungsressourcen verweisen, derer sich die Interaktanten in situativ-adaptierter Kombinatorik bedienen. Mit Verweis auf die Überblicksdarstellung zu Beginn des Kapitels werden wir nachfolgend nicht die Gesamtheit der Verfahren resümieren, sondern einerseits

spezifische Verfahren herausstellen⁶⁵, die im Rahmen des Untersuchungsvorhabens zentrale Erkenntnisse hervorgebracht haben und andererseits eine verfahrensübergreifende Einordnung zur Robustheit interaktionaler Systeme in einem abschließenden Fazit skizzieren.

Für den reibungslosen Übergang interaktionaler Aufgaben, in dem eine beiderseitige Wahrnehmbarkeit der Beendigung einer gegenwärtigen und Eröffnung einer neuen Interaktionsaufgabe sichergestellt werden muss, ließen sich empirische Hinweise auf die Orientierungsrelevanz verbalsprachlicher Koordinierungsverfahren gewinnen. Durch die wahrnehmbare Gestaltung interaktionaler Einheiten übernehmen Diskursmarker, verbalsprachliche Aufgabenratifikationen und -eröffnungen eine interaktionsstrukturierende Funktion. Sie bilden handlungsschematische Wegweiser der Orientierung und strukturieren die komplexen Koordinierungsanforderungen der interaktionalen Fortsetzung. In der analytisch rekonstruierten Bearbeitungslogik interaktionaler Teilaufgaben im sequentiellen Übergang auf ein nächstes "joint project" (Clark 1996) konnte gezeigt werden, dass sich die strukturierende Qualität der Diskursmarker in bidirektionalen Orientierungsfunktionen manifestiert (vgl. Abbildung 24).

Konstitutives Element dieser Übergangskoordination bildet die prospektive Orientierungsfunktion, die in unterschiedlichen Phänomenvariationen auftritt und von einzelnen Partikeln und Lexemen (mh::/ °hhh/ gut;/ okay;/ so::/ etc.) bis hin zu komplexen Kollokationen sowie syntagmatischen und phraseologischen Ausdrücken reicht (ja hier also-/ ich glaub wir müssen uns nen=bisschen beeilen/ dann hätten wir jetzt noch nen PA:AR sachen;/ ect.). Die prospektive Orientierungsfunktion interaktionsstrukturierender Markierungen geht dabei über eine rein kataphorische Referenz hinaus, da ihr nicht nur eine kohäsive, sondern vor allem eine projizierende Qualität zukommt, die sich aus einem funktionalen Zusammenwirken zweier Mechanismen begründet: Einerseits dienen diskursstrukturierende Ausdrücke der Sicherung eines gemeinsamen Bearbeitungsstatus (gut;/ okay;/ das haben=wir schon mal so;) und signalisieren damit den Einstieg in sequentielle Fortsetzungsoptionen. Andererseits dienen sie der konkretisierenden Eröffnung spezifischer Fortsetzungsoptionen (so::/ ja hier also-). Das Zusammenwirken dieser Funktionen ermöglicht den Interaktanten, sich innerhalb der handlungsschematischen Emergenz interaktional anstehenden Teilaufgaben bis zur Etablierung eines nächsten "joint projects" zu orientieren. Entscheidend ist hierbei, dass die Interaktanten ihre Handlungsaktivitäten auf den gegenwärtigen Bearbeitungsstatus abstimmen können, um den koordinierten Aufgabenbeginn gezielt einzuleiten (B-ver: dann hätten wir jetzt noch nen PA:AR sachen; (2.1) also: (1.3) – A-ver: ja hier also-).

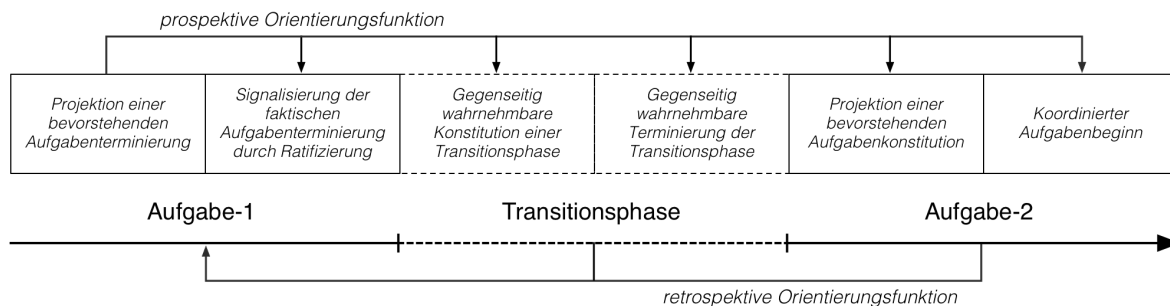


Abbildung 24: Grundordnung der Koordinierung im Übergang interaktionaler Aufgaben

Gegenüber der prospektiven Orientierungsfunktion ließen sich auch retrospektive Strukturmarkierungen aufzeigen (ja wobei::/ zumal/ na=ja::), die sich in einer Re-Orientierung des Ko-Partizipienten auf die gegenwärtige Aufgabe auswirken, sofern dieser seine Orientierung verlagert hat. Sie konstituieren optionale Strukturelemente, die im Fall eines ratifizierten Aufgabenabschlusses auf eine fortgesetzte oder expandierte Relevanzzuschreibung der vermeintlich abgeschlossenen Interaktionsaufgabe verweisen. In

⁶⁵ Auch hier ist anzumerken, dass die analysierten Verfahren in multimodalen Gestalten okkurieren, im Rahmen des AR-Sezierungsparadigmas jedoch in ihrer isolierten – aus der multimodalen Gestalt gelösten – Orientierungs- und Koordinierungsfunktion untersucht werden konnten, da den Rezipienten zu unterschiedlichen Zeitpunkten je unterschiedliche semiotische Einzelinformationen aus der Gesamtgestalt zur Verfügung stehen. Insofern ermöglicht das AR-System einzelne Ausdrucksmodalitäten der holistischen Gestaltform mit den entsprechenden Reaktionen des jeweiligen Ko-Partizipienten in Verbindung zu bringen.

der analytischen Rekonstruktion unterschiedlicher Fallalternationen des Verfahrens konnte am Beispiel des diskursstrukturierenden Partikels „wobei“ gezeigt werden, dass eine retrospektive Orientierungsfunktion nicht per se inskribiert ist, sondern sich erst in der Spezifik der sequentiellen Einbettung einer projizierten Aufgabenbeendigung entfaltet. Die alleinige ‚Abschlussratifikation‘ konstituiert somit kein hinreichendes Kriterium der Aufgabenbeendigung, was sich darin äußert, dass nach abschließenden Ratifikationen nicht selten expandierende Handlungen vermeintlich abgeschlossener Aufgabenbearbeitungen initiiert werden (z.B. durch Konzessivformen). Für die unterschiedlichen Teilnehmerinterpretationen (Aufgaben-terminierung vs. Aufgabenexpansion) konnten zwei Ursachenfaktoren herauskristallisiert werden: Einerseits ließ sich feststellen, dass eine abschließende Aufgabenratifikation in der unmittelbaren sequentiellen Nachfolge ohne weiteres, durch retrospektive Strukturmarkierungen, ‚annulliert‘ werden kann (ja wobei:, etc.). Andererseits wird aus vorzeitigen Orientierungsverlagerungen erkennbar, dass Ratifikationen von den Interaktanten nicht präzise einer Aufgabenbeendigung zugeordnet werden können (Wurde ein abschließender Vorschlag, ein Bearbeitungsstatus oder die Aufgabenbeendigung ratifiziert? Wurde vielleicht als Reaktion auf eine Ratifikation ratifiziert?). In diesem interaktionalen Entscheidungsproblem manifestiert sich die Orientierungsrelevanz struktureller Bearbeitungsabfolgen von Teilaufgaben, die nicht als isolierte, sondern nur als relationierte Einheiten einen Bearbeitungsstand sichern, der von beiden Parteien gleichermaßen interpretiert werden kann: Erst aus der Projektion eines nachfolgenden Bearbeitungsschrittes bestätigt sich den Teilnehmern die faktische Beendigung des vorigen Bearbeitungsschrittes.

Fazit: Robustheit interaktionaler Systeme im kontrastiven Vergleich der Bedingungen

Die empirisch beobachtbare Robustheit interaktionaler Verfahren, mit der die Interaktanten die Bearbeitung lokal anstehender Aufgaben koordinieren und das Handlungsgeschehen vorantreiben, ist auf ein Zusammenwirken unterschiedlicher Faktoren zurückzuführen. Die Aufmerksamkeitsmanipulation des Ko-Partizipanten und die Herstellung von Ko-Orientierung auf relevante Raumentitäten konstituieren Kernaufgaben der sozialen Interaktion. Soziale Handlungen sind auf ihre Interpretierbarkeit ausgelegt und werden unter dieser Anforderung kontinuierlich mit Orientierungshinweisen auf gegenwärtige Handlungsaktivitäten versehen. Unsere Analyseergebnisse zeigen, dass die Herstellung von Ko-Orientierung auf eine spezifische Entität im Kontext der gegenwärtigen Interaktionsaufgabe angekündigt, vorbereitet und etabliert wird. Ko-Orientierung bildet kein ad hoc Phänomen, sondern wird graduell, in sequentiell-simultanen Verlaufsformen mittels semiotisch differenter Orientierungshinweise auf allen verfügbaren Modalitätsebenen hervorgebracht. Relevant erscheint insbesondere, dass sich diese Orientierungshinweise nicht nur gegenseitig elaborieren⁶⁶, sondern vor allem auch in ihrem wechselseitigen Zusammenwirken stabilisieren. Das heißt, dass sich die semiotischen Hinweise auf den einzelnen Modalitätsebenen zwar zu einer holistischen Gestaltform zusammenschließen („gegenseitige Elaboration“), jedoch gleichermaßen selbstreferentiell auf ihre semiotischen Nachbarn innerhalb der Gestaltform verweisen („wechselseitige Stabilisation“), so dass partikuläre Ressourcen und Konstituenten der Ausdrucksgestalt absent sein können, da sie inferentiell erschließbar werden. Aus der vollumfänglichen ‚Ausnutzung‘ sequentieller und simultaner Gestaltdimensionen resultiert eine Stabilität, die es Rezipienten auch unter limitierten Zugangsbedingungen erlaubt, Orientierungshandlungen nachzuvollziehen. Insofern stellt interaktionale Robustheit immer auch das Produkt einer zeitlichen und räumlichen Entfaltung dar, die der Spezifik der interaktionalen Bedingung angepasst wird.

Wenn wir auf die analytischen Ergebnisse der F2F-Bedingung zurückblicken, ließ sich in unterschiedlichen Phänomeninstanzen eine massive multimodale Verdichtung von Orientierungshinweisen beobachten (z.B. Fragment 4), in denen die sequentielle Entfaltung der Ko-Orientierung auf ein Minimum beschränkt ist. Eine solche Verdichtung scheint durchaus unter Bedingungen möglich, in denen das kinesische Ausdrucksarsenal in seinem simultan-semiotischen Signalisierungspotential in vollem Umfang genutzt werden kann. Betrachten wir demgegenüber das analytisch rekonstruierte Handlungsprinzip der „wait & see“-Strategie (vgl. 3.3) und die empirischen Hinweise auf eine ‚entschleunigte Kommunikation‘ in der AR-Bedingung („slowing-down-communication“, vgl. 3.2.3), wird erkennbar, dass die sequentielle Entfaltung

⁶⁶ “[...] actions are both assembled and understood through a process in which different kinds of sign phenomena instantiated in diverse media, what I call semiotic fields, are juxtaposed in a way that enables them to mutually elaborate each other.“ (Goodwin 2009: 21; s.a. Goodwin 2010, 2013)

von Orientierungsverfahren auch Möglichkeiten bereitstellt, das System stabil zu halten. Denn gerade in der sukzessiven Herstellung von Ko-Orientierung kann auch die fragmentarische Rezeption einzelner Orientierungshinweise ein kohärentes Ganzes erzeugen, während bei zunehmender multimodaler Verdichtung eine strikte Aufmerksamkeitsbindung auf die verweisgebende Person vorausgesetzt werden muss. Eine ‚entschleunigte Kommunikation‘ erscheint somit im Vergleich zur F2F-Bedingung defizitär, konstituiert jedoch mitunter eine interaktionale Notwendigkeit, unter den eingeschränkten Bedingungen des *mutual monitoring* agieren zu können.

Ein zentraler Faktor der Robustheit interaktionaler Systeme liegt jedoch auch in dem interaktionalen Potential der Neuausrichtung von Relevanzsystemen und Bearbeitungspräferenzen begründet. Im Verlauf des Kapitels konnte an unterschiedlichen Phänomeninstanzen die Bedeutung objektmanipulativer Verfahren als stabilisierende Mechanismen der Orientierung aufgezeigt werden. Im kontrastiven Vergleich der untersuchten Bedingungen (F2F – AR-Obersee – AR-Exhibition) ließen sich empirische Hinweise auf eine Verschiebung der Relevanzsysteme gewinnen: Während in der F2F-Bedingung Objektvorschläge multireferentiell in argumentative Strukturen eingebunden und in komplexen Relationen zwischen Objekten und Entitäten der Karte elaboriert werden, lässt sich unter dem eingeschränkten *mutual monitoring* in der AR-Bedingung eine Tendenz zur objektzentrierten Bearbeitung feststellen. In unterschiedlichen Salienzkonfigurationen (Anheben, Drehen, Neigen, ostentatives Platzieren, etc.) erhöhen die Interaktanten den Figur-Grund-Kontrast (vgl. 3.4.3) und erzeugen so wahrnehmbare Relevanzkategorien (akute Relevanzen, Primär- und Sekundärrelevanzen im Fall konkurrierender Orientierungsangebote durch beidhändige Objektmanipulationen). In der Kontrastierung zum Exhibition-Design-Szenario (Kapitel 3.5) wurde deutlich, dass sich die Tendenz objektzentrierter Orientierungs- und Bearbeitungsstrategien in einer ausgeprägten Präferenz manifestiert, wenn die Interaktionsumgebung instabil ist. Durch die Aushandlung einer neuen Orientierungsroutine, die auf dem Prinzip der salienzkonfigurierenden Objektmanipulation beruht, ‚versuchen‘ die Interaktanten stabile Bearbeitungsstrukturen zurückzugewinnen. Insofern zeigt ein Vergleich der verschiedenen Bedingungen, dass mit steigenden Koordinierungsanforderungen sich eine Relevanzverschiebung einstellt, in der die Koordinierung über materielle Umgebungsressourcen als „centers of coordination” (vgl. Suchman 1997) einen sukzessiv höheren Stellenwert in der Bearbeitungsabfolge lokal anstehender Interaktionsaufgaben einnimmt.

4 Ko-Orientierung in handlungsschematischen Zusammenhängen

Der Untersuchungsgegenstand ‚Ko-Orientierung‘ beschreibt kein singuläres Phänomen, sondern einen Phänomenkomplex, der die unterschiedlichen Ebenen der Interaktion umfasst: Neben lokalen Herstellungsmechanismen im Sinne des „interactive achievement“, in denen die Interaktanten eine gemeinsame Orientierung auf Raumentitäten oder ein nächstes „joint project“ koordinieren, lassen sich in sozialen Interaktionsformen Orientierungslogiken erkennen, die sich auf übergreifende Handlungszusammenhänge beziehen. Die Analyse kann somit aus „unterschiedlichen Auflösungsstufen“ erfolgen, wie es Schmitt und Deppermann (2007: 38) am Beispiel des Untersuchungsgegenstandes „Koordination“ formulieren. Während in den vorausgehenden Kapiteln ‚Ko-Orientierung‘ vor allem aus einer „mikrophänomenologischen Perspektive“ analysiert wurde, fokussieren wir nachfolgend interaktionale Orientierungslogiken im Kontext „handlungsschematischer Zusammenhänge“, in denen das Wissen um „kernaktivitätsspezifische Abläufe“ den zentralen Bezugspunkt der Analyse bildet (ebd.).

Die Analyse greift hierbei auf eine empirische Beobachtung zurück, die wir im Rahmen der analytisch dargelegten Falltriade zu den „Sequenzierungsverfahren durch Direktiva“ (vgl. 2.3.7.2) gewinnen konnten. In den drei analysierten Fällen der Versuchspersonengruppe 03 (F2F-Bedingung) ließ sich eine fallübergreifende Dynamik feststellen, in der Interaktant B die Darstellungsgelegenheiten objektrelationierter Vorschlagskonstitutionen seines Ko-Partizipanten über einen längeren Zeitraum systematisch und mit steigender Explizitheit beeinträchtigte. Diese dynamische Intensivierung der „Störungsabwehr“ (Schmitt 2005: 33) parallelemergenter Handlungsansprüche haben wir terminologisch als eine emergente Abfolge aus „konzilianter“, „konsequenter“ und „insistenter Form“ zum Erhalt des eigenen Handlungsrechts konzeptualisiert. Ihre übergreifende Dynamik konstituiert eine strukturelle Einheit, in der Interaktant B seinen Handlungsanspruch mit steigender Vehemenz durchzusetzen versucht und seinen Ko-Partizipanten immer wieder in die Rolle des ‚partizipierenden Beobachters‘ seiner Handlungsentwürfe drängt.

Wenn wir uns für die zugrundeliegenden Konstitutionsbedingungen dieser interaktionalen Dynamik interessieren, stellen wir fest, dass sie auf unterschiedlichen strukturellen Ebenen der Interaktion emergiert. Einerseits ist sie in einen lokalen Sinnzusammenhang eingebunden, in dem sich die Teilnehmer am ‚Hier und Jetzt‘ der sequentiellen Emergenz orientieren. Aus dieser Perspektive ist festzustellen, dass A’s repetitiver Handlungsanspruch die Intensivierung der „Störungsabwehr“ seines Ko-Partizipanten forciert. Andererseits ist sie jedoch auch in einen globalen Sinnzusammenhang eingebunden, der die Ebene lokaler Herstellungsprinzipien übersteigt und nicht allein aus der sequentiellen Abfolge der jeweiligen Anschlussreaktionen erklärbar ist, sondern vielmehr aus dem, was aus den einzelnen Beitragsfolgen handlungsschematisch emergiert. Unter einer solchen ‚Emergenzperspektive auf Sequenzierung‘ lassen sich systematische, sowohl reflexive als auch funktionale Bezüge auf die präkontextuelle Einbettung der empirisch beobachtbaren interaktionalen Dynamik erkennen.

“Actions exist as rich, temporally unfolding process. Individual actions emerge from, and use, a consequential past shaped through chains of prior action, providing current participants with a dense, present environment, a rich now, containing many different kinds of resources that can be selectively decomposed, reused and transformed to build a next action, a proposal for how the future will be organized.” (Goodwin 2013: 21)

In Anlehnung an Goodwins Ausführungen einer starken Abhängigkeit interaktionaler Handlungen zu ihrer präkontextuellen Einbettung („chains of prior action“) werden wir nachfolgend ein Methodenprogramm vorstellen, das es erlaubt, Umschlagssequenzen in eine neue interaktionale Dynamik auf makrostruktureller Ebene zu identifizieren. Unter einer umschlagenden interaktionalen Dynamik fassen wir dabei analytisch rekonstruierbare, eruptive und präkontextuell initiierte Veränderungen des multimodalen Beteiligungsniveaus der Interaktanten.

4.1 Diskursstrukturelle Einheiten als Forschungsgegenstand

Unsere Analyse steht in der Tradition zur Erforschung diskursstruktureller Einheiten (Labov 1972; Sacks 1974; Jefferson 1978; Wald 1976, 1978; Houtkoop & Mazeland 1985, Ford, Fox & Thompson 1996, Selting 2000; Ford 2004) und liefert einen Beitrag zum empirischen Hinweis “not all turn-taking behavior can be explained in terms of Sacks et al.’s model. There are stretches of talk where the completion points of turn-constructural units do not present themselves as transition-relevance places, i.e., as opportunities for turn-transfer” (Houtkoop & Mazeland 1985). Konversationelle Struktureinheiten (“larger projects”, “Discourse Units (DUs)”, ebd.), in denen ein „primärer Sprecher“ („Primary Speaker“, Wald 1976, 1978) ein über das turn-by-turn Prinzip hinausgehendes Beteiligungsformat projiziert (“Closed DU”⁶⁷, Houtkoop & Mazeland 1985) oder dieses Beteiligungsformat interaktiv etabliert wird (“Open DU”⁶⁸, ebd.), bilden dabei einen Untersuchungsgegenstand, der bislang nicht hinreichend erforscht ist. Ford (2004) postuliert: “[...] the drive toward abstract and discrete models for units and unit projection is potentially misleading” und verweist damit auf eine mangelnde Berücksichtigung des Kontingenzcharakters⁶⁹ interaktionaler Ereignisse bei der Erforschung von Diskurseinheiten (“[...] units are always open to riffs and spins, reshaping units as they unfold.”).

In Seltings Beitrag “Construction of Units in Conversational Talk” wird ähnlich argumentiert: “In principle, units are always flexible and expandable; hence the actual completion can be recognized only retrospectively” (2000: 511). Ihr empirisch fundierter Verweis, größere Diskurseinheiten nur retrospektiv erschließen zu können, verdeutlicht, dass jene Einheiten als teilnehmerrelevante Kategorie weniger bedeutsam sind als bei Houtkoop & Mazeland (1985) angenommen⁷⁰. Komplexe Turnkonstruktionseinheiten (“compound TCUs”, “big packages”, “larger projects”, ebd.: 481) bilden infolgedessen für sie ein Epiphänomen spezifischer Aktivitätstypen:

“Speaker and recipients in general do not orient to the production of TCUs as such, but rather to the organization of interpretable activities that are constituted with and via such units. The production of units is only an epiphenomenon of the production of activities, so it is not surprising that participants do not show a manifest verbal orientation to each single TCU of larger turns that constitute activities.” (Selting 2000: 511)

Der analytischen Konzeption in sich geschlossener, begrenzter Diskurseinheiten (Houtkoop & Mazeland 1985) stellt Selting den Begriff “far-reaching projection” als teilnehmerrelevante Kategorie gegenüber und trägt damit der prinzipiellen Offenheit projizierter Aktivitätseinheiten Rechnung. Unsere empirischen Befunde aus 2.3.7 und 2.3.8 liefern hier einen Beitrag zum Verständnis der Kontingenzcharakteristik interaktionaler Phänomene: Das primäre Handlungsprinzip Interagierender scheint nicht in der Herstellung geordneter Strukturen zu liegen (Geordnetheit als „Epiphänomen“), sondern darin, ob die jeweiligen Handlungsaktivitäten der Teilnehmer Anschlussmöglichkeiten generieren oder der Bearbeitung von Fortsetzungsproblemen dienlich sind. Kontingenz bildet dabei eine Grundvoraussetzung zur sequentiellen Fortsetzung interaktionaler Aktivitäten.

⁶⁷ Eine mögliche Projektionsform für eine geschlossene Diskurseinheit (Closed DU) ist in der Konzeption des “story entry device” erfasst (Jefferson 1978). Durch einleitende Diskursmarker (z.B. “I once saw...”) (Houtkoop & Mazeland 1985: 598) etabliert sich ein Akteur als „primärer Sprecher“, um sich das Rederecht für eine längere Struktureinheit zu sichern. Der Diskursmarker dient damit dem jeweiligen Ko-Partizipanten als Hinweis, dass in den nachfolgenden Turnkonstruktionseinheiten (TCUs) innerhalb der Narration das Rederecht nicht neu verhandelt werden muss.

⁶⁸ In Abgrenzung zu geschlossenen Diskurseinheiten stellen Houtkoop & Mazeland (1985) für offene Diskurseinheiten (Open DUs) am Beispiel der Kategorie „Ratschläge“ (“advice”) fest: “Continuation of previous speaker is not in the first place projected by the speaker him-/herself, as is the case in the production of Closed DU’s, but it is an interactionally managed construction, due to a specific type of recipient reaction in a typical sequential position. Recipient reactions as, e.g. continuers, have another quality when they are produced – or omitted – on the kind of sequential positions described for proper nexts of advices. In the course of the production of closed DUs they signal that Primary Speaker can go on with his project, whereas in the post-advice constellation they negotiate primary speakership of advice-giver.”

⁶⁹ Kontingenz bezeichnet ein Prinzip der Offenheit und Unbestimmtheit. Lokale und globale Zugzwänge setzen zwar bestimmte Folgehandlungen konditionell relevant, legen sie jedoch nicht deterministisch fest. Interaktanten haben immer die Möglichkeit anders zu reagieren oder sich der konditionell relevant gesetzten, präferierten Folgehandlung zu entziehen (Bsp. Frage – Gegenfrage). Schegloff (1996b: 22) betont: “There is virtually nothing in talk-in-interaction which can get done unilaterally, and virtually nothing which is thoroughly pre-scripted.”

⁷⁰ Houtkoop & Mazeland (1985: 597): “We believe that, in addition to turn-units, there are larger units of talk of a different level of discourse and that conversationalists display sensitivity to such units in their turn-taking behavior.”

In der Literatur zur Erforschung diskursstruktureller Einheiten fällt auf, dass der Fokus auf einer Präzisierung der bislang nicht hinlänglich definierten Konzeption von “Turn-Constructional Units” (TCUs) und ihrer Relation zu “Transition Relevance Places” (TRPs) liegt. Während das Modell des turn-taking nach Sacks et al. (1974) für einzelne “single-unit turns” auf syntaktischer Ebene⁷¹ (“single-word”, “single-phrase”, “single-clause constructions”, Sacks et. al. 1974: 702) eine hohe empirische und analytische Evidenz bietet, stellt es für “multi-unit turns” keine hinreichenden Beschreibungskategorien zur Verfügung. Ursächlich dafür ist die empirische Beobachtung: “In single-unit turns, a TCU always ends in a TRP. In multi-unit turns, however, there are both non-final and final TCUs within the turn” (Selting 2000: 490). Infolgedessen ist ein übergeordnetes Forschungsinteresse an multimodalen Projektionsformaten zur Ankündigung und Aufrechterhaltung einfacher wie komplexer Struktureinheiten feststellbar.

Ausgehend von der historisch tradierten Fokussierung auf syntaktische Faktoren (Sacks et al. 1974) haben nachfolgende Arbeiten dabei einerseits zu einer analytischen Präzisierung syntaktischer Einheitenbildung beigetragen (Schegloff 1996a; Auer 2005), andererseits weitere Beteiligungsressourcen in den Blick genommen (Prosodie: Sacks 1992; Auer 1996; Couper-Kuhlen 1996, 2001, 2004 – Gestik: Schegloff 1984; Streeck & Hartge 1992 – Blick: Goodwin 1980, 1981; Lerner 2003; Rossano 2005) und multimodale Konstitutionsformen aufgezeigt (z.B. Ford, Fox & Thompson 1996; Goodwin 2000; Mondada 2007b). Die analytischen Ergebnisse dieser Forschung deuten auf multimodale Projektionsgestalten, in denen unterschiedliche verbale und nonverbale Ressourcen bei der Ankündigung und Aufrechterhaltung einfacher wie auch komplexer Struktureinheiten beteiligt sind. Die syntaktische Kategorie der Turnkonstruktionseinheit bildet folglich kein „Alleinstellungsmerkmal“ für die Konstitution und Rezeption komplexer Struktureinheiten.

4.2 Ableitung des Forschungsdesiderates

Während ein Querschnitt durch die einschlägige Literatur auf umfangreiche und empirisch fundierte Analyseergebnisse zu multimodalen Projektionsformen konversationeller Struktureinheiten (sowohl “single-unit turns” als auch “multi-unit turns”) deutet, liegen nur wenige Erkenntnisse zur Emergenz und präkontextuellen Verankerung jener Einheiten vor. Dabei stellen wir fest, dass Arbeiten zur strukturellen Einheitenbildung den “multi-unit turn” als maximale Größe des Phänomensbereichs fokussiert haben, der kontextuellen Dynamik aufeinanderfolgender “multi-unit turns” ko-präsentierender Akteure bislang jedoch kaum in systematischer Weise Rechnung getragen wurde. Erste Hinweise finden sich bei Sacks (1974) und Ryave (1978), die auf Grundlage verbalsprachlicher Daten empirische Tendenzen für die interpersonelle Reihung von “joke”- (Sacks) und “story” -Sequenzen (Ryave) aufgezeigt haben (s.a. Houtkoop & Mazeland 1985). Die erste “joke”- oder “story”-Einheit fungiert dabei als präkontextuelle Konfiguration einer nachfolgenden und projiziert die Möglichkeit einer Anschlusshandlung, die nicht explizit als Einheit der spezifischen Gattung “joke” oder “story” angekündigt werden muss. Unter einer multimodalen Perspektive auf Interaktion widmen wir uns den folgenden Forschungsfragen:

- Welche organisatorische und teilnehmerrelevante Funktion haben narrative Struktureinheiten in unseren Daten, in denen der Aktivitätstyp als kooperative Maximierung individueller Handlungsziele charakterisierbar ist?
- Ist die serielle Emergenz als repetitive Aneinanderreihung kategorialer Einheiten zu betrachten oder lassen sich systematische Bezüge feststellen, in denen die erste Einheit – als präkontextuelle Konfiguration – Auswirkungen auf die multimodalen Verfahren der sequentiellen Folgeinheit hat?

Ein zweites Forschungsdesiderat ergibt sich aus der Feststellung, dass komplexe Struktureinheiten vornehmlich als “larger projects” untersucht wurden, in denen ein “Primary Speaker” – unter der aktiven Mitgestaltungsmöglichkeit des Ko-Partizipanten – ein über das turn-by-turn Prinzip hinausgehendes

⁷¹ Selting weist in ihrer kritischen Reflexion zur “turn-taking machinery” (Sacks et al. 1974) darauf hin, dass die Definition von TCUs in exklusiver Weise auf den Kriterien “syntactic structure” und “projectability” beruht. Dabei stellt sie ein unterschiedliches Verständnis konversationeller Einheiten zwischen Sacks et al. und heutigen Forschungsansätzen fest: “[...]the unit Sacks and colleagues had in mind was fundamentally different from the units that other researchers, among them linguists, have generally looked for: units that display ‘self-determined, independent, recognizable completeness’” (Selting 2000: 480).

Beitragsformat etabliert. Auch wenn die Kategorie des “Primary Speaker” (Wald 1978) unter einer multimodalen Perspektive auf Interaktion durch den Hinweis auf das übergeordnete Prinzip der Teilnehmerorientierung auf „interpretierbare Aktivitäten“ präzisiert wurde (s.o.: Selting 2000: 511), legen sowohl die auf Verbalität wie auch die auf Multimodalität beruhenden Ansätze eine Differenzhypothese zwischen Sprecher und Zuhöreraktivitäten bei der Bestimmung von “multi-unit turns” zugrunde. Das heißt, die Existenz größerer Struktureinheiten ist in den untersuchten Fällen immer an eine primär handelnde Person gebunden, die eine strukturelle Expansion des Rede- und Handlungsrechts anzeigt.

Für unsere Daten stellen wir jedoch fest, dass Akteure, die sich innerhalb einer Diskurseinheit als “Primary Speaker” etablieren, nicht notwendigerweise ein prädominantes Aktivitätsniveau in den restlichen Handlungsmodalitäten zeigen. In Abhängigkeit zum spezifischen Aktivitätstyp der Interaktionsaufgabe kann eine hohe Verbalaktivität des “Primary Speaker” somit auch zu einer gesteigerten nonverbalen Aktivität des Ko-Partizipanten führen, da kontemporäre Handlungsziele und Interaktionsaufgaben über den verbalen Kanal nicht verfolgt und bearbeitet werden können. Die Bestimmung kontextueller Grenzen einer Diskurseinheit ist in solchen Fällen nicht eindeutig. Vor diesem Hintergrund stellen sich die Forschungsfragen:

- Ist die Bestimmung von Diskurseinheiten auf Basis einer primären Handlungsträgerschaft (vgl. “Primary Speaker”) sinnvoll?
- Bildet die Diskurseinheit – als klar begrenzte Kategorie mit Start- und Endpunkt – unter einer multimodalen Perspektive auf Interaktion eine analytisch sinnvolle Beschreibungsdimension?

In Reaktion auf die aufgezeigten Forschungslücken werden wir nachfolgend ein Methodenprogramm vorstellen, das nicht die Diskurseinheit ins definitorische Zentrum der Analyse rückt, sondern die Diskurs-Transformation im Sinne einer Veränderung der interaktionalen Dynamik berücksichtigt. Insofern stellen wir der primären Handlungsträgerschaft (“Primary Speaker”) die Konzeption dynamisch emergierender Beteiligungsniveaus gegenüber und fokussieren präkontextuelle Bedingungen, in deren sequentieller Nachfolge sich eruptiv Veränderungen des multimodalen Beteiligungsniveaus der Interaktanten abzeichnen. Im Zuge dessen werden wir analytisch zeigen, dass jene Umschlagsdynamiken interaktiv emergieren, im Zusammenhang diskursrelevanter Aushandlungen stehen und ihnen eine initiiierende präkontextuelle Bedingung vorausgeht, die eine Breitenwirkung auf nachfolgende Ereignisse ausübt. Diese Wirksamkeit erstreckt sich dabei nicht nur auf eine semantische Relation zwischen Präkontext und nachfolgender Handlung: “Building new action by performing operations on an existing public substrate⁷² is central to how participants grasp the meaningfulness of subsequent talk (Goodwin 2013: 9)”, sondern konstituiert sich als Abhängigkeitscluster, das auch eine Einflussnahme auf die in der Nachfolge stehenden interaktionalen Verfahren erkennen lässt.

Eine umfassende analytische Rekonstruktion interaktionaler Phänomene muss folglich unter Berücksichtigung relationierter Ereignisse des Präkontextes geschehen, die eine initiale Konfiguration für das im analytischen Fokus stehende Phänomen bilden (können):

“A speaker’s action is context-shaped in that its contribution to an on-going sequence of actions cannot adequately be understood except by reference to its context – including, especially, the immediately preceding configuration of actions – in which it participates. [...]. Since every ‘current’ action will itself form the immediate context for some ‘next’ action in a sequence, it will inevitably contribute to the framework in terms of which the next action will be understood.” (Heritage 1984: 242)

Dieser Argumentation folgend sind “contextual configurations” (Goodwin 2000a, 2003), die als holistische Einheiten in ihrem Zusammenwirken Projektions- und Bedeutungsdivergenzen zu ihren Einzelkomponenten aufweisen (Goodwin 2011), nicht ausschließlich auf die Domäne multimodaler Äußerungsgestalten zu beziehen. – Das gestaltpsychologische Prinzip *das Ganze ist mehr als die Summe seiner Teile* ist auch für Ereignisse des Präkontextes zu prüfen, die im Falle einer identifizierbaren Relation einen konstituierenden Teil des Phänomenbereichs darstellen.

⁷² “[...] I am choosing to use the term substrate to describe the local public configuration of action [...] that is operated on (frequently through processes of decomposition, reuse and indexical incorporation) to build a next action.” (Goodwin 2013: 11)

Unser Ansatz adressiert damit ein dringliches Desiderat der multimodalen Interaktionsforschung, in der die Komplexität der Phänomene oftmals eine isolierte, aus der präkontextuellen Emergenz losgelöste Phänomenbetrachtung bedingt, deren Postulate einer idealisierten Analytik gleichen. Auch wenn ein idealisierter Phänomenbereich eine analytische Rekonstruktion der situativ verwendeten multimodalen Verfahren und ihres interaktiven Zusammenwirkens beim Handlungsvollzug erlaubt, kann eine Berücksichtigung aller Initialbedingungen des Phänomens den konversationsanalytischen Forschungsfokus erweitern und Hinweise liefern, warum bestimmte Handlungen mit spezifischen Verfahren konstituiert werden. Wenn wir die Interaktionsgeschichte eines im analytischen Interesse stehenden Phänomens nach Goodwin als “proposal of how the future will be organized” (Goodwin 2013: 21) betrachten, muss eine umfassende Analyse diejenigen Initialbedingungen des Präkontextes identifizieren, in denen die Interaktanten ihre interaktionale Umwelt maßgeblich – d.h. mit potentieller Wirksamkeit auf Nachfolgehandlungen – verändert haben oder die Veränderung ihrer interaktionalen Umwelt reflexiv als richtungsweisendes Ereignis konstituieren.

4.3 Sequenzstrukturelle Auffälligkeiten der Verbalebene

4.3.1 Methodische Vorbemerkung zur Verbalannotation

Abbildung 25 zeigt die Sequenzstruktur der Verbalannotation vom Beginn bis zum Ende der dyadischen Aushandlungsphase in Versuchspersonengruppe 03 der face-to-face Bedingung. Die Annotationswerte wurden dabei von meinem Projektkollegen Alexander Neumann automatisch aus den separat aufgezeichneten Sprachsignalen der Teilnehmer generiert (Verwendung von Headsets, vgl. Abbildung 2). Das optimale Ergebnis zur Segmentierung des Sprachsignals ließ sich mit einem *threshold* von -15dB erzielen (s.a. Neumann & Hermann 2013). Da eine semantisch -prosodische Verbalannotation auf Korpusbreite für unser Forschungsziel keine dringliche Analyseressource bildet, kann für die Versuchspersonengruppe 03 keine Abweichung zwischen einer manuellen und automatischen Annotation des Sprachsignals angegeben werden. Für Versuchspersonengruppe 06 der AR-Bedingung liegt jedoch eine manuelle Annotation der Verbalebene vor, so dass wir die Genauigkeit der automatischen Sprachsegmentierung im Verhältnis zur manuellen Annotation bestimmen können:

Aushandlungsphase AR-VP06: 18:53 min (=1133 sec.)

Gesamtwert der manuellen Verbalannotation (A-ver_manus + B-ver_manus): $x_r \approx 1002,5$ sec.

Gesamtwert der automatischen Verbalannotation (A-ver_auto + B-ver_auto): $x_a \approx 1076,6$ sec.

absolute Abweichung: $F = x_a - x_r \approx 74,1$ sec.

relative Abweichung: $f = \frac{x_a - x_r}{x_r} \times 100\% \approx 7,4\%$

Die für Versuchspersonengruppe 06 der AR-Bedingung bestimmte Abweichung zeigt, dass der Gesamtwert der automatischen Verbalannotation um mehr als eine Minute (ca. 74 sec.) über dem der manuellen Annotation liegt. Der höhere Wert lässt sich dabei im Wesentlichen auf zwei Faktoren zurückführen: Zum einen auf die Erfassung nicht-sprachlicher akustischer Signale (z.B. lautstarkes Aufschlagen der Holzklötze auf die Tischplatte etc.). Zum anderen auf Episoden hoher Schallintensität, die durch das jeweilige Mikrophon des Ko-Akteurs zusätzlich erfasst werden, auch wenn dieser nicht Produzent der Äußerung ist (z.B. lautstark inszenierte Rollenkonflikte etc.). Die Abweichung verweist im Zuge dessen auf das methodische Problem, den *threshold* so zu wählen, dass möglichst auch schallarme Sprachsegmente erkannt, aber nicht-sprachliche Segmente (*noise*) herausgefiltert werden. Die relative Abweichung von 7,4% (ca. 4 sec./min) ist für eine erste makrostrukturelle Bewertung potentieller Umschlagsdynamiken jedoch ausreichend. Gleichwohl ist methodologisch darauf hinzuweisen, dass es sich nur um eine erste Sichtung episodisch auffälliger Sequenzstrukturen handeln kann, die dann im nachfolgenden Schritt am Datenmaterial überprüft werden müssen.

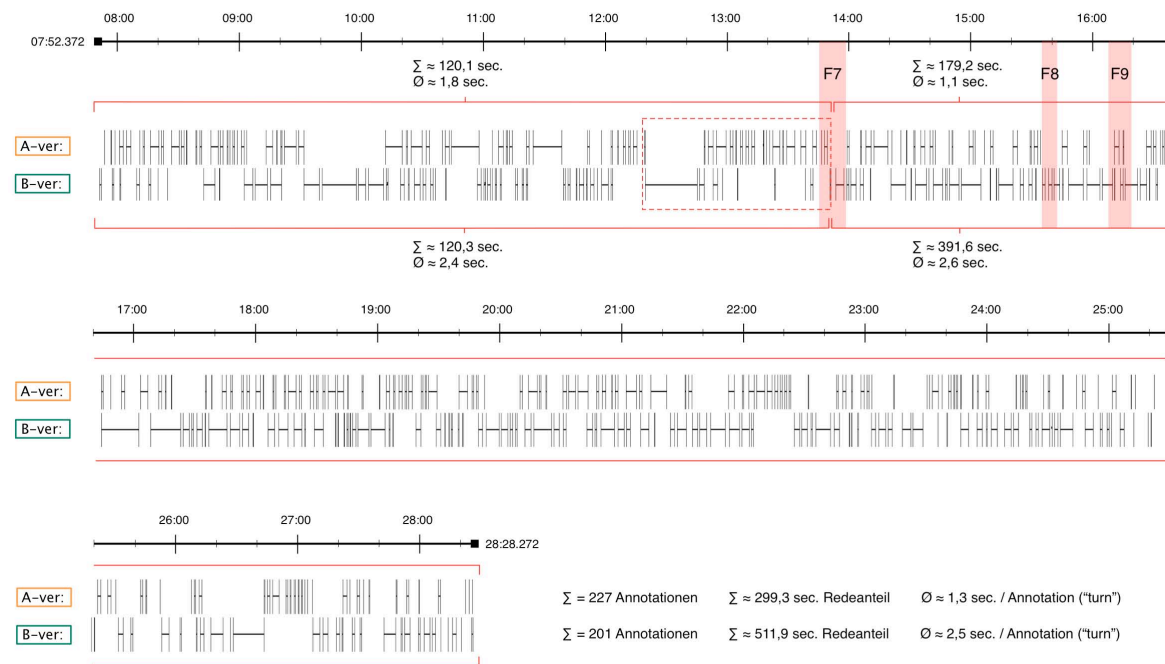


Abbildung 25: Sequenzstruktur der Verbalannotation (VP-Gr. 03, F2F)

4.3.2 Sequenzstrukturelle Auffälligkeiten und Hypothesenbildung

In der Sequenzstruktur der Verbalannotation fällt auf, dass die Rederechtsverteilung beider Versuchspersonen in den ersten sechs Minuten des Versuchs (07:52-13:50) ausgewogen ist (VP-A: 120,1 sec./ VP-B: 120,3 sec.). Beide Interaktanten haben hier einen Redeanteil von ca. 20 Sekunden pro Minute, wobei die durchschnittliche Länge der Redebeiträge zwischen den Versuchspersonen leicht variiert (VP-A: 1,8 sec./ VP-B: 2,4 sec.).

In den verbleibenden Versuchsminuten (13:50-28:28) hingegen ändert sich dieses Verhältnis: Während VP-A seinen Redeanteil auf 12,2 Sekunden pro Minute verringert, steigert VP-B diesen auf 26,8 Sekunden pro Minute. Mit einem summarischen Wert von 391,6 Sekunden ist B's Redeanteil in dieser Interaktionsphase damit mehr als doppelt so hoch wie der seines Ko-Partizipanten VP-A (179,2 Sekunden). Die grundlegende Veränderung der Rederechtsverteilung spiegelt sich auch in der durchschnittlichen Länge der Redebeiträge wider: Während A's Redebeiträge kürzer werden (1,1 Sekunden), nehmen B's Redebeiträge ab Minute 13:50 im Durchschnitt leicht zu (2,6 Sekunden).

Im Vorfeld der hier im analytischen Fokus stehenden Phänomene zur Sequentialisierung durch Direktiva (F7-F9) ist eine sequenzstrukturell auffällige Episode beobachtbar (gestrichelte Markierung), die eine Transitionsphase zwischen der ausgewogenen und der divergenten Rederechtsverteilung bildet. Ab Minute 12:20 beansprucht B das Rederecht kontinuierlich und ohne detektierbare Pausen für 25,5 Sekunden. Währenddessen ist bei A keine verbalsprachliche Aktivität zu verzeichnen. Erst nach Beendigung der monologischen Sequenz startet A eine Verbalaktivität, die sich über einen Zeitraum von 60,4 Sekunden erstreckt (12:50.000-13:50.400). In dieser Zeitspanne hat A einen absoluten Redeanteil von 26,5 Sekunden, während B's vereinzelte Sprachaktivität einen summarischen Wert von 6,3 Sekunden erzielt.

Zusammenfassend lassen sich unter einer bislang rein verbalsprachlichen Fokussierung die Beobachtungen zum Beteiligungsniveau der Interaktanten schematisch wie folgt zusammenfassen und beschreiben:

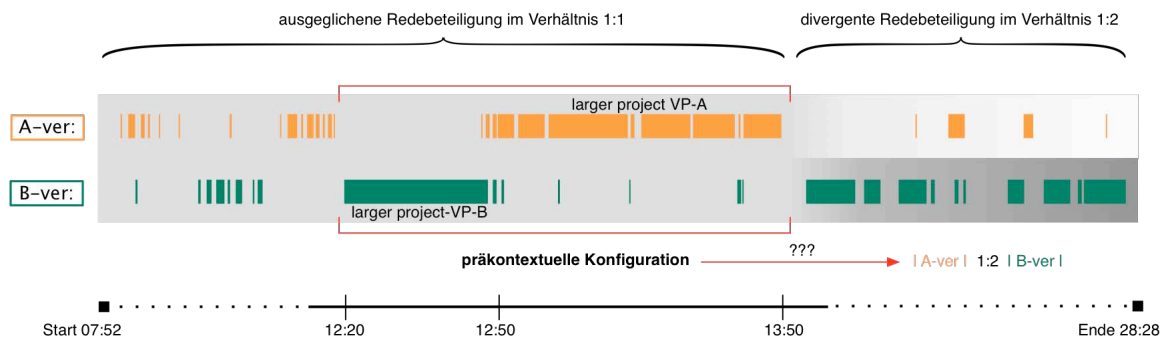


Abbildung 26: Schematisierung der sequenzstrukturellen Auffälligkeiten der Verbalebene (VP-Gr. 03, F2F)

- i. Beide Interaktanten haben in der präkontextuellen Konfiguration einen überproportional hohen Redeanteil.
- ii. Die monologische Sprachaktivität B's (25,5 sec.) wird vom Ko-Partizipienten A weder unterbrochen noch durch verbalsprachliche Aktivitäten begleitet (beispielsweise durch "back-channel-behavior" (Yngve 1970), "continuers" (Schegloff 1982), etc.).
- iii. Die Anschlussreaktion A's weist mit 26,5 Sekunden einen summarisch ähnlich langen Redeanteil zu B's voriger Sprachaktivität auf, ist jedoch über einen Zeitraum von 60,4 Sekunden verteilt.
- iv. Innerhalb des 60,4 sekundigen Zeitraums ist bei B nur eine geringfügige Sprachaktivität zu verzeichnen (6,3 sec.).
- v. Nach der sequentiellen Abfolge der größeren Struktureinheiten ändert sich das Redebeitragsverhältnis zwischen den Interaktanten maßgeblich.

Aus der sequenzstrukturellen Auffälligkeit und den Beobachtungen einer steigenden Explizitheit in den Sequentialisierungsverfahren durch Direktiva (F7-F9) leiten wir die Hypothese ab:

Die dynamische Abfolge interpersoneller Struktureinheiten bildet eine präkontextuelle Bedingung für den Umschlag in eine neue interaktionale Dynamik und übt als "proposal of how the future will be organized" (Goodwin 2013: 21) eine breite Wirksamkeit auf die in der Nachfolge stehenden Verfahren zur Sequentialisierung durch Direktiva aus.

4.4 Hypothesenverifizierung

4.4.1 Präzisierung durch Visualisierung der Redebeteiligung im Interaktionsverlauf

Sofern sich unsere Hypothese als haltbar erweist, stellt sich die Frage, wie sich der Einfluss des potentiell relevanten Präkontextes in der unmittelbaren Nachfolge ausgestaltet. Anhand der Sequenzstruktur der Verbalannotation in Abbildung 25 ist folglich nicht ableitbar, ob sich die Veränderung der Rederechtsverteilung ab Minute 13:50 in der unmittelbaren Nachfolge der sequenzstrukturell auffälligen Episode abzeichnet. Grund dafür ist, dass die für eine feingraduelle Analytik notwendige Präzision auf der Annotationsebene eine Datenfülle erzeugt, in der Veränderungen des Beteiligungsniveaus visuell kaum identifiziert werden können. Unser Ziel ist es also, feingraduelle Strukturen mittels deskriptiv statistischer Methoden in größeren Zusammenhängen zu erkennen.

Da die Bearbeitung (nicht der Bearbeitungserfolg) individueller Handlungsziele an die Aktivität der Objektmanipulation geknüpft ist, können die im Untersuchungsfokus stehenden Verfahren zur Auflösung von Handlungsparallelitäten (Fragment 7-9) durch eine analytische Erweiterung auf ihre Einbettung in die kontextuelle Dynamik sich verändernder intra- und interpersoneller Beteiligungsniveaus nicht nur präzisiert, sondern in ihrer Emergenzcharakteristik analytisch durchdrungen werden.

Um Veränderungen des Beteiligungsniveaus in der Nachfolge der als Präkontext identifizierten Konfiguration zu erschließen, betrachten wir die Redeanteile beider Interaktanten kontrastiv für Minutenintervalle des Versuchs (vgl. Abbildung 28). Die Datenpunkte im Diagramm repräsentieren folglich gebinnete Werte, in denen die zeitliche Dauer der Sprachaktivität für die einzelnen Versuchspersonen pro Minutenintervall (1Bin=60s) summiert wurde.

4.4.1.1 Methodische Vorbemerkung zum statistischen Verfahren des “Binning”

Der Fokus des statistischen Teils unserer Analyse liegt in einer adäquaten makrostrukturellen Beschreibungsform der interaktionalen Dynamik als eine sich stetig verändernde Konfiguration intra- und interpersoneller Beteiligungsniveaus auf unterschiedlichen Handlungsmodalitäten. Da in den einzelnen Handlungsmodalitäten die Annotationswerte kein numerisches Gleichgewicht bilden, müssen sie für einen kontrastiven Vergleich, der die deskriptive Darstellung sich verändernder Aktivitätsniveaus zum Ziel hat, auf eine gemeinsame Zeiteinheit bezogen werden. Die Größe des Intervalls richtet sich dabei nach dem Zweck der Darstellung. Einminütige Intervalle bilden für uns eine zweckdienliche Repräsentation, in der die Umschlagsdynamik sichtbar wird und die Versuchslänge im Diagramm räumlich darstellbar bleibt.

Unter einer multimodalen Perspektive auf Interaktion, in der ein praxeologischer Ansatz zur Rekonstruktion feingradueller, interaktiv und sequentiell hervorgebrachter Teilnehmerverfahren verfolgt wird, stellt sich die Frage nach dem faktischen Repräsentationsgehalt intra- und interpersoneller Beteiligungsniveaus für einminütige Intervalle. Einerseits stellen einminütige Intervalle eine relativ große Bezugsgröße dar, andererseits ist die makrostrukturelle Einteilung eines Interaktionsverlaufs in Intervalle arbiträr und gesprächsstrukturell nicht begründbar, da der interaktive und zeitlich lineare Entstehungsprozess keine Berücksichtigung findet. Demzufolge bildet eine makrostrukturelle Einteilung eine Repräsentationsform, die der Teilnehmerperspektive und dem Wissen der Interaktanten zum jeweiligen Zeitpunkt im Interaktionsverlauf im Sinne des sequenzanalytischen turn-by-turn Prinzips keine Rechnung trägt.

Dennoch scheint es in sozialen Interaktionssituationen auch größere, über das turn-by-turn Prinzip hinausgehende Struktureinheiten zu geben, die teilnehmerrelevante Orientierungsinstanzen bilden. Struktureinheiten wie “joke”- (Sacks 1974) oder “story” -Sequenzen (Ryave 1978), die die Fortsetzungsmöglichkeit generieren, typäquivalente Folgeeinheiten ohne “entry device” und gattungsspezifischer Markierung zu versehen, bilden dabei nur einen spezifischen Aktivitätstyp (Selting 2000), an dem die Teilnehmerorientierung empirisch beobachtbar ist. In unseren Daten lässt sich ein weiterer Typ identifizieren, der an die strukturelle Qualität der Vorschlagsunterbreitung gebunden ist (vgl. Beispiel I). Einerseits stellt die Vorschlagskonstitution oftmals selbst eine zeitlich expandierte Handlungsstruktur dar, die durch “pre- und postproposal accounts” (Houtkoop-Steenstra 1990) angekündigt und argumentativ gestützt wird, um die präferierte Anschlussreaktion der Vorschlagsakzeptanz und -ratifikation zu erzielen. Andererseits wird die Vorschlagsunterbreitung vielfach vom Ko-Partizipanten argumentativ gewendet, um die dyspräferierte Reaktion – seine Gegenargumentation – zu untermauern. Anders als für “joke”- (Sacks 1974) und “story” -Sequenzen (Ryave 1978) beschrieben, konstituiert die Gegenargumentation keine typäquivalente Folgeeinheit und erfährt infolgedessen zumeist eine pre-sequentielle Markierung („ähm:: DAzu kann ich was SA:GEN“). Dabei ist die Ausführung der Gegenargumentation, äquivalent zur Vorschlagskonstitution, an eine “account-Verdichtung” (Schmitt 2005) gekoppelt, um wiederum eine Präferenzstruktur zu erzeugen, die nun in der Einsicht zur Vorschlagsaufgabe und dem Abbruch des Handlungsvollzugs liegt:

Beispiel-I: VP-Gr. 03 (F2F): 09:23 - 09:46 + [argumentative Fortsetzung]

01 A-ver: ICH finde wir brauchen gar nicht so viel WALDfläche=ich finde eigentlich könnten wir die QUADfläche- also !HIER! is so=nen WA:LDgebiet das wird quasi GAR NICHT touristisch genutzt- da könnte man eigentlich super eine QUAD(.)rennstrecke reinbauen-

02 B-ver: ähm:: DAzu kann ich was SA:GEN und zwar kann man es nämlich WUNDERBAR touristisch nutzen- °h mit solchen ähm schönen WANDERwegen =die zerstören auch die naTUR nicht so sehr =weil die sind so nen bisschen erhöht (-) ne,=DA kann man LANGwandern OHNE das man durch die naTUR STRATZT [argumentative Fortsetzung]

In der sequentiellen Abfolge der Struktureinheiten „Vorschlagskonstitution“ und „Gegenargumentation“ orientiert sich die konterkariierende Partei somit auch an den argumentativen Teilstrukturen, die konstituierender Bestandteil der Vorschlagshandlung sind. Sie bilden nicht nur Gestaltungsmöglichkeiten der gegenargumentativen Darstellung, sondern manifestieren sich in strukturellen Zugzwängen der Bezugnahme. Das heißt, der kommunikative Erfolg der Gegenargumentation (i.S. der präferierten Anschlussreaktion) steigt und fällt mit der Reproduktion und argumentativen Wendung der vorschlagskonstituierenden Teilstrukturen. Vor diesem Hintergrund erklären sich iterative Abfolgen der Struktureinheiten „Vorschlagskonstitution“ und „Gegenargumentation“, in der die konstituierenden Teilaspekte der ersten Einheit in ihrer raum-zeitlichen Dimension in der Nachfolgeeinheit reproduziert, semantisch transformiert und durch individuelle, argumentative Teilstrukturen angereichert werden⁷³.

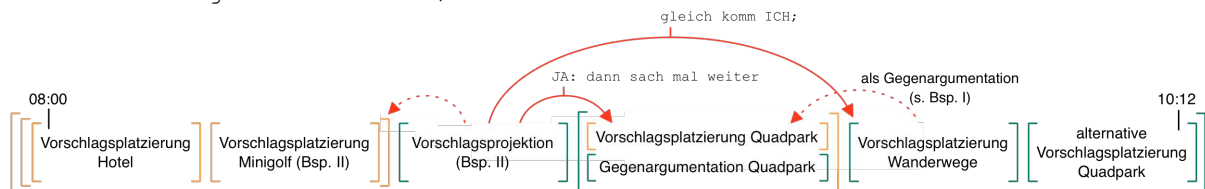
Ein anderes Format, das für unseren Aufgabentyp der kooperativ ausgehandelten Verwirklichung eigener Handlungsziele kennzeichnend ist, besteht in der Maximierung des individuellen Handlungsraums (vgl. Beispiel II). Die Interaktanten orientieren sich in unserem Szenario zum Teil auch an den „Vollzugsfenstern“, die der jeweilige Ko-Partizipant bei der Bearbeitung individueller Handlungsziele beansprucht. Dies ist vor allem dann beobachtbar, wenn die zeitliche Dimension intrapersonell vollzogener Handlungen metasprachlich thematisiert wird, um die eigene Beanspruchung des Handlungsraums in der sequentiellen Folgeeinheit anzukündigen und gleichermaßen zu legitimieren:

Beispiel-II: VP-Gr. 03 (F2F): 08:29 - 08:47

((im Vorfeld hat A bereits das Objekt "Hotel" auf der Karte platziert und schließt nachfolgend einen weiteren Vorschlag an))

01 A-ver: wir RICHTen DAFür noch EXTRA: ähm eine MINIGolfbahn ein-
=die wird dann (--) HIER so liegen- (1.8) das ist (1.8)
auch SEHR ähm:: (-) ke=KUNdenattraktiv- ne,

02 B-ver: JA: dann sach mal weiter =was hast du für VORSchläge noch =HIER,
=gleich komm ICH;



Beide Beispiele veranschaulichen einen grundlegenden strukturellen Unterschied zu den empirischen Beobachtungen der interpersonellen Emergenz von „joke“- (Sacks 1974) und „story“-Sequenzen (Ryave 1978), der gleichermaßen unser gewähltes methodisches Vorgehen begründet: Während die bei Sacks und Ryave beschriebenen Fortsetzungsmöglichkeiten sich auf Anschlussreaktionen beziehen, die aufgrund einer Typäquivalenz ohne gattungsspezifische Markierungen auskommen und in Aushandlungsprozessen der sozialen Beziehungskonstellation begründet liegen⁷⁴, ist die serielle Abfolge größerer Struktureinheiten in unseren Daten an das aufgabenspezifische Format der Vorschlagshandlung gebunden. Die Reihung

⁷³ Ein Anschauungsbeispiel des Phänomens stellen politische Debatten dar, die in der medialen Adaption des „Kanzlerduells“ dem Effekt der „Reproduktion und argumentativen Wendung der vorschlagskonstituierenden Teilstrukturen“ durch die Reglementierung der Redebeitragslänge der Kandidaten auf 60 Sekunden entgegenwirken. Der rhetorische Niederschlag (franz. *débatte*) ist somit kaum möglich, da die unterschiedlichen argumentativen Teilstrukturen in der gegenargumentativen Darstellung nur in Ansätzen aufgegriffen werden können. Die Kandidaten müssen folglich abwägen, wieviel Zeit sie für die Entkräftung fremder und Darstellung eigener Inhalte aufwenden.

⁷⁴ Es sei darauf hingewiesen, dass die interpersonelle Emergenz von „joke“ und „story“-Sequenzen anhand „authentischer Daten“ natürlichsprachlicher face-to-face Interaktionen beobachtet wurde, in denen die Interaktanten – im Gegensatz zu unseren Daten – a priori kein übergeordnetes Handlungsziel verfolgen („smalltalk“). In derartigen Situationen werden soziale Beziehungskonstellationen ausgelotet (z.B. durch die interpersonelle Abfolge von Witzen, die normative Erwartungen verletzen und je nach bestehender Beziehungskonstellation als „normverletzend“ angekündigt werden („ich kenn auch einen, der ist politisch aber nicht ganz korrekt“)).

struktureller Einheiten bildet somit innerhalb unseres Szenarios kein Randphänomen, sondern ist Konstitutionsmerkmal der Aufgabenbearbeitung.

Mit Perspektive auf die Größe der Struktureinheiten „Vorschlagskonstitution“ und „Gegenargumentation“ ist festzustellen, dass in den Einheiten unterschiedliche Teilaufgaben bearbeitet werden müssen, die je nach spezifischer kontextueller Bedingung Divergenzen in der raum-zeitlichen Beanspruchung der in den Einheiten bearbeiteten Aktivitäten bedingen. Die Größe der Struktureinheiten variiert damit stark. Einflussfaktoren stellen hier – neben intrinsisch motivationalen Aspekten – Affordanzen des spezifischen Objekts und des vorgeschlagenen Platzierungsortes der Karte⁷⁵, das Weltwissen und die Fähigkeit zur argumentativen Verwertbarkeit (Naturschutz vs. Investments) und die situative Einbettung⁷⁶ dar. Ein weiterer Faktor, der auf die raum-zeitliche Variationsbreite der Struktureinheiten wirkt, mag zudem in der Beobachtung begründet liegen, dass in gegenargumentativen Darstellungen sowohl die zentralen Argumentationsweisen der Vorschlagskonstitution aufgenommen und argumentativ gewendet als auch eigene Vorschlag-Argument-Kombinationen integriert werden.

Mit Verweis auf die über zweiminütige „Projektfolge“ in Bsp. II lassen sich folglich heterogene Struktureinheiten identifizieren, die sich in verschachtelten Binnenstrukturen über große Zeiteinheiten entfalten. Die einzelnen Strukturformate bilden dabei teilnehmerrelevante Einheiten, die über das turn-by-turn Prinzip hinausgehende Orientierungsinstanzen bilden und durch die intra- und interpersonelle Dynamik verändernder Beteiligungsniveaus auf Makrostrukturebene repräsentiert werden können. Die Veränderung multimodaler Beteiligungsniveaus über den Zeitverlauf stellt somit keine reine Auswirkung der interaktionalen Aufgabenbearbeitung dar, sondern – wie die Teilnehmerkoordinierung an Vollzugsfenstern verdeutlicht – eine orientierungsrelevante Bezugsgröße. Es handelt sich damit um ein bidirektionales Ursache-Wirkungsgefüge.

Während sich die Begründung einer makrostrukturellen Betrachtung aus dem spezifischen Aufgabenformat und der Bearbeitung durch größere Struktureinheiten wie „Vorschlagskonstitution“ und „Gegenargumentation“ erschließt, muss hervorgehoben werden, dass es sich bei unserem methodischen Ansatz um eine makrostrukturelle Repräsentationsform – eine Emergenzperspektive auf Sequenzierung – handelt, bei der die Größe des gewählten Intervalls die Granularität der Merkmalsausprägungen bestimmt (vgl. Abbildung 28). Eine makrostrukturelle Repräsentation in Intervallen ermöglicht somit nur Aussagen in Bezug auf das Intervall, während empirische Beobachtungen zu einzelnen Zeitpunkten innerhalb des Intervalls erst unter einer mikroanalytischen Präzisierung sinnvoll sind. Grund dafür sind arbiträre, gesprächsstrukturell nicht begründbare Intervallgrenzen.

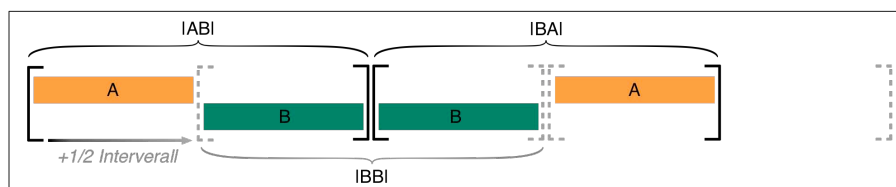


Abbildung 27: Intervallverschiebung zur Bestimmung der maximalen Abweichung

Abbildung 27 veranschaulicht den „worst case“, der in Abhängigkeit zum Startpunkt des Intervalls unterschiedliche Merkmalsausprägung der participationsstruktur in der Größenordnung eines halben Intervalls erzeugt. Es handelt sich dabei um den theoretischen Maximalwert der Abweichung. Participationsabfolgen von A-B-B-A, in denen die Interaktanten eine kontinuierliche Aktivität von je 30 Sekunden zeigen, sind für einminütige Intervalle jedoch gleichermaßen unwahrscheinlich wie faktisch nicht existent. Die tatsächliche Abweichung liegt folglich deutlich unter dem Betrag eines halben Intervalls. Demnach ist es entscheidend, die participationsstruktur als Tendenz über die Zeit zu verstehen, mikro- und makroskopische Ansätze als sich komplementär ergänzende Prinzipien zu betrachten und nicht die der spezifischen

⁷⁵ Stellt das ausgewählte Objekt oder die Kartenregion aufgrund seiner Merkmalspezifikationen eher einen „problematischen“ oder „unproblematischen“ Planungsvorschlag dar?

⁷⁶ Wurden dem jeweiligen Ko-Partizipanten gegenwärtig viele Zugeständnisse gemacht oder vermehrt Planungsvorschläge abgelehnt? Ist eine nächste Vorschlagshandlung bereits projiziert worden (vgl. Bsp. II: „gleich komm ICH;“)? Lässt sich die partikuläre Vorschlagseinheit mit integrierten Planungskonzepten relationieren und begründen?

Methodik innewohnende Beschreibungsmöglichkeiten als qualitativen Maßstab der jeweils anderen Methodik anzulegen.

Zusammenfassend nutzen wir das statistische Verfahren des Binning somit als retroaktive Rekonstruktionsform, um die interaktionale Dynamik der Beteiligungsressourcen über den Zeitverlauf analytisch zugänglich und kontrastiv vergleichbar zu machen. Binning dient damit als Vergrößerungsfilter, als eine Möglichkeit, die feingraduellen Strukturen auf Turnebene in größeren Zusammenhängen zu erkennen.

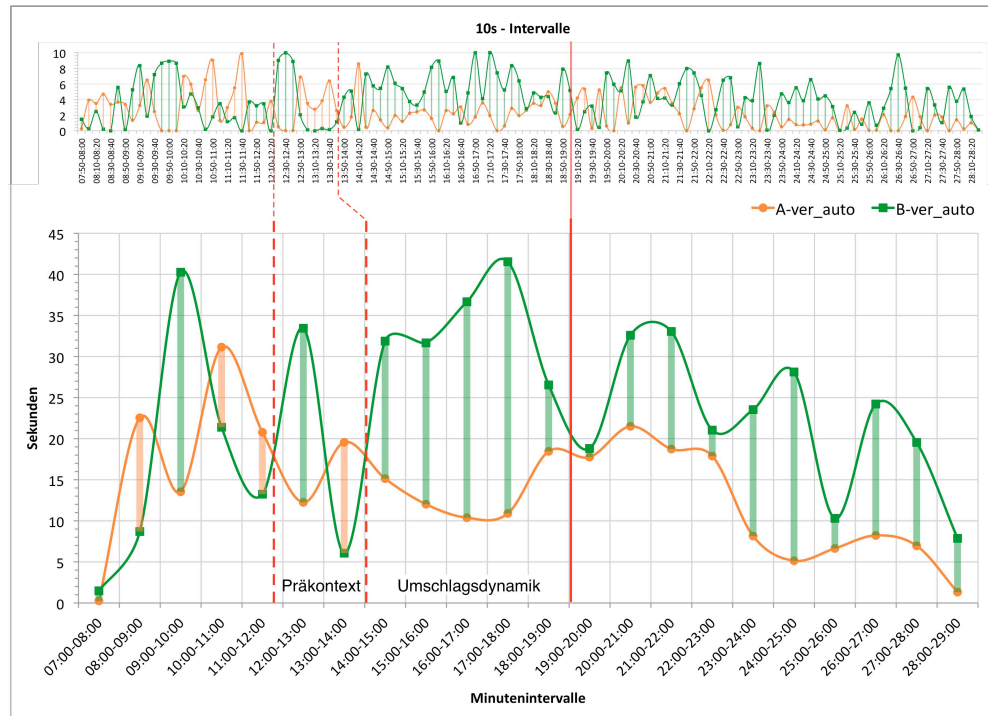


Abbildung 28: Redeanteil der Versuchspersonen pro Minutenintervall (VP-Gr. 03, F2F)

4.4.1.2 Beobachtungen zur Redebeitragsverteilung

Die ausgewogene Redevertelung in den ersten sechs Minuten des Versuchs (vgl. Abbildung 25) resultiert aus einer wechselseitigen Abfolge hoher und niedriger Verbalanteile der jeweiligen Versuchsperson in den Minutenintervallen. Der verbalsprachliche Interaktionsverlauf ähnelt in diesem Zeitintervall einer Sinusschwingung, in der die Gesprächspartner ihre summarische Redezeit pro Minute in (quasi-) harmonischen Wechseln koordinieren. Die zuvor als sequenzstrukturell auffällig identifizierte Episode (Abbildung 25), ist in Abbildung 28 durch die Minutenintervalle 12:00-13:00 und 13:00-14:00 repräsentiert (gestrichelte Markierung): Hier fällt auf, dass VP-B's hoher Redeanteil von 33,4 Sekunden (12:00-13:00) im zweiten Minutenintervall (13:00-14:00) in einen vergleichsweise niedrigen Wert umschlägt, der bezogen auf die automatische Annotation der Verbalebene „VP-B“ mit 6,0 Sekunden den Minimalwert darstellt (beim Versuchsbeginn handelt es sich nicht um ein vollständiges Minutenintervall, da die erste Verbalannotation erst bei 07:52.372 beginnt (vgl. Abbildung 25)).

Der Umschlag in eine neue interaktionale Dynamik, die auf verbalsprachlicher Ebene durch eine prädominante Redebeteiligung B's markiert ist, zeichnet sich im unmittelbaren Anschluss an die als potentiell relevanter Präkontext identifizierte Episode ab: Aus der Konfiguration einer minimalen verbalsprachlichen Beteiligung steigert B seinen Redeanteil im Minutenintervall 14:00-15:00 auf 31,9 Sekunden und hält diesen fortan auf einem Niveau, das deutlich über der Redebeteiligung seines Ko-Partizipanten liegt. Dabei fällt auf, dass insbesondere in den nachfolgenden vier Minuten (14:00-18:00) B's wachsende Redebeteiligung auf den Maximalwert von 41,5 Sekunden im Minutenintervall 17:00-18:00 zu einer persistenten Abnahme der Redebeteiligung A's führt. Der mögliche Raum, sich im sequentiellen Vollzug verbalsprachlich zu äußern wird folglich von VP-B okkupiert. Die lokale und endogene Transformation in eine neue

Ordnungsstruktur verdichtet im Zuge dessen unsere Annahme, dass die zuvor identifizierte sequenzstrukturell auffällige Episode als präkontextuelle Bedingung für den Umschlag in eine neue interaktionale Dynamik zu kennzeichnen ist und den Teilnehmern als “proposal of how the future will be organized” (Goodwin 2013: 21) dient.

4.4.2 Präzisierung durch Visualisierung weiterer Beteiligungsressourcen: Objektmanipulative Handlungen

Die Annahme einer durch eine spezifische präkontextuelle Bedingung initiierten Umschlagsdynamik manifestiert sich, wenn wir uns die objektmanipulativen Handlungen der Versuchspersonen für die einzelnen Minutenintervalle anschauen (vgl. Abbildung 29). Als objektmanipulative Handlungen wurden solche Aktivitäten erfasst, in denen die Hände der Versuchspersonen manipulierend in einen relationalen Zusammenhang zu einzelnen Objekten oder Objektgruppen gebracht wurden. Auch wenn eine zum Objekt geführte Greifhand in seiner indexikalischen Funktion einer deiktischen Verweishandlung (z.B. mit ausgestrecktem Zeigefinger) ähnelt, wurden deiktische Gesten als separate Kategorie behandelt. Wie wir später sehen werden, bilden sie in unserem Setting ein strukturell andere Beteiligungsressource, die separat betrachtet und bewertet werden muss.

Für die Darstellung im Liniendiagramm sind die einzeln annotierten Handlungsaktivitäten „Hand wird zum Objekt geführt“ (>), „Hand wird vom Objekt weggeführt“ (<), „Hand befindet sich am Objekt auf der Karte“ (@), „Hand bewegt Objekt auf der Karte“ (~@), „Hand befindet sich am Objekt in der Luft in ruhender Haltung“ (^@) und „Hand befindet sich am Objekt in der Luft in Bewegung“ (~^@) für die jeweiligen Minutenintervalle summiert worden. Da linkshändige und rechtshändige Aktivitäten der Versuchspersonen im Korpus durch einzelne *tiers* erfasst sind, wurden sie für die Darstellung verodert, um einer beidhändigen Aktivität keine höhere Gewichtung beizumessen. Infolgedessen bildet das Diagramm komparativ aktueursspezifische Aktivitäten, die im Zuge manipulativer Handlungen an Objekten ausgeführt wurden, für gebinnte Minutenintervalle des Versuchs ab.

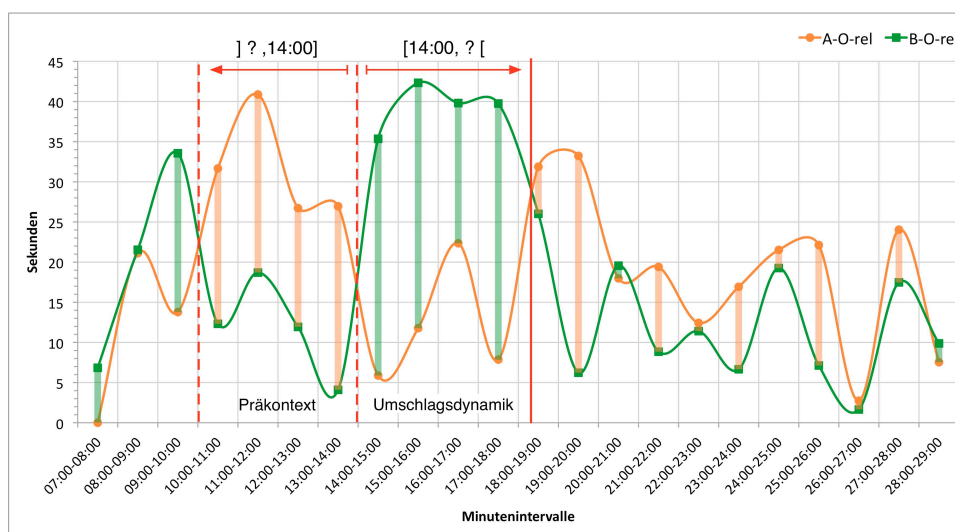


Abbildung 29: Objektmanipulative Handlungen der Versuchspersonen pro Minutenintervall (VP-Gr. 03, F2F)

Die Verteilung der objektmanipulativen Handlungen verweist auf ähnliche strukturelle Besonderheiten, die bereits zuvor als initiiierende kontextuelle Bedingung einer nachfolgenden Umschlagsdynamik identifiziert wurden. Im direkten Vergleich zur verbalsprachlichen Aktivität ist jedoch eine längere Episode von ca. vier Minuten (10:00-14:00) der umschlagenden Dynamik vorgestellt (gestrichelte Markierung), in der VP-A eine deutlich höhere Beteiligungsform zeigt und VP-B – im unmittelbaren Vorfeld des Umschlags (13:00-14:00) – auf ein niedriges Niveau in der objektmanipulativen Beteiligung „zwingt“. VP-B’s niedriges Beteiligungsniveau resultiert dabei aus einer Konfiguration, in der A’s summarische Handlungsaktivität

(A-ver+A-O-rel) im besagten Minutenintervall (13:00-14:00) auf einem Wert von 46,5 Sekunden liegt. Interaktant B muss sich folglich, sofern er gewillt ist, normative Erwartungen zu erfüllen, in eine eher defensive Rollenbeteiligung begeben.

Im kontrastiven Vergleich der Beteiligungsformate (A/B-ver und A/B-O-rel) wird deutlich, dass die eigentliche Umschlagsphase als eruptive Transformation eines multimodalen Beteiligungs-niveaus sowohl auf verbalsprachlicher als auch objektrelatierter Handlungsebene identifizierbar ist (in beiden Beteiligungsformaten zeichnet sich der Umschlag bei Minute ~14:00 ab), die zeitliche Dimensionierung des Präkontextes makrostrukturell jedoch nicht präzise ermittelt werden kann. Aus der Divergenz der Merkmalsausprägungen ist der Präkontext infolgedessen auf Makrostrukturebene als linksoffenes Intervall zu kennzeichnen ([14:00, ? , 14:00]), dessen initiale Konfiguration erst nach einer sequenzstrukturellen Analyse auf Mikroebene semantisch identifizierbar ist.

Äquivalent zur verbalsprachlichen Redebeteiligung zeichnet sich eine Umschlagsdynamik ab, die der präkontextuellen Episode unmittelbar nachgestellt ist (ab Minute 14:00). Auch hier steigert VP-B seine Handlungsaktivität aus einem niedrigen Beteiligungs-niveau von 4,1 Sekunden im Minutenintervall 13:00-14:00 auf ein überproportional hohes Beteiligungsformat von 35,4 Sekunden im Minutenintervall 14:00-15:00, das im nachfolgenden Minutenintervall mit 42,3 Sekunden den Maximalwert im Interaktionsverlauf erreicht. Das hohe Beteiligungs-niveau der objektrelatierten Handlungsaktivitäten erstreckt sich dabei über einen Zeitraum von ca. vier Minuten (14:00-18:00). Infolgedessen finden sich für diese Zeitspanne sowohl auf verbalsprachlicher als auch nonverbaler Handlungsebene Hinweise auf eine umschlagende Dynamik.

Betrachten wir hingegen erneut kontrastiv beide Beteiligungsformen (A/B-ver und A/B-O-rel) ab Minute 18:00, so zeigt sich, dass B's verbalsprachlicher Redeanteil bis zum Ende des Versuchs deutlich über dem Niveau seines Ko-Partizipanten liegt, seine objektrelatierten Handlungsaktivitäten jedoch auf ein Beteiligungs-niveau leicht unterhalb der seines Ko-Partizipanten sinken. Aus dieser Beobachtung resultiert die Frage nach der zeitlichen Dimension der Umschlagsdynamik und dem damit verbundenen „Wirkungsgrad“ des Präkontextes.

Unter einer rein verbalsprachlichen Fokussierung wäre die Annahme eines sich bis zum Versuchsende dimensionierten Wirkungsgrades plausibel, da sich für diese Zeitspanne B's konsistent hoher Redeanteil als „Effekt“ der umschlagenden Dynamik abzeichnet (vgl. Abbildung 28). Unter Berücksichtigung der objektrelatierten Handlungen erscheint die Annahme hingegen aufgrund der höheren Beteiligung A's nicht mehr plausibel (vgl. Abbildung 29). Dabei verdeutlicht die Divergenz zwischen den Merkmalsausprägungen die methodischen Grenzen des vorgeschlagenen Programms zur Identifikation umschlagender Dynamiken: Äquivalent zur zeitlichen Dimensionierung des Präkontextes ist der faktische Wirkungsgrad des Präkontextes ohne eine semantische Beurteilung nicht eindeutig bestimmbar und somit als rechtsoffenes Intervall zu kennzeichnen ([14:00, ? []).

Aus der Schwierigkeit der zeitlichen Dimensionierung treffen wir die Unterscheidung zwischen einem „unmittelbaren“ und „potentiellen Wirkungsbereich“. Die empirischen Hinweise verdichten die Annahme, dass der Präkontext – als interaktiv relevant gesetztes Ereignis – sich unmittelbar als homomorphe Ausprägung der umschlagenden Dynamik zwischen A/B-ver und A/B-O-rel im Zeitintervall ~14:00-18:00 abzeichnet (vgl. Abbildung 28 und Abbildung 29). Dabei kann jedoch nicht ausgeschlossen werden, dass auch in der Nachfolge des „unmittelbaren Wirkungsbereichs“ reflexive Sinngabeprozesse mit Bezug zum Präkontext hergestellt werden („potentieller Wirkungsbereich“). Das bedeutet, dass ein weiterer Umschlag in eine neue interaktionale Dynamik (das Minutenintervall 19:00-20:00 stellt hier einen potentiellen Kandidaten dar) zwar den „unmittelbaren“, nicht jedoch notwendigerweise den „potentiellen“ Wirkungsbereich des identifizierten Präkontextes terminiert, da dieser als fossilisiertes Element der Interaktionsgeschichte durch reflexiv- sinnstiftende Handlungsaktivitäten der Interaktanten „reaktiviert“ werden kann. Für unsere Analyse ist die notwendige Unterscheidung der Wirkungsbereiche jedoch an dieser Stelle nicht von Bedeutung, da die fokussierten Phänomene im „unmittelbaren Wirkungsbereich“ der präkontextuell initiierten Umschlagsdynamik liegen (vgl. Abbildung 25).

Der beobachtbare Unterschied in den Beteiligungsformaten ab Minute 18:00 verdeutlicht, dass der kontrastive Vergleich eine wichtige analytische Bewertungsressource darstellt und als epistemisches Instrument für den entwickelten Methodenkomplex zur Eingrenzung eines „unmittelbaren Wirkungsbereichs“ unabdingbar ist. Ferner bestätigt sich dabei auch auf makrostruktureller Ebene die Unzulänglichkeit einer logozentrischen Perspektive als Bewertungsressource, da Veränderungen der interaktionalen

Dynamik durch Übergangsphasen gekennzeichnet sind, die sich zu unterschiedlichen Zeitpunkten in unterschiedlichen Modalitätsebenen manifestieren (vgl. Ausführung zur oben genannten „verbalsprachlichen Fokussierung“). Eine adäquate Bewertung unterschiedlicher Merkmalsausprägungen ist folglich nur unter einer vergleichenden Analyse der verschiedenen Beteiligungsressourcen und genauer Kenntnis des Datenmaterials möglich.

4.4.3 Abweichende Beteiligungsressourcen: Deiktische Gesten

Die Notwendigkeit einer umfangreichen Datenkenntnis bei der analytischen Bewertung kontrastiver Merkmalsausprägungen wird in unserem Setting besonders deutlich, wenn wir die deiktischen Gesten für die einzelnen Minutenintervalle des Versuchs betrachten (vgl. Abbildung 30). In der makrostrukturellen Visualisierung des Beteiligungsformats wurden die einzeln annotierten Handlungsaktivitäten „deiktische Gesten auf Objekte“ (@objects), „deiktische Gesten auf Bereiche der Karte“ (@map) und „deiktische Gesten auf den Ko-Partizipanten“ (@VP-A/B) summarisch für beide Versuchspersonen pro Minutenintervall erfasst.

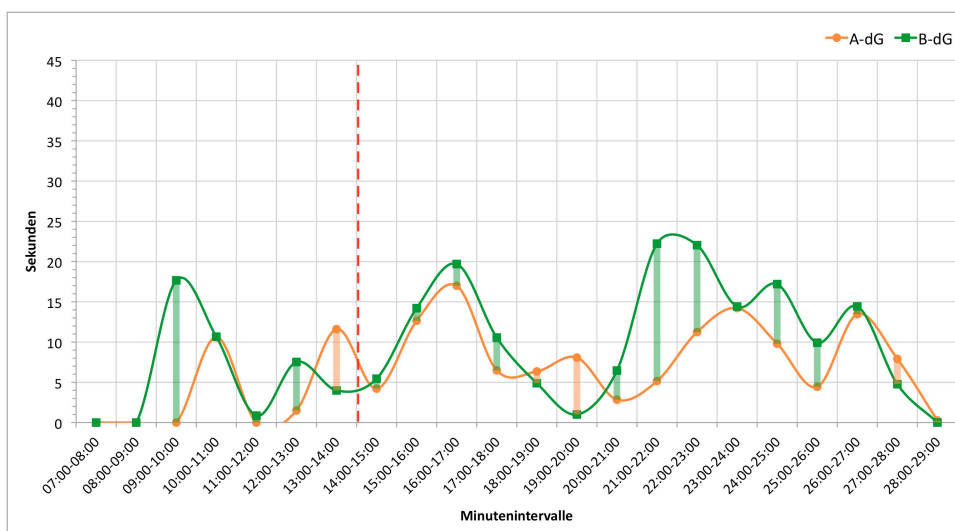


Abbildung 30: Deiktische Gesten (@objects+@map+@VP-A/B) pro Minutenintervall (VP-Gr. 03, F2F)

In einem kontrastiven Vergleich zu den verbalsprachlichen und objektrelatierten Handlungsaktivitäten zeigt das Beteiligungsformat der deiktischen Gesten auffallend divergente Merkmalsausprägungen, in der sich keine eindeutige Entsprechung der bisherigen Ergebnisse zu einer präkontextuell initiierten Umschlagsdynamik abzeichnet. Im Umfeld des zuvor identifizierten Umschlags (gestrichelte Markierung) zeigt sich – äquivalent zu den verbalsprachlichen und objektrelatierten Handlungsaktivitäten – im Präkontext zwar auch ein geringfügig höheres Beteiligungs-niveau VP-A's, in der eigentlichen Umschlagsdynamik ist jedoch kein signifikant höheres Beteiligungs-niveau B's zu verzeichnen. Stattdessen ähneln sich die Merkmalsausprägung im Beteiligungs-niveau der deiktischen Gesten zwischen beiden Versuchspersonen (~14:00-18:00). Infolgedessen stellt sich die Frage, ob die auf verbaler und objektrelatiertes Handlungsebene empirisch begründbare Hypothese verworfen werden muss oder ob deiktische Gesten in ihrer sequenzstrukturellen Funktion von den Teilnehmern anders verwendet werden?

4.4.4 Aktivitätsbasierte Klassifikation der Beteiligungsressourcen

Der Umschlag in eine neue interaktionale Dynamik hat sich unter der kontrastiven Analyse verbaler und objektrelatiertes Handlungen als reziproke Umkehrung aktorspezifischer Aktivitäten im Beteiligungs-niveau abgezeichnet. Aus diesem Grund ist es eine epistemische Notwendigkeit, der Frage nachzugehen, inwieweit die einzelnen Handlungsmodalitäten repräsentative Ressourcen zur Bestimmung eines

Aktivitätsniveaus bilden und inwieweit sich die Handlungsmodalitäten voneinander unterscheiden. Die genaue Kenntnis des Datenmaterials bezieht sich folglich nicht auf eine umfangreiche Erfassung der einzelnen Handlungsabläufe, sondern auf eine präzise Kenntnis des zentralen Aufgabenformats und seiner interaktiven Handhabung durch die Teilnehmer. Das heißt, insbesondere dann, wenn eine Beteiligungsressource auffallend divergente Merkmalsausprägungen aufweist, die sich den bisherigen Ergebnissen nicht zuordnen lassen, sind weitere Formen der Datenexploration notwendig, um die potentiellen Ursachen zu eruieren oder ggf. die Anfangshypothese zu verwerfen. Eine erste Annäherung zur Klassifikation der Beteiligungsressourcen „verbale, objektrelationierte und deiktische Handlungsaktivität“ kann dabei durch einen makrostrukturellen Vergleich erfolgen.

4.4.4.1 Makrostruktureller Vergleich der Beteiligungsressourcen

In Abbildung 31 sind die zentralen Beteiligungsressourcen als dreidimensionales Oberflächendiagramm über den Zeitverlauf der Interaktion in gebinteten Minutenintervallen abgetragen. Beide Visualisierungen sind identisch und unterscheiden sich lediglich in der perspektivischen Darstellung.

Betrachten wir die Aktivitätspotentiale der verbalen und objektrelationierten Handlungen, so ist speziell in der unteren „heatmap“-Visualisierung die präkontextuell initiierte Umschlagsdynamik als reziproke Umkehrung akteurspezifischer Aktivitäten im Beteiligungs niveau erkennbar: A's hohes Beteiligungs niveau im Präkontext (HL) schlägt in ein niedriges Beteiligungs niveau um (LL), während B's niedriges Beteiligungs niveau (LL) im Minutenintervall 13:00-14:00 in ein überproportional hohes Beteiligungs niveau umschlägt (HL). Zwischen den verbalsprachlichen und objektmanipulativen Handlungsaktivitäten lässt sich dabei eine starke Merkmalskorrelation für beide Interaktanten feststellen (vgl. Abbildung 32 a, b). Dies deutet auf eine Form der Kooperativität, in der die objektmanipulativen Handlungen sprachbegleitend, also in einer für den jeweiligen Ko-Partizipanten semantisch nachvollziehbaren Weise markiert werden, wobei sich die Richtung der Relation aus der Kenntnis des Datenmaterials ergibt.

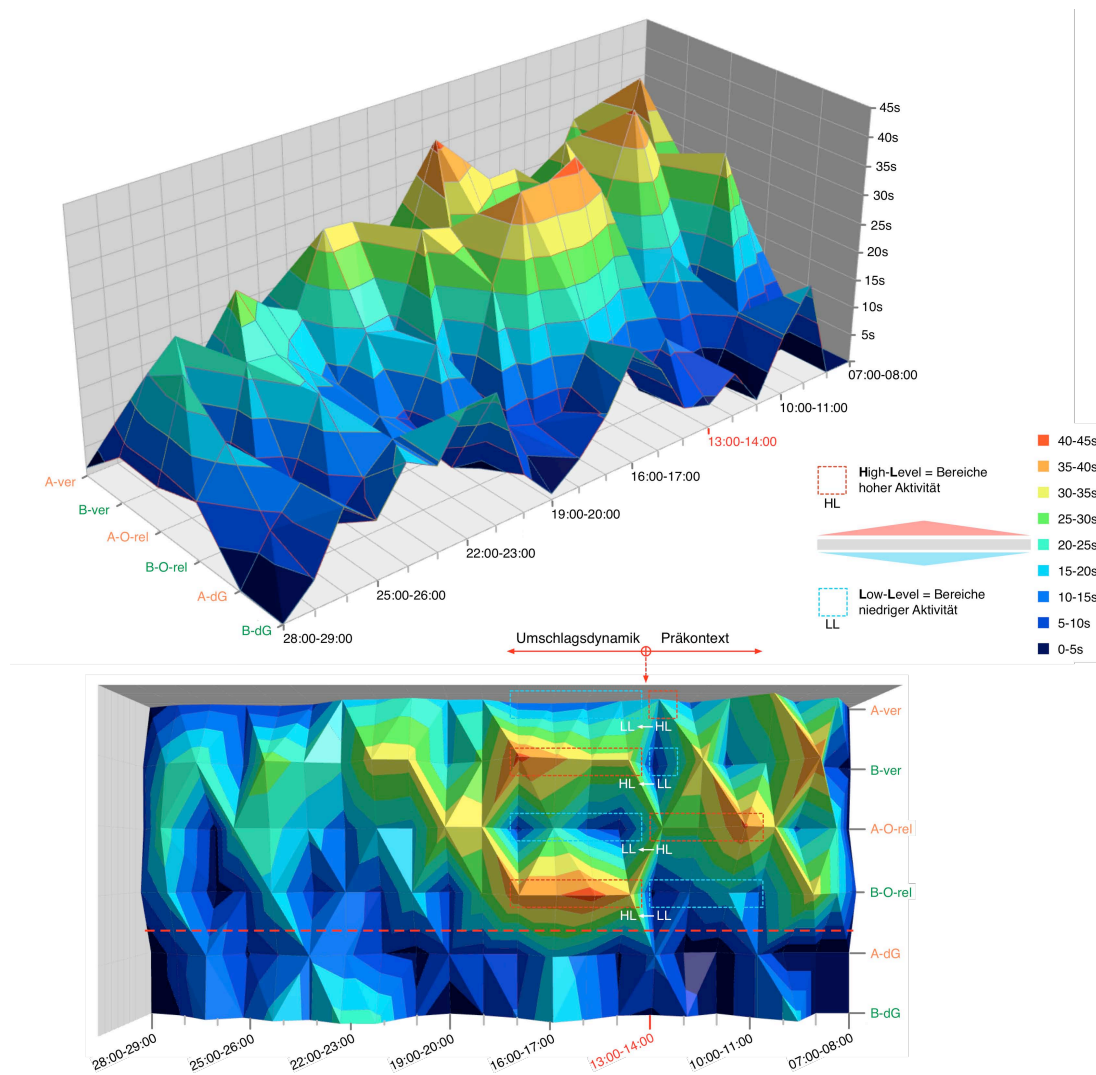


Abbildung 31: Aktivitätspotentiale der Beteiligungsressourcen A/B-ver, -O-rel und -dG pro Minutenintervall

Im Vergleich zur verbalsprachlichen und objektrelatierten Handlungsaktivität ist bei den deiktischen Gesten zum einen ein deutlich niedrigeres Aktivitätsniveau beobachtbar (unterhalb der gestrichelten Markierung). Zum anderen – wie bereits die Visualisierung der einzelnen Aktivitätspotentiale teilweise verdeutlicht – sind mit Ausnahme der stark korrelierten Merkmale B-dG und B-ver (vgl. Abbildung 32, e) keine signifikanten Merkmalskorrelationen zu den übrigen Beteiligungsressourcen feststellbar (Abbildung 32, c, d, f).

Unter einem makrostrukturellen Vergleich der Modalitäten lässt sich das korrelierte Auftreten deiktischer und verbalsprachlicher Handlungen bei Versuchsperson B (vgl. Abbildung 32, e) mit der umschlagenden Dynamik in ein neues Aktivitätsformat erklären. Wie wir bereits festgestellt haben, steigert B neben der objektrelatierten Handlungsaktivität ab Minute ~14:00 vor allem seine Redebeteiligung und hält diese fortan auf einem Niveau, das deutlich über der verbalsprachlichen Aktivität seines Ko-Partizipanten liegt (vgl. Abbildung 28). Vor diesem Hintergrund lässt sich die empirische Tendenz ableiten, dass B die für den Präkontext identifizierte Handlungsdominanz seines Ko-Partizipanten auszugleichen versucht, indem er seine nonverbalen Handlungsvollzüge mit einer erhöhten, sprachlich dominanten Aktivität kontextualisiert. Salopp ausgedrückt, reagiert B auf die präkontextuelle Handlungsdominanz seines Ko-Partizipanten mit dem Versuch, ihn „klein zu reden“, indem er den möglichen Raum, sich im sequentiellen Vollzug verbalsprachlich zu äußern, okkupiert. Die Ursache dieser umschlagenden Dynamik wird dabei noch auf semantischer Ebene zu klären sein.

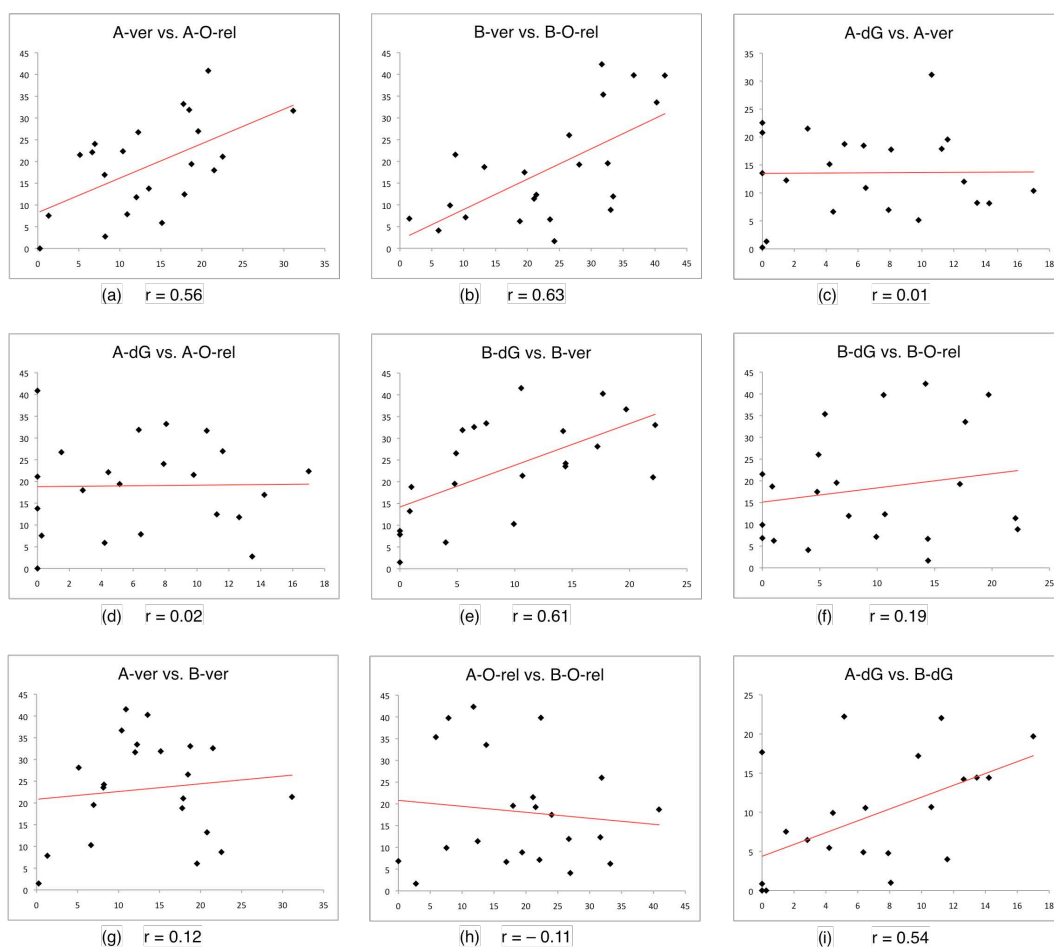


Abbildung 32: Scatter Plots und Regressionsgrade für $n=22$ Fälle: Die Pearson-Korrelation ist in Fall (a), (b), (e) und (i) signifikant ($\alpha = 0,05$)

Während die hohe Korrelation der Merkmale B-dG und B-ver aus der Emergenz der umschlagenden interaktionalen Dynamik heraus plausibel erscheint, bleibt unklar, warum die Verwendung deiktischer Gesten überwiegend in nicht korrelierter Form zu den übrigen – untereinander stark korrelierenden – Beteiligungsressourcen auftritt und ob diese Beobachtung als Hinweis auf ein strukturell anderes Beteiligungsformat zu deuten ist. Demgegenüber ist im interpersonellen Vergleich der deiktischen Gesten (A-dG vs. B-dG) ein korreliertes Auftreten für gebinnte Minutenintervalle im Interaktionsverlauf nachweisbar (vgl. Abbildung 32, i). Da ein solches Emergenzformat, in dem ein hohes Beteiligungs-niveau des einen Akteurs mit einem hohen Beteiligungs-niveau des anderen einhergeht (positive Kovarianz), weder für verbalsprachliche (A-ver vs. B-ver) noch für objektrelationierte (A-O-rel vs. B-O-rel) Handlungen nachgewiesen werden kann (vgl. Abbildung 32, g, h), stellt sich die Frage: Repräsentiert die interpersonell korrelierte Dauer deiktischer Gesten für gebinnte Minutenintervalle ein Beteiligungsformat, das primär in strukturell alignierter Form (Zeigehandlung evoziert Zeigehandlung des jeweiligen Ko-Partizipanten) auftritt und sich somit gesprächsorganisatorisch von den anderen Beteiligungsressourcen zur Bearbeitung interaktionaler Aufgaben abgrenzt?

An dieser Stelle sei darauf hingewiesen, dass es sich um eine generalisierende Einordnung der Beteiligungsressourcen handelt. Wie vor allem die Arbeiten Goodwins zur multimodalen Bedeutungskonstitution in Gesprächen mit einem Aphasiker gezeigt haben (z.B. 1995, 2000b, 2003c, 2010, 2013), können die einzelnen Handlungsmodi (Blick, Gestik, Positur, Prosodie etc.) im jeweils spezifischen situativen Kontext unterschiedliche Funktionen übernehmen. Der Nachweis einer gesprächsorganisatorisch

divergenten Handlungsressource bezieht sich folglich nicht auf den empirischen Einzelfall, sondern fokussiert ein interaktiv etabliertes Handlungsschema, das von den Teilnehmern zur Bearbeitung der Interaktionsaufgabe routinisiert genutzt wird.

4.4.4.2 Mesostruktureller Vergleich der einzelnen Handlungskategorien der Modalität „deiktische Gesten“

Um die gesprächsorganisatorische Funktion der abweichenden Beteiligungsressource „deiktische Gesten“ näher zu ergründen, wechseln wir von der Makro- auf die Mesostrukturebene. Unter einer mesostrukturellen Analyse fassen wir eine Datenexploration, in der die Annotations-/Handlungskategorien „@objects“, „@map“ und „@VP-A/B“ der Modalität einzeln fokussiert werden.

In der Verteilung der Handlungskategorien (vgl. Abbildung 33) fällt auf, dass Zeigehandlungen auf den jeweiligen Ko-Partizipanten ein Randphänomen bilden. Im Hinblick auf eine generalisierende Einordnung der Beteiligungsressourcen sind die drei von VP-B initiierten Vorkommen mit einer Gesamtdauer von ca. 2,2 Sekunden als Konstitutionsmerkmal nicht relevant.

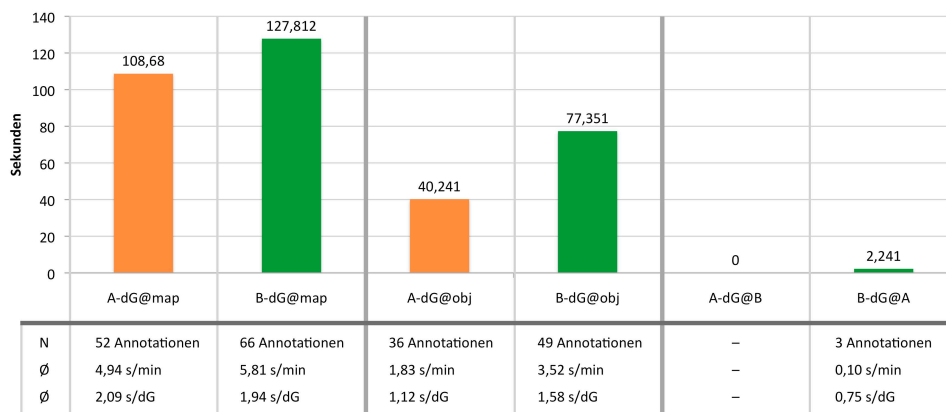


Abbildung 33: Gesamtverteilung der deiktischen Gesten (VP-Gr. 03, F2F)

In den konstitutionsrelevanten Handlungskategorien „A/B-dG@map“ und „A/B-dG@obj“ ist für beide Versuchspersonen ein deutlich höheres Beteiligungs-niveau an deiktischen Verweisen auf die Karte (@map) als auf einzelne Objekte oder Objektgruppen (@obj) zu verzeichnen. Dabei fällt auf, dass deiktische Verweishandlungen der Kategorie „@map“ zwischen den Akteuren ähnliche Merkmalsausprägungen in der Gesamtdauer (A: 108,7s /B: 127,8s) und durchschnittlich aufgewendeten Zeit – sowohl pro Minute (A: 4,9s /B: 5,8s) als auch Gestenhandlung (A: 2,1s /B: 1,9s) – aufweisen, für deiktische Verweishandlungen an Objekten hingegen eine deutliche Handlungsdominanz B's feststellbar ist (A: 40,2s /B: 77,4s). Die Divergenz im Aktivitätsniveau schlägt sich hier sowohl in der durchschnittlich aufgewendeten Zeit pro Minute (A: 1,8s /B: 3,5s) als auch in der mittleren Ausführungszeit deiktischer Gestenhandlungen⁷⁷ nieder (A: 1,1s /B: 1,6s).

Während die Beobachtung ähnlicher interpersoneller Merkmalsausprägungen bei deiktischen Gesten der Kategorie „@map“ nicht als zufälliger, sondern systematischer Effekt zu kennzeichnen ist (vgl. Abbildung 35, j), lässt sich für deiktische Gesten auf Objekte keine signifikante Merkmalskorrelation für

⁷⁷ Die Berechnung erfolgte nicht auf Grundlage einzelner Gesten (i.S. objektspezifischer Referenzierungen), sondern auf „Gestenhandlungen“, innerhalb derer auch solche Sequenzen erfasst sind, in denen verschiedene Objekte oder Lokalisationsorte der Karte sequentiell nacheinander – ohne Unterbrechung der Ausführungskontinuität – referenziert wurden. Das heißt, die unmittelbare Abfolge von Zeigehandlungen auf unterschiedliche Entitäten |@Ent.-1|@Ent.-2|, ist als eine zusammenfassende Handlung |...@Ent.-1-2...| betrachtet worden. Grund dafür ist, dass eine kontinuierlich ausgeführte Gestenhandlung auf verschiedene Lokalisationsorte kaum präzise definiert werden kann. Eine Vergleichbarkeit zwischen den Merkmalen „@map“ und „@obj“ ist folglich nur auf Basis durchschnittlicher Gestenhandlungen sinnvoll, die unabhängig von singulären oder multiplen Referenzierungen anhand ihrer initiiierenden (Onset) und terminierenden (Retraktion) Charakteristika bestimmbar sind.

eine interpersonelle Emergenz nachweisen (vgl. Abbildung 35, k). Infolgedessen ist die Beobachtung einer korrelierten interpersonellen Emergenz deiktischer Gesten für gebinnte Minutenintervalle im Interaktionsverlauf (A-dG vs. B-dG, vgl. Abbildung 35, i) auf die signifikante Merkmalskorrelation interpersoneller Verweisaktivitäten der Handlungskategorie „@map“ (vgl. Abbildung 35, j) zurückzuführen.

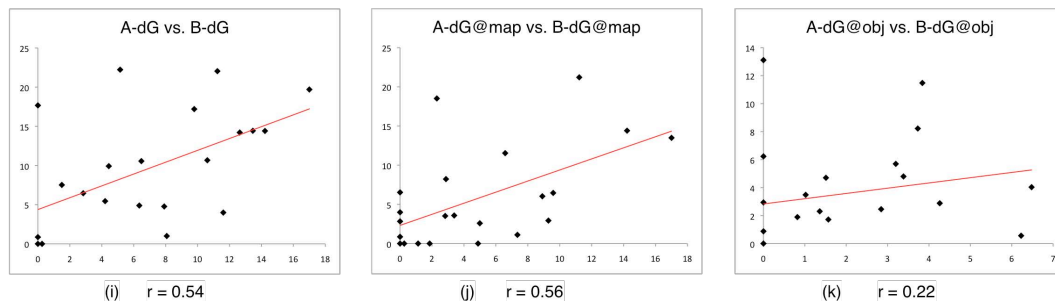


Abbildung 34: Scatter Plots und Regressionsgrade der deiktischen Gesten für n=22 Fälle: Die Pearson-Korrelation ist in Fall (i) und (j) signifikant ($\alpha = 0,05$)

In einem kontrastiven Vergleich der konstitutionsrelevanten Handlungskategorien „A/B-dG@obj“ (Abbildung 35) und „A/B-dG@map“ (Abbildung 36) für gebinnte Minutenintervalle manifestiert sich die Annahme strukturell unterschiedlicher Emergenzformate der Handlungskategorien „@obj“ und „@map“.

Für deiktische Verweise auf einzelne Objekte/Objektgruppen lässt sich unter der phänomenologischen Fokussierung auf eine umschlagende Interaktionsdynamik feststellen, dass sich der auf Grundlage verbalsprachlicher und objektmanipulativer Handlungsaktivitäten definierte Präkontext – als initiiierende Bedingung einer reziproken Umkehrung aktorspezifischer Aktivitäten im Beteiligungslevel – nicht abzeichnet. Deiktische Verweise auf Objekte bilden folglich als konstituierender Faktor der präkontextuellen Konfiguration eine eher untergeordnete Beteiligungskategorie.

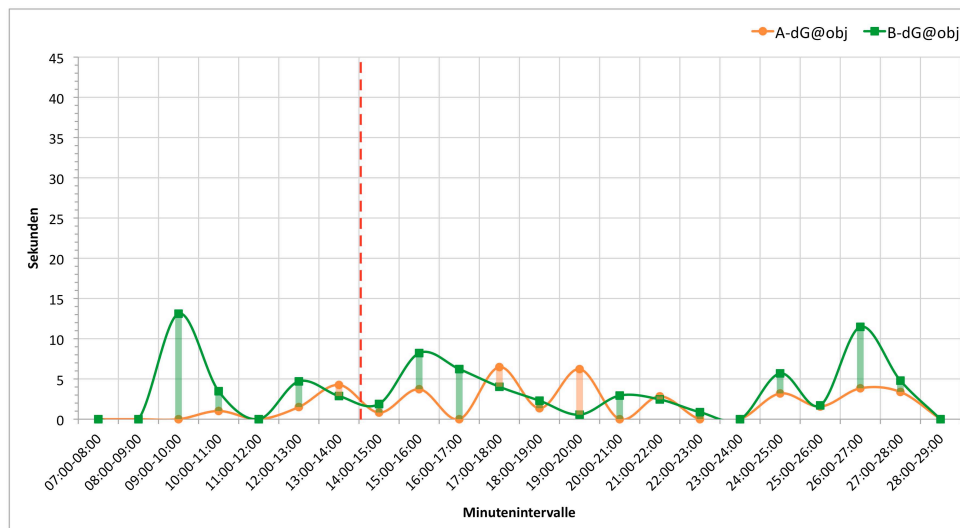


Abbildung 35: Deiktische Gesten (@objects) pro Minutenintervall (VP-Gr. 03, F2F)

Demgegenüber scheint der primär durch verbalsprachliche und objektmanipulative Handlungsaktivitäten konstituierte Präkontext auch einen Effekt auf das Beteiligungslevel der Objektreferenzierungen in der identifizierten Umschlagsdynamik (ca. 14:00-18:00) auszuüben. Die Aktivitätssteigerung VP-B's zeichnet sich jedoch weniger deutlich als in den Handlungskategorien „A/B-ver“ und A/B-O-rel“ ab, da gestische Verweishandlungen auf Objekte generell ein niedrigeres Aktivitätsniveau zeigen (vgl. Abbildung 33, durchschnittliche Zeit an gestischen Verweishandlungen auf Objekte pro Minute: VP-A: 1,8s/VP-B: 3,5s).

Die physische Präsenz der Objekte im gemeinsamen Handlungsraum („shared workspace“) scheint dabei eine verstärkte Objektreferenzierung durch manipulative Aktivitäten zu bedingen. In einer abschließenden mikrostrukturellen Analyse werden wir daher der Frage nachgehen, welche Funktion deiktische Gesten innerhalb des spezifischen Aufgabenformats haben.

Im Vergleich zu den objektreferentiellen Gesten lässt sich für Verweishandlungen der Kategorie „@map“ im Präkontext der umschlagenden Interaktionsdynamik eine ähnliche Merkmalsausprägung feststellen, wie wir sie bereits für verbalsprachliche und objektmanipulative Handlungen analytisch rekonstruiert haben. Die prädominante Handlungsaktivität VP-A's ist jedoch auch hier weniger stark ausgeprägt.

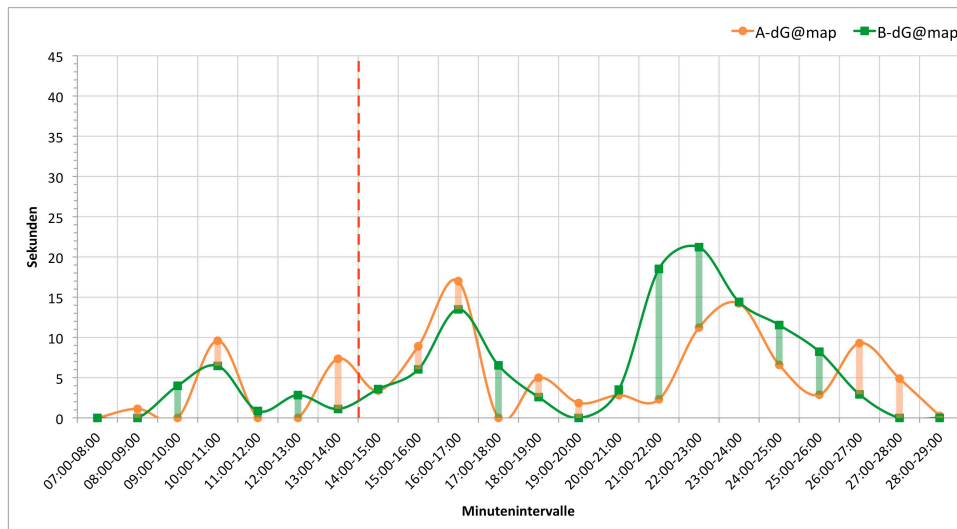


Abbildung 36: Deiktische Gesten (@map) pro Minutenintervall (VP-Gr. 03, F2F)

Betrachten wir hingegen das Zeitintervall der umschlagenden Interaktionsdynamik (ca. 14:00-18:00), die in Bezug auf die verbalsprachlichen, objektmanipulativen und objektreferenziellen Handlungsaktivitäten als eruptive Steigerung des Beteiligungsniveaus VP-B's identifiziert wurde, ist für deiktische Gesten der Kategorie „@map“ keine Handlungsdominanz B's beobachtbar. Trotz einer erkennbaren Aktivitätssteigerung liegt die Beteiligung des Ko-Partizipanten VP-A auf einem ähnlichen, teils prädominanten Niveau (zwischen den Minutenintervallen ~14:00-17:00).

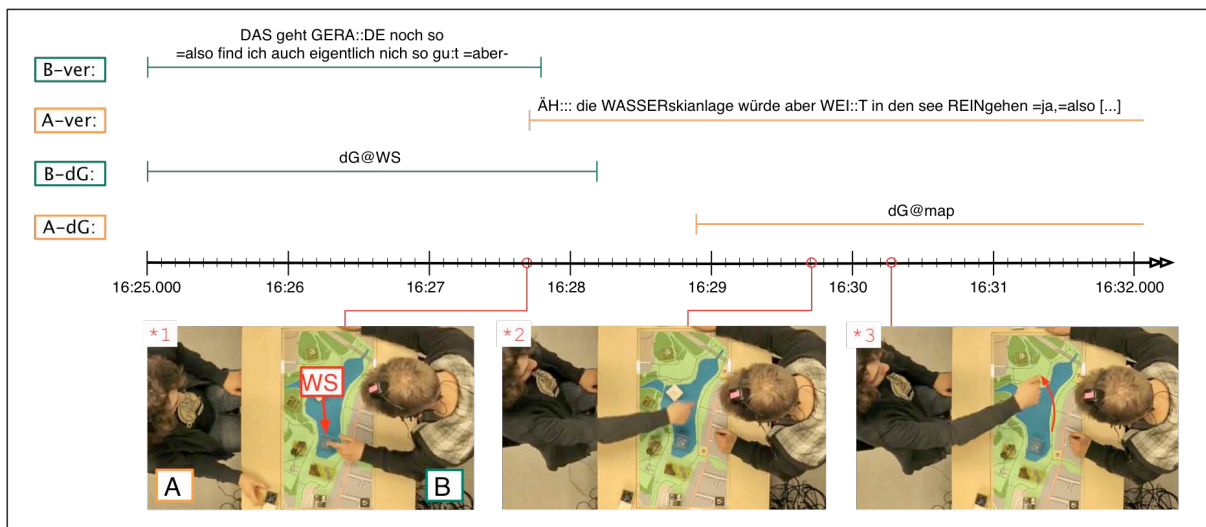
4.4.4.3 Mikrostrukturelle Analyse zur Bestimmung der interaktionalen Funktion strukturell divergenter Handlungskategorien

Während uns die mesostrukturelle Analyse empirisch fundierte Hinweise auf eine strukturelle Divergenz der Handlungskategorien „A/B-dG@map“ und „A/B-dG@obj“ liefert, kann eine funktionale Bestimmung der abweichenden Kategorien nur mikroanalytisch – auf Handlungsebene der Interaktanten – erfolgen. Entscheidend ist hierbei die Frage, ob die auf makro-/mesostruktureller Ebene festgestellte phänomenologische Divergenz der Handlungskategorien „from the data themselves“ (Schegloff & Sacks 1973: 70f.; Sacks 1984: 27) analytisch rekonstruierbar ist und Hinweise auf phänomenologische Ursachen der strukturellen Divergenz liefert. Die Masse der annotierten Vorkommen der Handlungskategorien („@map“: 118 Annotationen; „@obj“: 85 Annotationen; vgl. Abbildung 33) zwingt uns im Folgenden zu einer ergebnisorientierten Darstellung, in der die analytische Rekonstruktion zur funktionalen Bestimmung der Kategorien auf exemplarische Phänomeninstanzen beschränkt wird. Sie dienen der Veranschaulichung und Unterfütterung zentraler Ergebnisse, während auf die funktionale Diversität der Handlungskategorien – im Sinne einer ausdifferenzierenden Analytik – im Rahmen dieser Arbeit nur verwiesen werden kann.

1) Interaktionale Funktionen objektrelationierter Verweishandlungen der Kategorie „@obj“

Auf mesostruktureller Ebene haben wir gezeigt, dass sich für deiktische Gesten auf Objekte oder Objektgruppen keine signifikante Merkmalskorrelation der zeitlichen Dauer für gebinnte Minutenintervalle im Interaktionsverlauf zwischen den Versuchspersonen feststellen lässt (vgl. Abbildung 34). Eine mikroanalytische Rekonstruktion jener Sequenzen, in denen deiktische Gestenhandlungen auf Objekte („dG@obj“) eine semantisch rekonstruierbare und handlungsimplicative Verweisform für sequentiell nachfolgende Objektreferenzierungen („dG@obj“) des Ko-Partizipanten bildet, bestätigt dieses Muster auf sequenzanalytischer turn-by-turn Ebene: Mit lediglich acht identifizierten Vorkommen handelt es sich um eine Verweisform, auf die der Ko-Partizipant primär nicht durch einen Handlungsvollzug der gleichen Kategorie reagiert. Betrachten wir jene acht identifizierten Vorkommen, in denen eine objektreferentielle Geste durch eine typäquivalente Verweisform bearbeitet wird, fällt auf, dass diese interpersonellen Verweiskombinationen in sequenzstrukturell minimalistischer Form als Element der Ko-Referenzierung und/oder der Gegenargumentation auftreten (i.S.v. Objekt X stört Objekt Y). Dementsprechend sind vornehmlich Instanzen der Verweisform „dG@obj“ beobachtbar, in denen erzeugte Handlungsimplicationen vom Ko-Partizipanten in einer typdivergenten Kategorie bearbeitet werden, wie im nachfolgenden Beispiel veranschaulicht werden kann:

FRAGMENT 35 (F2F-BASELINE, VP-GRUPPE 03, VP-A: INVESTOR / VP-B: NATURSCHÜTZER):



Das Fragment setzt ein, unmittelbar nachdem VP-B den von ihm aus rechtseitig liegenden Bereich des Sees als „Wasserschutzgebiet“ ausgewiesen und mit dem entsprechenden Objekt „WaterProtectionArea“ (WPA) markiert hat. Der Beginn des Fragments zeigt die sich sequentiell anschließende Fortsetzungsaktivität, in der Partizipant B mit einer deiktischen Geste auf das Objekt „WaterSki“ (WS) verweist (*1) und konzessiv zum Ausdruck bringt, dass er zwar prinzipiell gegen den Bau einer Wasserskianlage ist, diese jedoch „GERA:DE noch so“ akzeptieren kann. Noch bevor B seine Handlungsaktivität fortsetzen kann („gu:t=aber-“), unterbricht ihn sein Ko-Partizipant VP-A mit „ÄH:: die WASSERSkianlage würde aber WEI:T in den see REINGehen [...]“ und markiert die imaginierte räumliche Ausdehnung der Anlage durch eine deiktische Geste, die den gesamten Bereich, einschließlich des von B ausgewiesenen „Wasserschutzgebietes“, umfasst (*2 & *3).

Dieses kurze Fragment ist insofern charakteristisch, als es zwei zentrale Beobachtungen objektreferentieller Gesten abbildet. Einerseits wird deutlich, dass gewissermaßen keine interaktionale Notwendigkeit besteht, ein deiktisch instantiiertes Objekt erneut unmissverständlich durch eine deiktische Geste auszuweisen. Vielmehr wird eine Angebotsstruktur erzeugt, in der etwaige Aushandlungsbedarfe auf das etablierte Referenzobjekt bezogen werden können (im Beispiel durch die verbalsprachliche Referenz „WASSERSkianlage“). Andererseits wird im Beispiel eine zweite zentrale Beobachtung objektreferentieller Gesten erkennbar: Sie sind vorwiegend nicht in Präferenzstrukturen eingebunden, in denen eine unmittelbare

Reaktion des Ko-Partizipienten konditionell relevant gesetzt wird. Im Beispiel dient die Referenz auf das Objekt „Wasserski“ als ein Zugeständnis im „pre-proposal account“ zum weiteren Ausbau des „Wasserschutzgebietes“: Während B linkshändig auf das Objekt zeigt, manipuliert er bereits rechtshändig das Objekt „FishSign“, welches er nach der Unterbrechung seines Ko-Partizipienten und der daraus emergierenden Aushandlungsphase einführt (deutlich nach der im Fragment skizzierten Episode). Für diese Versuchspersonengruppe lässt sich dementsprechend feststellen, dass das Zeigen auf Objekte überwiegend keine direkten Handlungsimplicationen erzeugt, da objektreferentielle Verweise lediglich ein Bearbeitungsstadium oder eine Prämisse zum eigentlichen Handlungsziel bilden – einer Objektplatzierung und Integration in das Planungskonzept des Naherholungsareals. Handlungsimplicationen werden somit zumeist auf die wahrnehmbare Gestaltung und Beendigung der eigentlichen Zielhandlung verschoben, da die verweisgebende Partei weiter in ihrem Handlungsschema fortfährt und etwaige Handlungsimplicationen durch die Fortsetzungsaktivitäten des zeigenden Akteurs suspendiert werden, bis dieser sein Kernanliegen vollendet hat. Objektreferentielle Gesten treten somit vornehmlich im Rahmen initialer Objektidentifikationen und Argumentstrukturen auf (z.B. 12:59-13:04: „das passt hier ganz GUT zusammen“ + deiktische Gesten auf drei verschiedene Objekte), in denen Vorschläge zur Platzierung konkreter Objekte entworfen werden. In dieser zentralen Handlungsfunktion manifestiert sich die mesostrukturelle Beobachtung, dass sich für deiktische Gesten auf Objekte keine signifikante Merkmalskorrelation der zeitlichen Dauer für gebinnte Minutenintervalle im Interaktionsverlauf zwischen den Versuchspersonen feststellen lässt.

2) Interaktionale Funktionen von Verweishandlungen der Kategorie „@map“

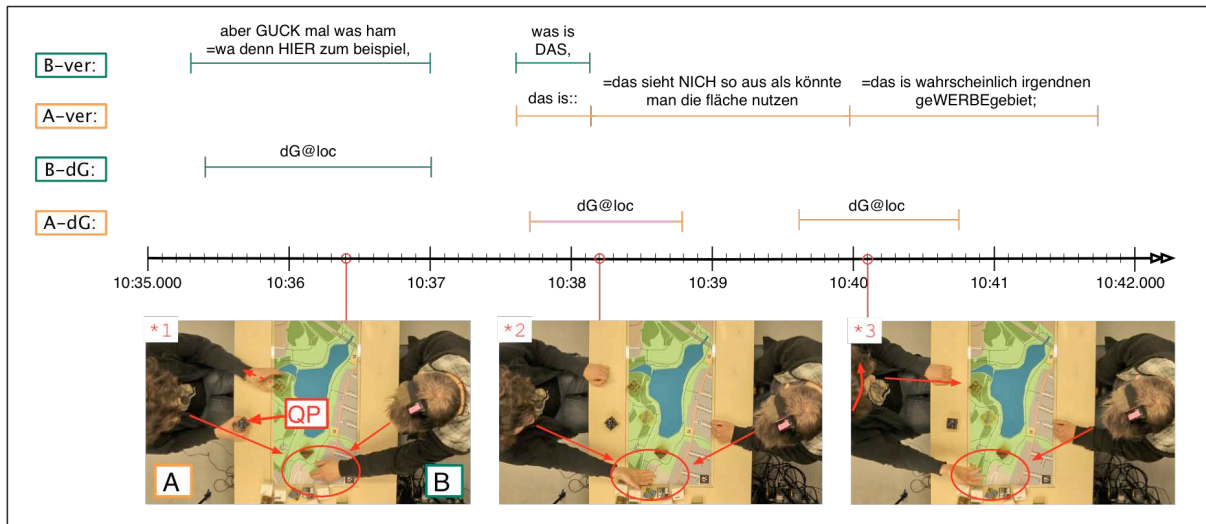
Gegenüber den objektreferentiellen Gesten („dG@obj“, Punkt 1) ließ sich für deiktische Gesten der Kategorie „@map“ eine signifikante Merkmalskorrelation der zeitlichen Dauer für gebinnte Minutenintervalle im Interaktionsverlauf zwischen den Versuchspersonen feststellen (vgl. Abbildung 34). Insofern handelt es sich um eine Handlungskategorie, die gewissermaßen einen ‚Sonderstatus‘ einnimmt, in der sich die umschlagende Interaktionsdynamik ab ca. Minute 14:00 in einer deutlich veränderten Form abzeichnet: Zwar steigert VP-B das Beteiligungsniveau der Handlungskategorie, aber im Gegensatz zu den verbalsprachlichen, objektmanipulativen und objektreferenziellen Handlungsaktivitäten geht diese Aktivitätssteigerung mit einer ähnlichen, teils prädominanten Aktivitätssteigerung des Ko-Partizipienten einher (vgl. Abbildung 36). Der divergenten Merkmalsausprägung der Handlungskategorie „@map“ muss folglich eine besondere analytische Beachtung zuteil werden, wenn wir die Emergenzcharakteristik der Aktivitätsniveaus über den Zeitverlauf verstehen wollen. Die Frage lautet: Warum zeichnet sich der rapide Anstieg im Aktivitätsniveau von Versuchsperson B, der sich deutlich für verbalsprachliche und objektmanipulative Handlungsweisen abzeichnet und in schwächerer Ausprägung auch für deiktische Gesten auf Objekte beobachtbar ist, nicht in einer äquivalenten Form für deiktische Gesten auf die Karte ab? - Also in einer Form, in der das hohe Aktivitätsniveau von VP-B ein niedriges Aktivitätsniveau seines Ko-Partizipienten VP-A bedingt.

Auf der sequenzanalytischen Ebene lassen sich empirische Hinweise auf eine funktional divergente Einbettung der Handlungskategorie in die Bearbeitung interaktionaler Aufgaben finden, die ein interpersonell korreliertes Auftreten erklären. Dabei stellen wir fest, dass deiktische Gesten auf Bereiche der Karte, die aufgrund ihrer symbolischen Topographie als ein Raumbereich semiotischer Vagheit und Vielfalt zu kennzeichnen ist, unterschiedlichste Möglichkeiten zur interaktionalen Fortsetzung generieren. Deiktische Kartenverweise erzeugen hierbei Angebotsstrukturen, die sich in vielgestaltigen Aushandlungsbedarfen und -anforderungen manifestieren. Dabei ist eine empirische Tendenz zur deiktischen mehrfach-Referenzierung beobachtbar, wenn die Referenzentität in ihrer semantischen und räumlich-figurativen Repräsentation nicht eindeutig ist oder von den Interaktanten unterschiedlich interpretiert wird. Das nachfolgende Fragment 36 veranschaulicht diese Praxis.

Im Vorfeld des dargestellten Transkriptausschnittes hatte B das von seinem Ko-Partizipienten A in Seenähe platzierte Objekt „QuadPark“ (QP) in einen weniger prominenten Randbereich der Karte versetzt und dafür sein präferiertes Objekt „NatureTrail“ (NT) platziert. Dieser Eingriff in die konzeptionellen Überlegungen seines Ko-Partizipienten wurde dabei auf verbalsprachlicher Ebene argumentativ legitimiert, indem B die Unvereinbarkeit der Aktivität ‚Quadfahren‘ mit einem Seeblick herausstellte („ich mein beim quad(.)fahren da haste nen HELM auf- da ACHtete auf das FAHREN=da guckste nich auf=n !SEE!“). In Reaktion darauf lehnte sein Ko-Partizipient A die vorgeschlagene Alternativplatzierung

mit dem Hinweis ab, dass es sich um eine bereits bebaute Fläche handele („ja ABER das is eine beNUTZTE FLÄCHE=SIEH: doch mal DA is die STRA:ßenANbindung=da is ein WOHNgebiet oder was weiß ich is DA;“). Der Transkriptausschnitt setzt hier ein und zeigt B’s unmittelbaren Anschlussvorschlag zur alternativen Platzierung des Aushandlungsobjektes “QuadPark” (QP), das aus seiner Sicht, aufgrund des Störpotentials für die Natur, nicht in der Nähe des Sees lokalisiert werden sollte.

FRAGMENT 36 (F2F-BASELINE, VP-GRUPPE 03, VP-A: INVESTOR / VP-B: NATURSCHÜTZER):



Den Ausgangspunkt des Fragments bildet somit eine Situation, in der Interaktant B abermals eine geeignete Fläche zur Replatzierung des Objektes “Quadpark” (QP) sucht und im Rahmen dieser interaktionalen Aufgabe einen spezifischen Ort der Karte vorschlägt, den er in einer multimodalen Gestalt aus verbalsprachlicher und gestischer Referenzierung als potentiellen Kandidaten ausweist (*1). Seine anschließende Nachfrage „was is DAS,“ bezieht sich hier auf die präkontextuelle Einbettung, in der sein initialer Alternativvorschlag von seinem Ko-Partizipienten A mit dem Hinweis einer bereits bebauten Fläche abgewiesen wurde. Im Rahmen der übergeordneten Zielsetzung, einen möglichst weit vom See entfernten Platzierungsort zu finden, kommt der Nachfrage B’s eine illokutive Kraft zu, in der nicht sachorientiert, sondern in einer direktiven Qualität angefragt wird, ob es sich denn bei der nun ausgewiesenen Fläche um eine bebaubare handelt. In A’s Reaktion auf B’s initiales Verweisformat [„aber GUCK mal was ham=wa denn HIER zum beispiel“] + [deiktische Geste] ist erkennbar, dass er zuerst mit einer Bearbeitung der durch die Frageform initiierten Präferenzstruktur beginnt („das is::“) und dann parenthetisch sein Kernanliegen⁷⁸, die ‚Abwehr‘ einer alternativen Platzierung und damit einhergehende Relevantsetzung seiner initialen Platzierung in Seenähe formuliert („=das sieht NICH so aus als könnte man die fläche nutzen“), bevor er die Präferenzstruktur mit der Beantwortung der Frage schließt („=das is wahrscheinlich irgendnen gewERBEgebiet;“). Auffällig ist hierbei, dass sowohl in der Parenthese als auch in der Bearbeitung der Präferenzstruktur deiktische Verweise auf die situativ relevante Referenzentität erfolgen (*2+*3).

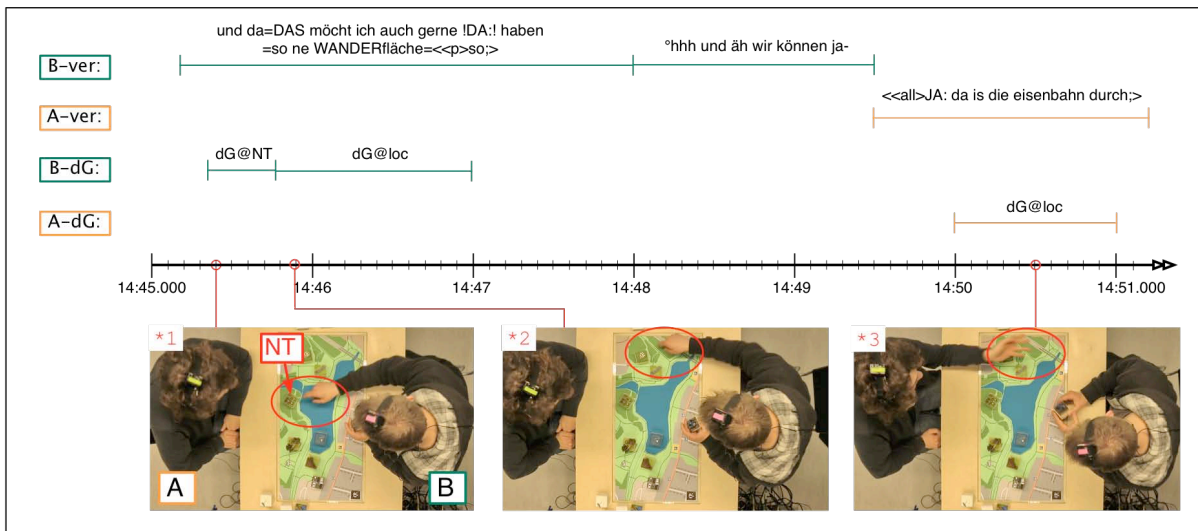
Hier manifestiert sich die zentrale empirische Beobachtung, dass mit einem offenen Aushandlungsstatus der Referenzentität eine tendenzielle mehrfach-Referenzierung durch deiktische Gesten einhergeht. Gegenüber objektreferentieller Gesten („@obj“, vgl. Punkt 1) ist die Raumstruktur der topographischen Karte in einem höheren Maße individuell auslegbar und anzweifelbar. Insofern kann die faktische Bedeutung der Referenzentität im Aushandlungsprozess bestimmt werden. Dieser Sinngebungsprozess scheint in einer haptisch zugänglichen und individuell auslegbaren Raumstruktur an die deiktisch-gestische Verweisgebung gebunden zu sein. Konfligierende Absichten und individuell divergente Auslegungen der

⁷⁸ Dieses Kernanliegen A’s, sich gegen die geplante Verlegung des Quadparks in den Randbereich zur Wehr zu setzen, bestätigt sich vor allem retrospektiv, in der unmittelbaren Anschlusssequenz des dargestellten Transkriptausschnittes: A macht B’s Objekttausch (“NT” anstelle von “QP” in Seenähe) rückgängig, indem er das Objekt “NT” abseits der Karte platziert und dafür wiederum das Objekt “QP” auf seinen initialen Platzierungsort in Seenähe stellt.

Raumstruktur werden dabei oftmals mit finiten gestischen Akzenten zur abschließenden Markierung der Referenzbestimmung versehen.

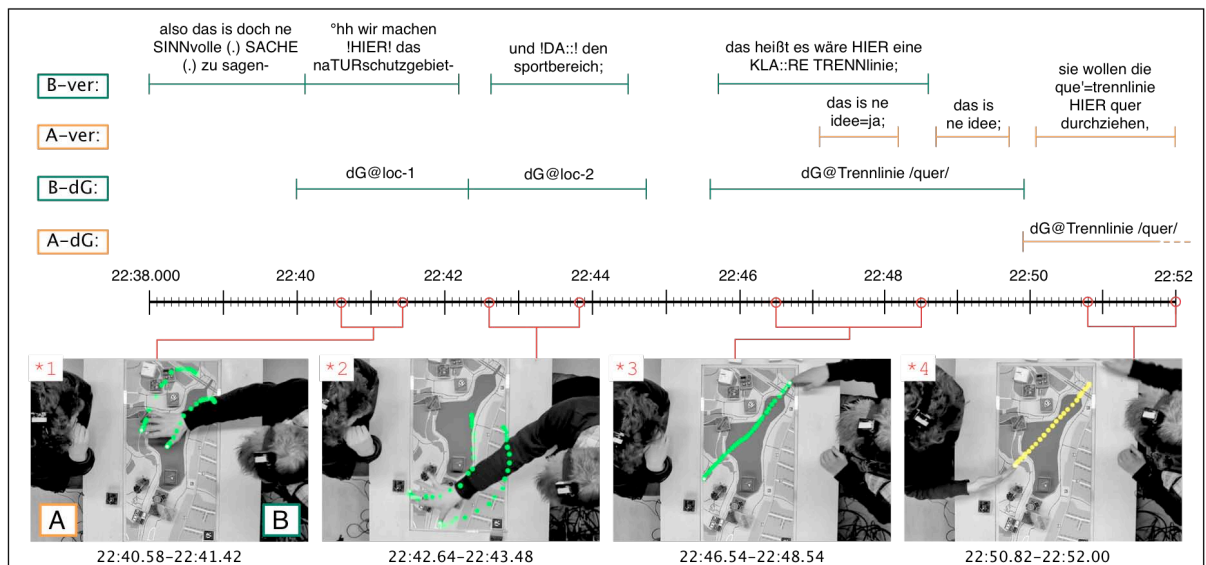
Eine weitere funktionale Verwendung deiktischer Gesten auf die Karte, die ein interpersonell korreliertes Auftreten erklärt, manifestiert sich im nachfolgenden Beispiel (Fragment 37). Es zeigt einen Ausschnitt aus dem fortgesetzten Dissens zwischen den Interaktanten bezüglich der strittigen Verwendung der in der Nähe des Sees befindlichen Fläche und der Frage, ob dort ein Quadpark (QP) oder ein Naturlehrpfad (NT) platziert werden soll (vgl. Fragment 36). Im Beispiel wird deutlich, dass die semiotische Angebotsstruktur der topographischen Karte argumentativ genutzt wird, um sich gegen formulierte Planungsvorhaben des jeweiligen Ko-Partizipanten zu stellen.

FRAGMENT 37 (F2F-BASELINE, VP-GRUPPE 03, VP-A: INVESTOR / VP-B: NATURSCHÜTZER):

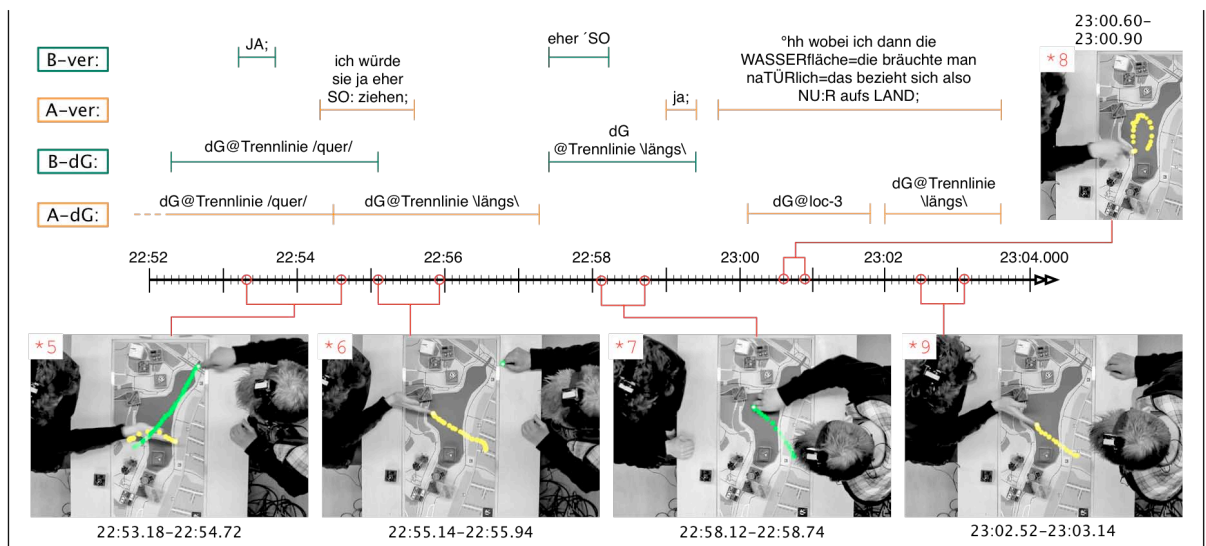


Zu Beginn des Fragments erweitert VP-B, der im Vorfeld erneut das Objekt “NatureTrail” (NT) im Austausch für den Quadpark (QP) auf die kontrovers diskutierte Fläche in Seenähe gestellt hatte, seine Forderung. In der erweiterten Forderung deutet B deiktisch auf sein Objekt “NatureTrail” (NT) („und da=DAS“, *1) und weist insistierend eine zweite Fläche als „WANDERfläche“ aus („möchte ich auch gerne !DA! haben“, *2). Während er in seinen Aktivitäten fortfährt, sich von der indizierten zweiten „Wanderfläche“ abwendet und ein weiteres Planungsvorhaben formuliert („°hhh und äh wir können ja-“), unterbricht ihn A, indem er mit einer deiktischen Geste auf die eingezeichnete Bahnstrecke verweist (*3) und diese als Gegenargument zur von B geforderten zweiten „Wanderfläche“ herausstellt („<<all>JA: da is die eisenbahn durch;>“).

In einem abschließenden Beispiel (Fragment 38) kann gezeigt werden, dass im geschilderten Handlungskontext der Dissens zum Teil sequentiell expandierte Aushandlungsphasen emergieren, in denen die Interaktanten wechselseitig ihre konzeptionellen Vorhaben und Ansichten auf der Karte ‚inskribieren‘.



Den Ausgangspunkt der Analyse stellt eine Situation dar, in der Partizipant B einen Einigungsvorschlag formuliert, der durch die Zweiteilung des Areals ein Nebeneinander der konfligierenden Planungskonzepte ermöglichen soll: In flächigen Gestengestalten („palm down open hand“) markiert er zuerst einen Bereich, den er verbalsprachlich mit „°hh wir machen !HIER! das naTURschutZgebiet-“ ausweist (*1), gefolgt von einem zweiten Bereich, auf den er verbalsprachlich mit „und !DA::! den sportbereich;“ deutet (*2). Anschließend grenzt er die beiden Bereiche durch eine gestische Markierung voneinander ab, die er seinem Ko-Partizipanten verbalsprachlich als „KLA::RE TRENNlinie;“ anbietet (*3). Darauf reagiert A durch eine positive Bestätigung des Vorschlags („das is ne idee=ja; (-) das is ne idee;“), der er eine gestische Markierung anschließt, in der B's ausgewiesene Trennlinie mimetisch nachvollzogen (*4) und zur Verständnissicherung verbalsprachlich mit „sie wollen die que=trennlinie HIER quer durchziehen,“ adressiert wird.



Auf die konditionelle Relevanz der Nachfrage reagiert B durch eine ratifizierende Gestalt aus verbalsprachlicher („JA;“) und gestischer Bestätigung, in der er die „Trennlinie“ erneut markiert (*5). Noch bevor B die Bearbeitung der erzeugten Relevanzstruktur durch die Rückführung der ausführenden Gestenhand in ihre *homeposition* beendet, setzt A mit einer individuellen Modifikation des Einigungsvorschlags ein,

in der er verbalsprachlich seine Idee präsentiert „ich würde sie ja eher SO: ziehen;“ und gestisch eine orthogonal versetzte „Trennlinie“ zieht (*5+*6). Retrospektiv erscheint A’s mimetische Inskription der von B vorgeschlagenen „Trennlinie“ hier weniger ein Element der Verständnissicherung als ein planerisches Füllelement zur eigenen Vorschlagskonstitution einzunehmen. B’s Anschlussreaktion, in der er äquivalent zur Strategie seines Ko-Partizipanten den Vorschlag durch eine Nachfrage („eher ‘SO‘“) und der mimetischen Darstellung der „Trennlinie“ bearbeitet (*7), lässt ebenfalls auf die Funktion der Planungsvorbereitung schließen. Auch hier werden die Aktivitäten der vorschlaggebenden Partei visuell fokussiert, so dass von ihrer Nachvollziehbarkeit auszugehen ist. Insofern dient die mimetische Inskription einer projizierten oder behaupteten Verständnissicherung, um Freiräume zur Erarbeitung individueller Vorschlagsmodifikationen zu schaffen, in denen die eigenen Interessen verfolgt werden können. Auf die in der projizierten/behaupteten Verständnissicherung erzeugte Relevanzstruktur reagiert A, indem er einerseits verbalsprachlich die mimetische Inskription des Alternativvorschlages ratifiziert („ja;“) und andererseits eine den Alternativvorschlag spezifizierende Modifikation anschließt: Dabei verweist A auf den Bereich des Sees (*8) und schränkt die Gültigkeit der „Trennlinie“ auf den Landbereich ein, bevor er das modifizierte Verständnis mit einer finalen Markierung der „Trennlinie“ gestisch akzentuiert (*9).

4.4.5 Zusammenfassung: Kernbeobachtungen und ‚Seiteneffekte‘

In den bisherigen Analysen haben sich empirische Hinweise auf eine präkontextuell initiierte Umschlagsdynamik verdichtet, in der VP-B ab ca. Minute 14:00 sein Beteiligungsniveaus ‚eruptiv‘ erhöht, was sich ‚disruptiv‘ auf das Beteiligungsniveaus seines Ko-Partizipanten VP-A auswirkt. Während sich die Veränderung im Beteiligungsniveaus der Interaktanten deutlich auf der verbalsprachlichen und objektmanipulativen Handlungsebene abzeichnet, tritt sie für „deiktische Gesten auf Objekte“ („@obj“) in verminderter und für „deiktische Gesten auf die Karte“ („@map“) in einer divergenten Merkmalsausprägung in Erscheinung (in positiver Kovarianz).

Um diese Unterschiedlichkeit in der Emergenzcharakteristik zu eruieren, haben wir in 4.4.4 eine aktivitätsbasierte Klassifikation der Beteiligungsressourcen vorgenommen. Dabei konnte gezeigt werden, dass die abweichenden Merkmalsausprägungen („Seiteneffekte“) kein zufälliges Produkt der interaktionalen Emergenz bilden, sondern einen systematischen Effekt der aktivitätsspezifischen Aufgabenbearbeitung. Die dynamischen Veränderungen im Beteiligungsniveau wirken sich besonders stark in der verbalsprachlichen und objektmanipulativen Handlungsmodalität aus, da sie – aus einer verfahrenübergreifenden, handlungsschematischen Perspektive betrachtet – das zentrale „action package“ (z.B. Goodwin 2007, 2010) zur Bearbeitung der interaktionalen Aufgabe konstituieren⁷⁹, in der Objekte vorgeschlagen und hinsichtlich einer geeigneten Platzierung ausgehandelt werden. Demgegenüber fungieren deiktische Gesten eher als spezifizierende Konstituenten im Rahmen argumentativer und gegenargumentativer Handlungsweisen, in denen multireferentiell die Eignung der Orte zur Objektplatzierung ausgehandelt wird. Diese funktionale Unterscheidung spiegelt sich in der Distribution der Beteiligungsniveaus verbalsprachlicher, objektmanipulativer und deiktischer Aktivitäten wider: So liegt der summarische Wert für verbalsprachliche und objektmanipulative Handlungen auf einem ähnlich hohen Niveau (A+B-ver: 811,2s = 36,9s/min vs. A+B-O-rel: 819,2s = 37,2s/min), während sich deiktische Handlungen in einem deutlich niedrigeren Aktivitätsniveau abzeichnen (A+B-dG: 355,8s = 16,2s/min). Die Besonderheit der positiv kovarianten Emergenz deiktischer Verweishandlungen der Kategorie „@map“ ließ sich im Zuge dessen auf einen spezifischen Angebotscharakter zurückführen, der nicht zuletzt durch die individuelle Auslegbarkeit des semiotischen Feldes ‚topographische Karte‘ unterschiedlichste Optionen zur gegenargumentativen Fortsetzung und Behauptung akteurspezifischer Interessen ermöglicht.

⁷⁹ Epistemisch interessant ist diese Beobachtung im Vergleich zur *mutual monitoring*-einschränkenden AR-Bedingung, da sie darauf hindeutet, dass die Verschiebung von Relevanzsystemen und Bearbeitungspräferenzen (vgl. 3.6) nicht zufällig in eine bestimmte Richtung läuft. Vielmehr ist anzunehmen, dass sich die Relevanzsteigerung in genau jenem „action package“ manifestiert, das für die Bearbeitung der interaktionalen Kernaufgaben konstitutiv ist (In unserem Szenario müssen die Interaktanten Objekte manipulieren, um sie auf der Karte zu verorten. Es besteht jedoch keine interaktionale Notwendigkeit, Objekte multireferentiell in relationalen Bezügen zu anderen Objekten und Entitäten der Karte zu elaborieren, was sich darin äußert, dass diese Praxis unter den restringierten AR-Bedingung in deutlich verminderter Ausprägung auftritt (vgl. Abbildung 16)). Ob es sich hierbei um einen systematischen Effekt handelt, kann jedoch nur durch weitere Untersuchungen mit strukturell unterschiedlichen Aufgabenanforderungen und Restriktionen der Interaktionsbedingung bestimmt werden.

4.5 Bestimmung der kernaktivitätsspezifischen Abläufe des Präkontextes

Während wir aus dem bisherigen Vorgehen ein empirisch fundiertes Abbild der präkontextuell initiierten Umschlagdynamik gewinnen konnten, in dem die ‚Seiteneffekte‘ als integraler Bestandteil der aktivitätsspezifischen Aufgabenbearbeitung erklärbar wurden, können wir uns in einem letzten Schritt der konversationsanalytischen Rekonstruktion der handlungsschematischen Ursachen der umschlagenden Dynamik widmen. Analyseleitend ist hierbei die Frage: Welche handlungsschematischen Zusammenhänge lassen sich im Vorfeld der umschlagenden Dynamik identifizieren, die die reziproke Umkehrung akteurspezifischer Aktivitäten im Beteiligungsniveau ab ca. Minute 14:00 erklärbar machen? Oder anders ausgedrückt: Welche empirischen Hinweise lassen sich im Vorfeld der umschlagenden Dynamik finden, aus denen sich die Umschlagsreaktion – oder die Konzeption eines „larger projects“ – von Interaktant B begründet? Die Teilnehmerorientierung an den „kernaktivitätsspezifischen Abläufen“ (Deppermann & Schmitt 2007: 38) bildet somit im Folgenden den zentralen Bezugspunkt der Analyse.

Im kontrastiven Vergleich der Beteiligungsformate (A/B-ver und A/B-O-rel) wurde deutlich, dass sich die eigentliche Umschlagsphase als eruptive Transformation eines multimodalen Beteiligungsniveaus sowohl auf verbalsprachlicher als auch auf objektrelationierter/-manipulativer Handlungsebene ab in etwa Minute 14:00 abzeichnet, die zeitliche Dimensionierung des Präkontextes makrostrukturell jedoch nicht präzise ermittelt werden kann (vgl. 4.4.2). Aus diesem Grund hatten wir den Präkontext auf der Makrostrukturebene als linksoffenes Intervall gekennzeichnet ([] ? , 14:00]). Auf praxeologischer Handlungsebene stellt sich nun die Frage, welche kernaktivitätsspezifischen Abläufe handlungsweisende Elemente der umschlagenden interaktionalen Dynamik bilden und zu welchem Zeitpunkt diese einsetzen. Aus dieser Problematik leitet sich das analytische Vorgehen ab. Ausgehend von den einleitend dargestellten Überlegungen Heritages betrachten wir die unmittelbar vorausgehende Handlungskonfiguration⁸⁰ und arbeiten uns retrospektiv – in Anwendung der Goodwin’schen Metapher “chains of prior action”⁸¹ – bis zur initialen Konfiguration vor.

Da wir in der nachfolgenden Darstellung dem sequentiellen Verlauf der Handlungsemergenz folgen, bedeutet das an dieser Stelle für den Leser, dass sich die präkontextuellen Ursachen der umschlagenden Dynamik erst später, in der Analyse der unmittelbar vorausgehenden Handlungskonfiguration, vollumfänglich eröffnen werden. Die Bedeutungskonstitution und Teilnehmerinterpretation der unmittelbar vorausgehenden Handlungskonfiguration ist jedoch maßgeblich von ihrer initialen Handlungskonfiguration abhängig. Aus diesem Grund werden wir uns in einem ersten analytischen Schritt sozusagen den initialen ‚Bindegliedern‘ der holistischen Handlungsverkettung widmen, bevor diese dann, in einem zweiten analytischen Schritt, auf die unmittelbar vorausgehende Handlungskonfiguration der umschlagenden Interaktionsdynamik bezogen werden können.

4.5.1 Initiale Handlungskonfiguration

Wenn wir der Handlungsverkettung bis zur initialen Konfiguration folgen, gelangen wir zu einer ca. zweiminütigen Aushandlungsepisode, die bereits relativ früh im Interaktionsverlauf mit der Vorschlagskonstitution des Objektes “QuadPark” (QP) in 08:46 (min:sec) einsetzt und für die unmittelbar vorausgehende Handlungskonfiguration bedeutungskonstitutiv ist.

Die kernaktivitätsspezifischen Abläufe dieser Aushandlungsepisode sind in Abbildung 37 dargestellt: Ausgangspunkt bildet A’s Vorschlagskonstitution im Zeitraum zwischen 08:46-09:03, in der er das Objekt “QuadPark” (QP) als Planungskonzept einführt und auf einem spezifischen Ort der Karte platziert (*1). Darauf reagiert B mit einer Gegenargumentation (09:03-09:14), in der er die Entfernung zum Objekt

⁸⁰ “A speaker’s action is context-shaped in that its contribution to an on-going sequence of actions cannot adequately be understood except by reference to its context – including, especially, the immediately preceding configuration of actions – in which it participates” (Heritage 1984: 242).

⁸¹ “Actions exist as rich, temporally unfolding process. Individual actions emerge from, and use, a consequential past shaped through chains of prior action [...]” (Goodwin 2013: 21).

“Hotel” als ungünstige Konstellation anmerkt („meinste NICH das das nen bisschen weit weg is, [...] ich meine wenn man jetzt HIER is=in dem HO:tel und man möchte gerne QUAD fahren“). A ratifiziert diesen gegenargumentativen Hinweis („sie haben VÖLLIG recht=wir RO:den den WALD und stellen HIER die quadbahn hin;“) und replaziert das Objekt “QP” auf einer Waldfläche in unmittelbarer Nähe zum See (09:14-09:17, „@loc-2(Wald)“, *2). In der konsekutiven Handlung gibt B zu verstehen, dass die Rodung des Waldes für ihn keine Option darstellt: Er greift das Objekt aus der manipulierenden Hand seines Ko-Partizipanten A und replaziert es mit den Worten „ich FÄnde es zum beispiel GAR nich schlecht ^h wir BRAUchen GAR nich so viel PA:Rkfläche- wenn wir das QUADding HIER=hin machen;“ auf einen alternativen Lokalisationsort (09:16-09:22, *3). A wendet daraufhin B’s Argumentation mit „ich FINde wir brauchen GAR nich so viel WALDfläche“, replaziert das Objekt “QP” auf das Waldstück in Seenähe („@loc-2(Wald)“) und ergänzt seine Argumentation im “post-proposal account” mit dem Hinweis „also HIER is so=nen WA:Ldgebiet das wird quasi GAR nich touristisch genutzt [...]“ (09:22-09:33, *4).

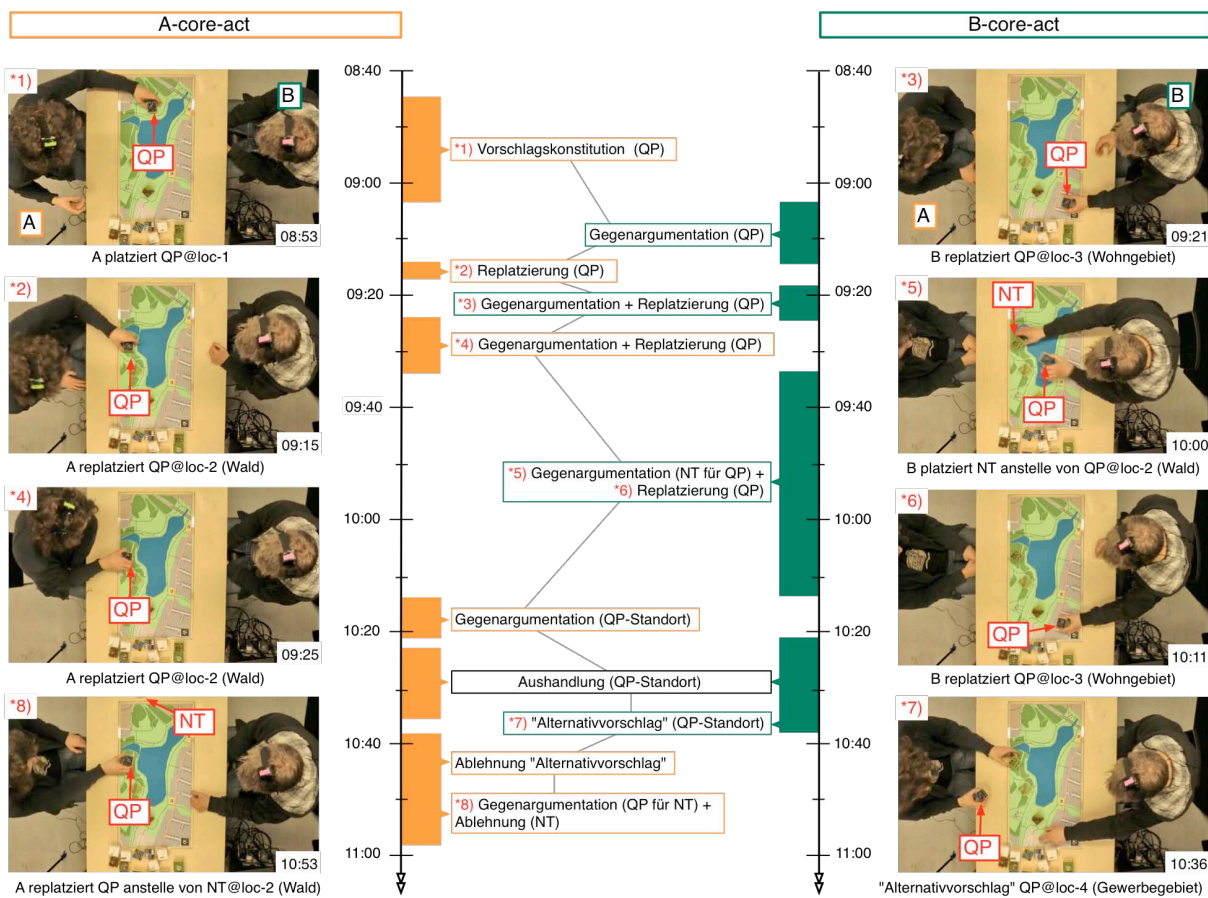


Abbildung 37: Initialkonfiguration der “chains of prior action” im Prädiktor der Umschlagsdynamik

In der sequentiellen Folgehandlung (09:33-10:13) ändert B seine Strategie, indem er den Aspekt der „nicht-touristischen Nutzung“ aufgreift und persuasiv mit einem neuen Vorschlag verknüpft, der auf die Konzeption des Objektes “NatureTrail” (NT) als idealtypisches Planungskonzept für die umstrittene Waldfläche abzielt („dazu kann ich was SA:gen- und zwar kann man es NÄMLICH WUNDERBAR touristisch nutzen (-) ^h mit solchen SCHÖnen WANDERwegen=die ähm zerstören auch die naTUR nich so sehr [...]“). Im Zuge dieser Handlungsstrategie platziert er das Objekt “NT” anstelle des Objektes “QP” auf die umstrittene Waldfläche („@loc-2(Wald)“, *5) und replaziert das Objekt “QP” auf den bereits zuvor (vgl. *3) als Alternativoption ausgewiesenen Lokalisationsort (*6). Dieser wird jedoch von A gegenargumentativ mit dem Hinweis abgelehnt, dass es sich um eine nicht bebaubare Fläche handle („ja ABER das is eine beNUTZTE FLÄCHE=SIEH: doch mal DA is die STRA:BenANbindung=da is ein WOHNgebiet

oder was weiß ich is DA;“, s.a. Fragment 36). Aus der sich sequentiell anschließenden Aushandlungsphase, in der die Interaktanten die aktuelle Bebauung und potentielle Nutzbarkeit unterschiedlicher Flächen der topographischen Karte diskutieren, entwickelt B einen weiteren Alternativvorschlag zur Verortung des Objektes “QP” (*7), der erneut von A mit dem Hinweis einer nicht bebaubaren Fläche abgelehnt wird (s.a. Fragment 36). Dieser Vorschlagsablehnung schließt A eine konterkarierende Handlung an, in der er B’s umgesetztes Planungsvorhaben, den Naturlehrpfad (NT) anstelle des Quadparks (QP) auf die umstrittene Waldfläche zu platzieren, rückgängig macht: Er stellt das Objekt “NT” zurück an den Kartenrand, replaziert das Objekt “QP” auf die Waldfläche („@loc-2(Wald), *8) und legitimiert seine Handlung argumentativ, indem er auf die Klientelunterschiede der Planungskonzepte “NT” und “QP” vor dem Hintergrund einer touristischen und profitablen Nutzung verweist („und außerdem=eine =SEHEN sie doch MAL=das is (-) nen WANDERweg für äh: RENTner=wir WOLLEN hier aber eine äh:: SPORTanlage raus machen [...] wir machen den WALD weg und bauen da ne QUADstrecke; [...]“). Auffällig ist hierbei die multimodal konstituierte Vehemenz, mit der VP-A sein Handlungsziel verfolgt: Er insistiert nicht nur verbalsprachlich auf die Beanspruchung der umstrittenen Waldfläche durch den Quadpark (QP), sondern agiert gleichermaßen resolut und rigoros, indem er das von seinem Ko-Partizipanten B vorgeschlagene Objekt “NatureTrail” aus dem relevanten Handlungsraum entfernt.

Zusammenfassend beschreibt die dargestellte Aushandlungsepisode die interaktive Emergenz einer Ausgangssituation, in der spezifische Objekte (QP & NT) und Bereiche der Karte (Waldfläche) “become a ‘container’ for locally produced meanings”, wie es Streeck (1996: 378) am Beispiel der symbolischen Transformation von Objekten (“things become *situated symbols*”) in einer Geschäftsverhandlung formuliert. Wie wir aus der retrospektiven Analyseperspektive feststellen können, werden in dieser Interaktionsepisode die Objekte „Quadpark“ und „Naturlehrpfad“ initial als symbolische Kontrahenten des Interessenkonfliktes angelegt und funktionalisiert, wengleich dieser Interessenkonflikt durch die aufoktroierte Rollenverteilung im Szenario keinen authentischen, sondern einen spielerischen Reproduktionscharakter hat. Die Handlungen an den materiellen Artefakten „Quadpark“, „Naturlehrpfad“ und der umstrittenen „Waldfläche“ inkorporieren eine “situated history” (ebd.), der die Interaktanten – gewissermaßen als ‚prototypische Dissensmarkierung‘ – eine übergeordnete Bedeutungsrelevanz in der Bearbeitung der interaktionalen Aufgabe und der damit verbundenen Durchsetzung individueller Interessen zusprechen.

Diese Relevanzzuschreibung spiegelt sich in der Gesamtdauer der objektmanipulativen Handlungen der einschlägigen Objekte “QP” und “NT” wider (vgl. Abbildung 38). Sie stellen – gemessen an der zeitlichen Dauer im gesamten Interaktionsverlauf – die ‚manipulationsstärksten‘ Objekte dar, wobei die absolute Manipulationsdauer des Objektes “QP” mit ca. 160 Sekunden in etwa doppelt so hoch ist wie die des Objektes “NT” mit knapp 80 Sekunden (linke Grafik).

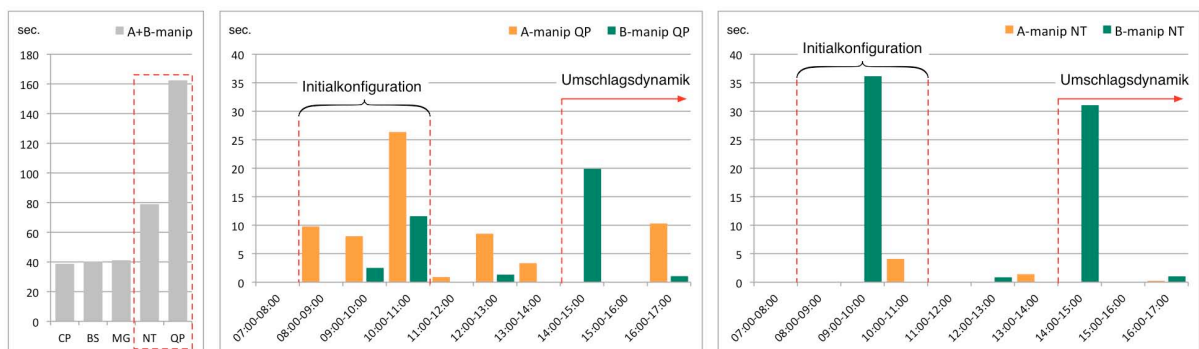
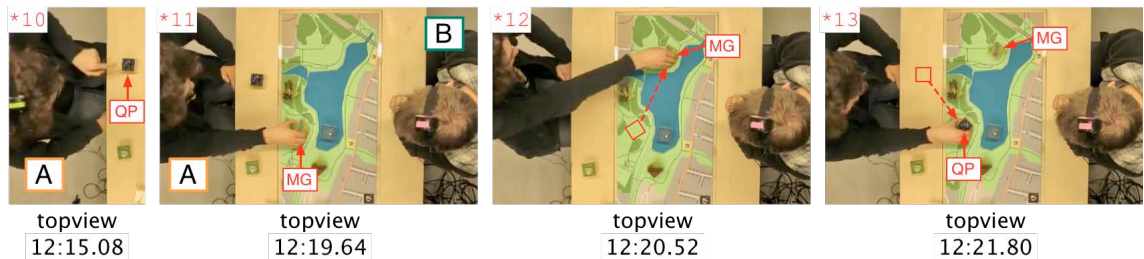


Abbildung 38: Objektmanipulationen (VP-Gr. 03, F2F) – Links: Gesamtdauer der fünf – gemessen an der zeitlichen Dauer – manipulationsstärksten Objekte im Interaktionsverlauf (VP-A und VP-B summiert); Mitte: Manipulationszeiten des Objektes “QuadPark” (QP) in den ersten zehn Versuchsminuten für gebinnte Minutenintervalle (VP-A und VP-B einzeln); Rechts: Manipulationszeiten des Objektes “NatureTrail” (NT) in den ersten zehn Versuchsminuten für gebinnte Minutenintervalle (VP-A und VP-B einzeln).

Betrachten wir die ersten zehn Minutenintervalle des Interaktionsverlaufs für objektmanipulative Handlungen am „Quadpark“ (Grafik in der Mitte) und am „Naturlehrpfad“ (rechte Grafik), zeichnet sich



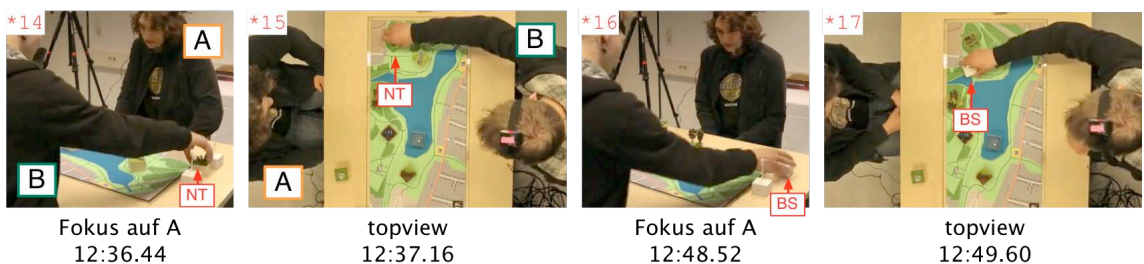
Die Platzierung des Quadparks (QP) in Entsprechung seiner „URsprünglichen IDEE“ (10) erscheint für A jedoch keine adäquate Option zu bilden. Unmittelbar nachdem er das Objekt in ‚Warteposition‘ am Kartenrand abgestellt hat, drückt er mit „ALLERDINGS- (-) mh::“ (12) eine nicht näher spezifizierte Einschränkung oder ein ‚inneres Zerwürfnis‘ seiner „URsprünglichen IDEE“ aus⁸². Im parenthetischen Einschub der projizierten Einschränkung markiert er jedoch – gewissermaßen trotz der angezeigten Problematik einer adäquaten Verortung – die Relevanz des Objektes, das er insistierend als einen der „wahrscheinlich !WICH!tigsten EINnahmefaktoren“ hervorhebt (12+13, *10). Daraufhin schließt A die Parenthese in der Umsetzung einer konkreten Planungsidee, in der sich retrospektiv das ‚innere Zerwürfnis‘ seiner „URsprünglichen IDEE“ manifestiert (14+15): Sein Lösungsansatz besteht darin, das Objekt „MiniatureGolf“ (MG) seiner „URsprünglichen IDEE“ (10) entsprechend auf die indizierte Fläche zu platzieren (*11+*12) und die dadurch entstehende Freifläche für den „Quadpark“ (QP) zu nutzen (*13). Gegen Ende der verbalsprachlichen Erläuterung seines Vorhabens setzt B mit „ALSO ich bin mir NICHT so SICHER;=ich GLAUbe ähm::“ (15) zur Gegenargumentation an, die in Reaktion auf die Behauptung A’s erfolgt, es handele sich um einen der „wahrscheinlich !WICH!tigsten EINnahmefaktoren“ (13).

- 16 B-ver: so=ne QUADstrecke kostet ziemlich viel WA:RTUNGSarbeiten; die
- 17 B-ver: quads gehen mal kaputt und so und die LEUTE haben HIER ZIEMlich
- 18 B-ver: viel alternatIVangebote =die können HIER durch die natUR (-)

- 19 B-ver: wandern °h ähm::|<<p>wegen unserer WANDERstrecke;> die ham HIER
- B-act: |((platziert NT vom stack auf die Karte zurück))
- *14 *15

- 20 B-ver: !SUPER! GEILE aktionen als ANgebote °h wir=ham !WASSER!SKI
- 21 B-ver: =das natürlich nur in=nem beSTIMMTEN bereich des sees is-
- 22 B-ver: damit ähm andere stellen- weil die vögel brauchen ihre RÜhe (-)

- 23 B-ver: |ähm HIER ne, VOGELschutzgebiet- und SO-
- B-act: |((platziert BS vom stack auf die Karte))
- *16 *17

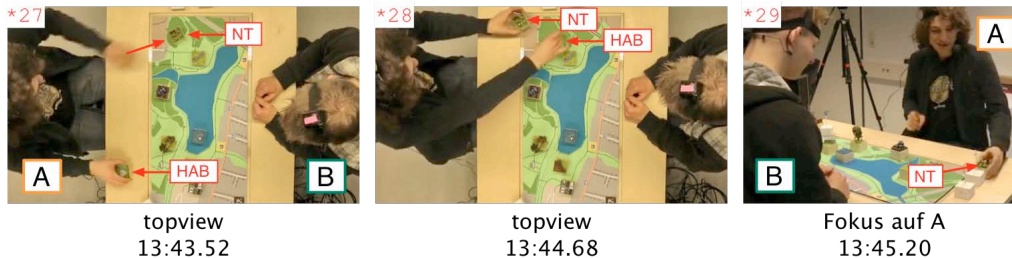


⁸² Auch wenn die genauen Gründe von A nicht weiter expliziert werden und somit einer analytischen Rekonstruktion unzugänglich bleiben, ist darauf hinzuweisen, dass bereits im Rahmen der initialen Vorschlagskonstitution Einwände gegen eine Platzierung innerhalb des indizierten Kartenbereichs von B geäußert („Entfernung zum Hotel“) und von A selbst ratifiziert wurden. Da diese ‚Initialkonfiguration‘ (vgl. Abbildung 37) hier in explizierter Form (A verweist selbst auf die ‚Ursprungsidee‘) die kontextuelle Rahmung des Handlungsgeschehens bildet, ist davon auszugehen, dass die Ergebnisse der initialen Objektaushandlung „QP“ auch in die kontemporäre Aufgabenkonstitution im Sinne des „common ground“ einfließen und als handlungsleitende Prinzipien ‚bedacht‘ werden.

jetzt SO-“ – in der Fortsetzung seines “longer turns” gewähren lässt und sich auf vereinzelte Rezeptionssignale beschränkt (vgl. Abbildung 26, Punkt iii+iv).

40 A-ver: |und |HIER kann man auch noch HEIßluftballon fliegen; (2.4) so;
 B-ver: |!ALSO! |
 A-act: |((platziert HAB anstelle von NT & stellt NT zurück in den stack))
*27 *28 *29

41 A-ver: das wär so meine IDEALvorstellung glaub ich;



In Bezug auf die kernaktivitätsspezifischen Abläufe erscheint vor allem auch der Abschluss dieses “larger projects” von VP-A eine übergeordnete Relevanz für die sich in der Nachfolge abzeichnende „Umschlagsreaktion“ einzunehmen. A ergänzt hierbei seine Konzeption um ein weiteres Planungskonzept, den “Hot-Air-Balloon” (HAB), den er im Austausch für das zentrale Objekt “NatureTrail” (NT) platziert (*27+*28). Im Zuge dessen entfernt er das bedeutungsträchtige, interaktionsgeschichtlich als “indexical monument” (Streeck 1996: 365) der Dissensmarkierung etablierte Objekt B’s (vgl. „Initialkonfiguration“, Abbildung 37) und stellt es zurück an den Kartenrand (*29). Auch wenn die initial aufgezeigte „Marschroute“ „wir machen das jetzt SO-“ im Nachklapp mit „das wär so meine IDEALvorstellung glaub ich;“ (41) eine konjunktivierte Relativierung erfährt, reagiert B in der sequentiellen Nachfolge mit einer massiven Steigerung seines Beteiligungsniveaus, indem er die Ansprüche auf das Rede- und Handlungsrecht seines Ko-Partizipanten mit den analytisch dargelegten „Sequentialisierungsverfahren durch Direktiva“ (vgl. 2.3.7.2) strategisch-lenkend elidiert.

4.6 Zusammenfassung und Fazit

Analytisch motiviert wurde das vorliegende Kapitel 4 aus den empirischen Beobachtungen, die wir im Rahmen einer typologisch besonderen Verfahrensgruppe zur Bearbeitung konfligierender Simultanaktivitäten, den „Sequentialisierungsverfahren durch Direktiva“ (vgl. 2.3.7.2), gewinnen konnten. Ihre typologische Besonderheit manifestiert sich darin, dass die analytisch rekonstruierten Sequentialisierungsverfahren nicht – in Entsprechung des verbalsprachlichen “overlap resolution device” – die Lösung eines systemischen “overloads” an Simultaneität zum Ziel haben (vgl. die analysierten Verfahren der „gestischen Sequentialisierung“, 2.3.7.1), sondern dem strategischen Ziel dienen, individuelle Handlungsentwürfe gegen parallelemergente Handlungsansprüche des Ko-Partizipanten im Sinne einer „Störungsabwehr“ (Schmitt 2005: 33) zu verteidigen. Dabei fiel auf, dass Interaktant B die Darstellungsgelegenheiten objektrelationierter Vorschlagskonstitutionen seines Ko-Partizipanten A über einen längeren Zeitraum systematisch und mit steigender Explizitheit beeinträchtigte. Die strategische Intensivierung des Handlungsanspruchs manifestierte sich in einer repetitiven Konsolidierung der Beteiligungsrollen (“participation framework”), in der B seinen Ko-Partizipanten A immer wieder und mit steigender Vehemenz in die Rolle des „partizipierenden Beobachters“ seiner Handlungsentwürfe zurückdrängte.

Von dieser mikrophenomenologischen Beobachtung ausgehend haben wir die drei analytisch rekonstruierten Fallalternationen der Verfahrensgruppe in den Kontext ihres „handlungsschematischen Zusammenhangs“ (vgl. Deppermann & Schmitt 2007: 38) gestellt, um empirische Hinweise auf die zugrundeliegenden Konstitutionsbedingungen der fallübergreifenden Dynamik zu gewinnen. Die epistemische Relevanz der Frage „Welche kontextuellen Umstände bedingen B’s systematische Beeinträchtigung der Darstellungsgelegenheiten seines Ko-Partizipanten?“ liegt jedoch nicht ausschließlich in der phänomenologischen Fallspezifik begründet, sondern manifestiert sich in einem allgemeinen, sich auf die Emergenz-

charakteristik der Interaktion beziehenden Erkenntnisinteresse: Welcher Konstitutionslogik folgt die Herausbildung interaktionaler Dynamiken? Wie lassen sich diese Dynamiken – verstanden als sequenzstrukturell auffällige Veränderungen im multimodalen Beteiligungsniveau der Interaktanten – analytisch fassen und abbilden? Welche Orientierungslogik liegt ihnen zugrunde, aus denen sich eruptive Veränderungen im Beteiligungsniveau eines Interaktanten – terminologisch mit dem Begriff der ‚Umschlagsdynamik‘ gefasst – abseits rein intrinsischer Motivationen⁸³ erklären lassen?

Dieses Erkenntnisinteresse bildete den Ausgangspunkt unserer Analyse, in der wir – unter einer ‚Emergenzperspektive auf Sequenzierung‘ – ein Methodenprogramm vorgeschlagen haben, das nicht die einzelne Diskurseinheit (‘‘multi-unit turn‘‘, ‘‘larger project‘‘, ‘‘big package‘‘, etc.) ins definitorische Zentrum der Analyse rückt, sondern die Diskurstransformation im Sinne einer Veränderung der interaktionalen Dynamik berücksichtigt. Solche transformatorischen Interaktionsprozesse konstituieren einen komplexen Phänomenbereich, der eine Methodenkombination erfordert, in der die unterschiedlichen Geltungsbereiche einer mikro- und makrostrukturellen Perspektive analytisch gewinnbringend vereint werden, um die epistemischen Grenzen der jeweiligen Ansätze zu überwinden. In besonders anschaulicher Weise manifestiert sich diese Anforderung in den Beobachtungen zu den sequenzstrukturellen Auffälligkeiten der Verbalebene (vgl. Abbildung 25): Unter einer konversationsanalytischen Methodik ist festzustellen, dass B seinem Ko-Partizipanten A über einen längeren Zeitraum das Handlungsrecht zur Vorschlagskonstitution strategisch entzieht. Darüber hinaus lässt sich sequenzanalytisch ermitteln, mit welchen Verfahren B den eigenen Handlungsanspruch durchzusetzen versucht. Einer mikrophänomenologischen Perspektive unzugänglich bleibt hingegen, dass mit dieser interaktionalen Auffälligkeit eine massive Veränderung des Beteiligungsniveaus der Interaktanten einhergeht, in der ein interpersonell ausgeglichenes Redebeitragsniveau in ein prädominantes Redebeitragsniveau im Verhältnis 2:1 umschlägt. Die makrostrukturelle Analyse kann somit als zusätzlicher Bewertungsmaßstab eingesetzt werden, um die interaktionale Relevanz der mikrophänomenologisch identifizierten Episode zu ergründen. Erst aus der perspektivischen Vereinigung wird erkennbar, dass die strategische Handlungsinitiative B's im Kontext einer strukturellen Veränderung des Interaktionsverlaufs steht, die ihre lokalen Herstellungsprinzipien überdauert.

Um diese strukturelle Veränderung und ihre präkontextuelle Verankerung näher zu spezifizieren, haben wir die interaktionale Dynamik der einzelnen Beteiligungsressourcen ‚Verbalität‘, ‚objektmanipulative Handlungen‘ und ‚deiktische Gesten‘ durch eine makrostrukturelle Repräsentationsform über den Zeitverlauf abgebildet (‘‘Binning’’). Im kontrastiven Vergleich der Beteiligungsressourcen fiel auf, dass sich der Umschlag in eine ‚neue‘ interaktionale Dynamik deutlich auf der verbalsprachlichen und objektmanipulativen Handlungsebene als reziproke Umkehrung aktorspezifischer Beteiligungsniveaus abzeichnet: Aus einem auffällig niedrigen Beteiligungsniveau im Vorfeld der umschlagenden Dynamik steigert VP-B sein Beteiligungsniveau eruptiv auf ein episodisches Maximum, das seinen Ko-Partizipanten VP-A auf ein niedriges Beteiligungsniveau ‚drängt‘. Demgegenüber ließ sich für deiktische Gesten eine auffallend divergente Merkmalsausprägung feststellen, in der sich keine eindeutige Entsprechung einer präkontextuell initiierten Umschlagsdynamik abzeichnete. Aus diesem Grund haben wir eine aktivitätsbasierte Klassifikation der Beteiligungsressourcen vorgenommen (vgl. 4.4.4), in der gezeigt werden konnte, dass die abweichenden Merkmalsausprägungen (insbesondere der Handlungskategorie ‚dG@map‘) kein zufälliges Produkt der interaktionalen Emergenz bilden, sondern einen systematischen Effekt der aktivitätsspezifischen Aufgabenbearbeitung (vgl. hierzu die Zusammenfassung aus 4.4.5).

In der abschließenden handlungsschematischen Rekonstruktion der kernaktivitätsspezifischen Abläufe des Präkontextes ließen sich empirische Hinweise auf die interaktionalen Ursachen der umschlagenden Dynamik gewinnen. Dabei wurde deutlich, dass die eruptive Steigerung im Beteiligungsniveau VP-B's nicht in ‚schlichter Anschlussreaktion‘ auf eine vorige Handlungsinitiative seines Ko-Partizipanten erfolgt, sondern aus einem Entwicklungsverlauf emergiert, dessen Initialkonfiguration bereits in einem interaktionsgeschichtlichen Frühstadium – durch die symbolische Transformation der Objekte ‘‘QP’’ und ‘‘NT’’ als prototypische Kontrahenten des Interessenkonflikts – angelegt ist. Im analytisch dargelegten Interessenkonflikt ging Interaktant A wiederholt (in der Initial- und unmittelbar vorausgehenden Handlungskonfiguration) als diejenige Partei hervor, die ihre individuellen Handlungsinteressen bezüglich der interaktio-

⁸³ Das bedeutet nicht, dass der individuellen Konstitution und Aufrechterhaltung einer extensiven Handlungsinitiative eine intrinsische Motivation abzuspochen wäre. Analyseleitend war jedoch vielmehr die Frage, wie sich diese Motivation aus der dynamischen Veränderung der Beteiligungsrollen über den zeitlichen Verlauf der Interaktion begründet.

nen Streitfrage, welches der beiden Objekte für die Platzierung auf der umstrittenen Waldfläche geeignet sei, durchsetzen und die der Gegenpartei konterkarieren konnte. Genau aus diesem wiederholten ‚Scheitern‘ begründet sich die umschlagende interaktionale Dynamik, in der VP-B die Handlungsansprüche seines Ko-Partizipanten mit steigender Vehemenz strategisch zu unterbinden versucht. Er reagiert somit auf eine Gesamtsituation, die aus dem interaktionalen Gleichgewicht der kooperativen Aufgabebearbeitung im Sinne der „joint projects“ gerät und seinem persönlichen Handlungsanspruch zuwiderläuft.

Fazit: Orientierungslogiken in handlungsschematischen Zusammenhängen

Im analytischen Vorgehen des aufgezeigten methodischen Programms ließen sich unterschiedliche empirische Erkenntnisse gewinnen, die abschließend zusammengefasst und vor dem Hintergrund der strukturellen Einheitenbildung diskutiert werden sollen.

Die zentrale Erkenntnis besteht darin, dass sich in Interaktionsverläufen eruptive dynamische Veränderungen zeigen lassen, die sich aus präkontextuellen Inputakkumulationen derjenigen kernaktivitätsspezifischen Handlungen begründen, denen die Interaktanten eine zentrale Relevanz in der Aufgabebearbeitung und Durchsetzung individueller Handlungsinteressen zuweisen. In dieser Beobachtung wird deutlich, dass sich die Interaktanten an handlungsschematischen Zusammenhängen orientieren und ihnen regulative Mechanismen zur Verfügung stehen, wenn aus der Interaktionsgeschichte eine interaktionale Fortsetzung erkennbar wird („proposal of how the future will be organized“, Goodwin 2013: 21), die ihren persönlichen Handlungsanspruch konterkariert. Auf empirischer Grundlage des analytisch dargelegten Falles ist nicht davon auszugehen, dass es sich hierbei um eine einseitige Orientierung handelt, in der Interaktant B aus der interaktionsgeschichtlichen Handlungsdominanz seines Ko-Partizipanten (die Durchsetzung individueller Interessen als Bewertungsmaßstab anlegend) eine aktive Veränderung der Ausgangslage herbeiführt. Im Handeln A's manifestiert sich die Orientierung am handlungsschematischen Zusammenhang gleichermaßen, wenn auch in einer subtileren Form: Einerseits ist erkennbar, dass er seine präkontextuell vorausgehende Handlungsinitiative „wir machen das jetzt SO-“ retrospektiv selbst als eine Normverletzung des kooperativen Handelns bewertet, indem er die aufgezeigte ‚Marschroute‘ im Nachklapp als eine „IDEALvorstellung“ relativiert und legitimiert. Andererseits zeigt sich in den Reaktionen auf die repetitiven Abbruchaufforderungen seines Ko-Partizipanten, dass er sich über einen längeren Zeitraum immer wieder ‚bereitwillig‘ in die Rolle des ‚partizipierenden Beobachters‘ zurückdrängen lässt. Dabei nimmt er diese Rolle nicht nur an, sondern ratifiziert sie in einer prototypischen (Selbst-)darstellung eines beobachtenden Akteurs (vgl. Fragment 7, *5a+b & Fragment 8, *7a+b). Es ist somit davon auszugehen, dass er die repetitiven Abbruchaufforderungen seines Ko-Partizipanten B vor dem Hintergrund seiner vorausgehenden Normverletzung bewertet und für einen spezifischen Zeitraum anerkennt.

Unter einer multimodalen Perspektive auf Interaktion, in der wir der primären Handlungsträgerschaft („Primary Speaker“) die Konzeption dynamisch emergierender Beteiligungsniveaus gegenübergestellt haben, ließen sich empirische Erkenntnisse zur strukturellen Einheitenbildung gewinnen, die das bestehende Paradigma erweitern und präzisieren. Die verbalsprachliche Konzeption eines primären Sprechers, dem das projizierte und beanspruchte Rederecht über einen „multi-unit turn“ zuerkannt wird, erscheint unter einer multimodalen Perspektive aus unterschiedlichen Gründen revisionsbedürftig. Während die übersatzmäßige Beanspruchung des Rederechts in narrativen Strukturen (vgl. 4.2, „joke“- und „story“-Sequenzen) durch Höflichkeitsmaxime gesichert wird, steht ein exklusiver Handlungsanspruch in semiotisch komplexen und situativ zugänglichen Raumsstrukturen unter einem deutlich höheren Kontingenz Einfluss. Die Gattung-, Kontext- und Aufgabenspezifika stellt hierbei eine entscheidende Einflussgröße auf die Herausbildung und Anerkennung struktureller Einheiten dar. Während „joke“- und „story“-Sequenzen die Orientierung der Teilnehmer exklusiv auf die verbalsprachliche Modalität und der in ihr ausgedrückten phantasmatischen Wirklichkeitsbezüge lenken, liegt in unserem Szenario eine materielle Realität vor, auf die sich die Interaktanten beziehen und mit der sie die interaktionale Fortsetzung in situ verändern können. Eine projizierte Struktureinheit im Zuge einer sequentiell elaborierten Argumentationsweise ist somit per se nicht gesichert. Sie liefert fortlaufend Anknüpfungspunkte zur gegenargumentativen Bezugnahme. Infolgedessen ist für die verschiedenen multimodalen Beteiligungsressourcen innerhalb struktureller Einheiten nicht zwangsläufig eine ähnliche Merkmalsausprägung im Beteiligungsniveau festzustellen. Für die Interaktionsform in unserem Szenario ließen sich dabei vor allem deiktische Kartenverweise als eine Handlungskategorie ausfindig machen, die auch innerhalb struktureller Argumentationseinheiten legitime

Anknüpfungspunkte zur Gegenargumentation liefern, aus denen heraus die Struktureinheit ‚aufgebrochen‘ und in ein anderes Bearbeitungsformat überführt werden kann.

In dieser Beobachtung manifestiert sich die empirische Erkenntnis, dass die strukturellen Grenzen eines „larger projects“ nicht a priori festgelegt sind, wie es in etwa bei der Gattung ‚Witz‘ durch die an die Pointe geknüpfte Erwartungshaltung der Fall ist. Die Aufrechterhaltung einer größeren Struktureinheit muss vielmehr aktiv eingefordert, repetitiv ratifiziert und strategisch durchgesetzt werden, da sich die Teilnehmer nicht an der Terminierung, sondern den Anschlussoptionen orientieren, die innerhalb einer strukturellen Einheit gegenargumentative Bezugnahmen ermöglichen. Diese lokale Orientierung an potentiellen Anschlussoptionen wird dabei durch eine globale Orientierungslogik ergänzt, in der sich die Interaktanten am endogenen Handlungskontext orientieren, der aus der dynamischen Abfolge struktureller Einheiten, bzw. der mit ihnen vollzogenen Handlungsinitiativen, emergiert. Epistemisch interessant ist hierbei, dass aus der dynamischen Abfolge solcher Handlungsinitiativen neue Struktureinheiten entstehen können, die einer anderen Konstitutions- und Orientierungslogik folgen: Während im Präkontext der umschlagenden interaktionalen Dynamik – einerseits mit B’s Gegenargumentation (Zeile 15-23) und andererseits mit A’s aufgezeigter ‚Marschroute‘ „wir machen das jetzt SO-“ (Zeile 24-41) – relativ klassische Struktureinheiten zu beobachten sind, verweist die Umschlagsdynamik selbst auf eine Einheitenbildung, die sich bisherigen Beschreibungskategorien entzieht. Die Handlungsinitiative B’s ist nicht dadurch bestimmt, dass A über Minuten hinweg ein Recht zu interaktionaler Beteiligung aberkannt wird, sondern dadurch, dass ihm systematisch das spezifische Handlungsrecht zur objektrelationalen Vorschlagskonstitution mit den analytisch dargelegten ‚Sequentialisierungsverfahren durch Direktiva‘ entzogen wird.

5 Schlussbetrachtung und Ausblick

Auf methodischer Grundlage eines analytischen Kontrastierungsansatzes zwischen Interaktionsbedingungen, die sich in ihrem Koordinierungspotential zum “mutual monitoring” unterscheiden, hatte sich die Arbeit zum Ziel gesetzt, soziale Orientierungs- und Wahrnehmungspraktiken am Beispiel der Ko-Orientierung und Ko-Referenzierung aus einer multimodalen Perspektive auf Interaktion zu erforschen. Damit sollte einerseits auf die prädominante Fokussierung visuell-perzeptiver Beteiligungsweisen im Rahmen der kognitionspsychologischen Perspektive auf “Joint Attention” reagiert werden. Andererseits sollte eine analytische Präzisierung der konversationsanalytischen Grundannahmen des “mutual monitoring” (Goffman 1963; M.H. Goodwin 1980; Clark & Krych 2004) und der holistischen Rezeption kommunikativer Gestalten (Goodwin 2000a, 2003a, 2009) erfolgen, in der deutlich wird, worauf die Interaktanten in spezifischen Interaktionssituationen reagieren und welche semiotischen Ressourcen sie innerhalb der Gestaltformen nutzen, wenn ihnen diese nicht ganzheitlich zur Verfügung stehen. Vor diesem Hintergrund stellte sich die Frage, inwieweit die gegenseitige Beobachtung der Interaktanten und die holistische Rezeption kommunikativer Gestalten eine grundlegende Voraussetzung zum Orientierungserfolg und zur kooperativen Bearbeitung lokal anstehender Interaktionsaufgaben bildet.

Die zentralen analytischen Erträge der bearbeiteten Themenkomplexe werden nachfolgend in konzentrierter Form zusammengefasst und im Hinblick auf ihre epistemische Relevanz und mögliche Anschlussoptionen für die Folgeforschung reflektiert.

5.1 Grundlagenkenntnisse: Orientierungslogiken der F2F-Interaktion

Die analytischen Resultate dieser Arbeit weisen darauf hin, dass die Herstellung von Ko-Orientierung kein ad hoc Phänomen bildet (“looking where someone else is looking” (Butterworth 1998: 171), sondern einen interaktiven Prozess, den die Interaktanten – trotz der teils geringen zeitlichen Expansion im Rahmen der multimodalen Verdichtung von Orientierungshinweisen – in sequentiell-simultanen Verlaufsformen und auf Grundlage von “mutual monitoring”-Aktivitäten kooperativ „erarbeiten“. Ko-Orientierung wird im Kontext der gegenwärtigen Interaktionsaufgabe angekündigt, vorbereitet und etabliert. Infolgedessen ist Ko-Orientierung auf eine Entität nicht präsent oder absent, sondern als gradueller Prozess zu verstehen, der in einer ganzheitlichen Ko-Orientierung auf eine materielle Raumentität oder einen interaktional relevanten handlungsschematischen Zusammenhang mündet. Die Referenzherstellung auf materielle Umgebungsressourcen ist hierbei nicht per se durch eine verbalsprachliche Bezugnahme definiert: Ko-Orientierung und Ko-Referenzierung beschreiben einen nahtlosen Übergang, bei dem die Referenzherstellung im Kern eine Relevanzmarkierung darstellt, die sowohl durch verbalsprachliche („Benennung“) als auch kinesische Salienzkonfigurationen (Drehen, Neigen, Kippen, etc.) des Referenzobjektes erbracht werden kann.

Die analytischen Ergebnisse der F2F-Bedingung weisen im Zuge dessen darauf hin, dass “mutual monitoring” in der Tat eine persistente Grundbedingung in natürlichen “face engagements” bildet (Goffman 1963). Sich am „Verhandlungstisch“ gegenüberstehende Interaktanten beobachten sich selbst bei starker Aufgabenfokussierung permanent und interpretieren die jeweiligen Aktivitäten des Ko-Partizipanten in situ, um sie mit den eigenen Handlungsentwürfen koordinieren zu können. Dieser Mechanismus ist so stark ausgeprägt, dass selbst peripher wahrgenommene Aktivitäten des Ko-Partizipanten (vgl. “peripheral monitoring” (Heath et al. 1995)) auf ihre handlungsimplicativen Wirkungen überprüft werden, um gegebenenfalls Maßnahmen der „Störabwehr“ (Schmitt 2005) einleiten und den eigenen Handlungsanspruch behaupten zu können. Über die empirische Bestätigung dieser Grundannahme hinaus ließ sich jedoch zeigen, dass “mutual monitoring” in seiner interaktionalen Präsenz und situativen Funktion in der Herstellung gemeinsamer Orientierungen von ganz unterschiedlichen Faktoren beeinflusst wird.

Ein grundlegender Faktor manifestiert sich in der Assemblierungs- und Situierungsqualität kommunikativer Orientierungsgestalten. Unter einer massiven Steigerung des analytischen Auflösungs-niveaus, in der die reziproken Minimalreaktionen der Interaktanten aus einer mikrophänomenologischen Analytik untersucht wurden, ließen sich phänomenologische Unterschiede in typäquivalenten Aktivitäten aufdecken. Am Beispiel des Greifens von Objekten in Übergangsphasen interaktionaler Aufgaben konnte

gezeigt werden, dass Orientierungsgestalten in ihrer aufmerksamkeitslenkenden Funktion unterschiedliche Qualitäten besitzen, die sich auf einem Kontinuum zwischen aufmerksamkeitsbindend („Attraktorfunktion“) und aufmerksamkeitsvermittelnd („Mediatorfunktion“) einordnen lassen (vgl. 2.2). Während in der Attraktorfunktion die gestische Verlaufsform sukzessive vom Onset bis zum Stroke auf den Referenten verfolgt wird, fungiert in der Mediatorfunktion bereits der Onset der gestischen Verlaufsform als Orientierungshinweis auf den Suchraum. Die mediatisierende Funktion unterscheidet sich dabei hinsichtlich zweier Rezeptionsstrategien: Entweder ‚wartet‘ die rezipierende Partei auf die Referenzspezifikation in der Strokephase oder sie antizipiert das Referenzobjekt, sofern sie es inferentiell aus der präkontextuellen Einbettung der Orientierungshandlung erschließen kann.

Von dieser empirischen Beobachtung ausgehend ließen sich die zugrundeliegenden Konstitutionsbedingungen der unterschiedlichen Orientierungsformate aus den multimodalen Ablaufformaten der Präkontexte spezifizieren. Dabei konnte gezeigt werden, dass die jeweilige Orientierungsfunktion (Attraktor- vs. Mediatorfunktion) im Wesentlichen von zwei Faktoren abhängt. Der erste Faktor manifestiert sich in der kontextuellen Einbettung der Orientierungsgestalt und der Frage, inwieweit eine Situation vorliegt, in der sich die interaktionale Fortsetzung in einem konversationellen Problem tradiert hat (Ist die gegenwärtige Interaktionsaufgabe wirklich beendet? Ist es legitim, eine neue Interaktionsaufgabe zu eröffnen?). Der zweite Faktor liegt in der Assemblierungsqualität der Orientierungsgestalt begründet, dieses Unklarheitsproblem („situative Ambiguität“) in einer für beide Parteien wahrnehmbaren Form zu lösen („situative Desambiguierung“). Dabei fiel auf, dass sich situativ tradierte Unklarheitsbedingungen (z.B. eine projizierte, jedoch nicht fortgesetzte Aufgabenbearbeitung) in verdichteten „mutual monitoring“ Aktivitäten der Interaktanten abzeichnen. Aus dieser Konfiguration heraus hat die orientierende Gestalt eine aufmerksamkeitsbindende Funktion. Der Ko-Partizipant lässt sich gewissermaßen aus dem Unklarheitsproblem auf eine neue interaktionale Aufgabe ziehen (Attraktorfunktion). Demgegenüber ist die aufmerksamkeitsvermittelnde Orientierungsfunktion (Mediatorfunktion) an die Voraussetzung gekoppelt, dieses Unklarheitsproblem gar nicht erst entstehen zu lassen. Erst aus der instantanen situativen Desambiguierung, in der eine gegenwärtige Interaktionsaufgabe eindeutig abgewählt und eine neue ostentativ angekündigt wird, sind antizipatorische Inferenzprozesse auf den Suchraum oder den Referenten der Orientierungsgestalt möglich.

Neben der phänomenologischen Differenzierung manifestieren sich in den empirischen Beobachtungen zur attrahierenden und mediatisierenden Orientierungsfunktion vor allem zwei zentrale Erkenntnisse, die das bestehende Paradigma erweitern. Der erste Erkenntnisgewinn äußert sich in einem erweiterten Verständnis des indexikalen Charakters orientierender Handlungen: Die kontextuelle Einbettung liefert nicht nur den interpretativen Rahmen einer Orientierungshandlung (z.B. die Erfordernis der situativen Raumwahrnehmung zur Äußerungszeit lokaldeiktischer Ausdrücke), sondern wirkt sich direkt auf die „monitoring“-Aktivitäten der Interaktanten und die Orientierungsfunktion der Gestaltform aus (Attraktor- vs. Mediatorfunktion). Der zweite Erkenntnisgewinn manifestiert sich in einer konzeptionellen Präzisierung der koordinativen Anforderungen: Die Herstellung der Ko-Orientierung liegt nicht allein in der Anforderung begründet, die für die Interaktion gerade relevanten Wahrnehmungen im Raum auszuwählen und zu etablieren (vgl. Hausendorf 2010: 178), sondern im Wesentlichen auch darin, die persistenten Aufmerksamkeitskonkurrenzen der jeweiligen Situation hinsichtlich einer faktischen interaktionalen Fortsetzung aufzulösen. Dieser Aspekt ist deshalb so relevant, da es sich bei der rezipierenden Partei nicht um einen Akteur handelt, der erst durch die eingesetzten Orientierungsverfahren orientiert wird, sondern um einen Akteur, der bereits orientiert ist – und zwar an dem, was die jeweilige Situation an Orientierungsmöglichkeiten bereitstellt. Die Herstellung einer gemeinsamen Orientierung ist somit an den Erfolg der situativen Desambiguierung gekoppelt.

Eine kontextuelle Bedingung, in der die Interaktanten in besonderem Maße gefordert sind, situative Aufmerksamkeitskonkurrenzen hinsichtlich einer gemeinsamen interaktionalen Fortsetzung aufzulösen, stellen parallelemergente Simultanaktivitäten dar. Sie bildeten den empirischen Ausgangspunkt der nachfolgenden Analysesektion (vgl. 2.3), in der wir der Frage nachgegangen sind, mit welchen multimodalen Verfahren Interaktanten das Problem ko-okkurrenter Fortsetzungsoptionen bewältigen. In der empirischen Ausdifferenzierung der alltagspraktischen Desambiguierungsverfahren ließen sich verschiedene Typen von Simultaneität feststellen, die wir Anlehnung an die verbalsprachliche „overlap“-Klassifikation nach Schegloff (2000) zu einer Topographie kinesischer Simultanaktivitäten systematisiert haben. Die Basisdifferenzierung in „kontingente zeitliche Verhältnisse“ und „zeitlich konstituierte Relationen“ (Stukenbrock 2012)

begründete sich aus der empirischen Beobachtung, dass aus einer etischen Perspektive auf multimodale Interaktion das Interaktionsgeschehen zwar als permanente Gleichzeitigkeit von Aktivitäten sichtbar wird, aus einer emischen Perspektive diese jedoch nicht permanent relevant ist. Unmarkiertes körperliches Verhalten bildet eine Normalformervartung, die durch „mutual monitoring“ Aktivitäten begleitet wird, selbst jedoch keine Koordinierungsanforderungen stellt.

Von dieser Basisdifferenzierung ausgehend ließen sich mit der „funktionalen und konfligierenden Simultaneität“ zwei typdivergente Gleichzeitigkeitsformen „zeitlich konstituierter Relationen“ bestimmen, in denen Fortsetzungsprobleme unterschiedlich bearbeitet werden: Während funktionale Simultanaktivitäten durch instantane situative Desambiguierung Ordnungsstrukturen schaffen, in denen ko-okkurente Fortsetzungsoptionen zugunsten einer faktischen Fortsetzungsaktivität elidiert werden, bilden konfligierende Simultanaktivitäten selbst eine Situation ambiger Fortsetzungsoptionen, die sich in einer Störung der kontemporären Aufgabenbearbeitung auswirkt. Erst die Reparaturverfahren konfligierender Simultanaktivitäten haben eine situationsdesambiguierende Funktion. In beiden Verfahrensgruppen verwenden Interaktanten Mechanismen zur Linearisierung komplexer Mehrdeutigkeitsituationen. Während die situativen Desambiguierungsverfahren der funktionalen Simultaneität jedoch proaktiv wirksam sind, indem einer Diskontinuitätsbedingung ein konkretes Fortsetzungsangebot in „sequentieller Lesart“ gegenübergestellt wird, ließ sich für die Bearbeitungsverfahren konfligierender Simultaneität eine reaktive Wirksamkeit feststellen, in der auf die interaktive Störausprägung einer initialen Diskontinuitätsbedingung durch Verfahren der Sequentialisierung reagiert wird. Für die Bearbeitung dieses Fortsetzungsproblems ließen sich unterschiedliche Verfahren rekonstruieren („gestische Sequentialisierung“, „Sequentialisierung durch Direktiva“, „Sequentialisierung durch Blickkoordination“), mit denen Interaktanten emergente Handlungsparallelitäten in ein sequentielles Ablaufformat überführen, das in der Ko-Orientierung auf ein „joint project“ (Clark 1996) mündet.

Aus den empirischen Beobachtungen zur funktionalen und konfligierenden Simultaneität konnten neue Erkenntnisse zur Beziehungskonstellation der Strukturbildungsmechanismen „Sequentialität“ und „Simultaneität“ gewonnen werden, aus denen sich ein grundlegendes Prinzip zur Wahrnehmungsanforderung kommunikativer Gestalten herauskristallisiert hat (Wie wird mittels der Strukturbildungsmechanismen interaktive Ordnung in einer für die Interaktanten nachvollziehbaren Weise hergestellt?). Die bislang empirisch nicht ausdifferenzierte Relation „gleichberechtigt nebeneinander tretender Strukturprinzipien“ (Schmitt 2005; Deppermann und Schmitt 2007) ließ sich aus der mikrophänomenologischen Beobachtung präzisieren, dass die permanente Gleichzeitigkeit simultaner Aktivitäten einer bestimmten Logik folgt. Die Grundordnung des interaktiven Geschehens ist durch gerichtete Ablaufformate gekennzeichnet, die wir konzeptionell als „sequentiellen Simultaneität“ gefasst haben: Interaktanten strukturieren ihre Orientierungs- und Äußerungsgestalten durch die sukzessive Nominierung der Ausdrucksressourcen in „sequentiellen Lesarten“, die vom jeweiligen Ko-Partizipanten für ein sinngemäßes Verständnis erschlossen werden müssen. Im Vergleich zur Verbalsprache handelt es sich hierbei jedoch nicht um „reine Nachzeitigkeit“ (Deppermann & Schmitt 2007), sondern um Mikroformen der Nachzeitigkeit, die auf dem Prinzip der „Teilsimultaneität“ (Stukenbrock 2012) beruhen. In diesen gerichteten Ablaufformaten erscheint die Ko-Okkurrenz verschiedener intra- und interpersoneller Ausdrucksmodi im Gegensatz zur verbalsprachlichen „overlap“-Konzeption relativ unproblematisch. Simultaneitätsprobleme entstehen erst, wenn das Prinzip der Teilsimultaneität verletzt wird. In den Analysen zur konfligierenden Simultaneität wurde deutlich, dass dies insbesondere dann der Fall ist, wenn die intrapersonellen Simultanaktivitäten der Interaktanten zeitgleich einsetzen oder in einer teilsimultan nicht mehr auflösbaren Differenz auf gleicher Modalitäts- und Fokusebene miteinander konkurrieren („echte Simultaneität“). In diesen Fällen annullieren sich die sequentiellen Lesarten der aufeinandertreffenden Teilnehmeraktivitäten gegenseitig und müssen – äquivalent zur physikalischen Ko-Okkurrenz akustischer Signale im verbalsprachlichen „overlap“ – durch Verfahren der Sequentialisierung bearbeitet werden.

5.2 Auswirkungen eines eingeschränkten „mutual monitoring“

Die empirischen Beobachtungen zur Emergenz und Auflösung konkurrierender Aktivitäten im kontrastiven Vergleich der AR- und F2F-Bedingung (vgl. 2.4) haben ein phänomenologisches Unterscheidungskriterium hervorgebracht, das die interaktionalen Auswirkungen des eingeschränkten „mutual monitoring“

beschreibt: Während Interaktanten unter der Persistenz gegenseitiger Beobachtung über ein ausgeprägtes Repertoire an multimodalen Verfahren verfügen, um konfligierende Simultanaktivitäten in sequentielle Ordnungsstrukturen zu überführen, sind sie unter eingeschränktem “mutual monitoring” nur vermindert fähig, ihre Aktivitäten kleinschrittig zu koordinieren, um emergierender Handlungsparallelität entgegenzuwirken. Das Prinzip “order possible at any point” (Schegloff 2005) ist in seiner Wirksamkeit gehemmt. Eine sich verlagernde Orientierung des Rezipienten bildet unter diesen Bedingungen kein stabiles Signalisierungssystem, das dem gegenwärtigen Sprecher erlaubt, Reparaturen zur Rückgewinnung der Aufmerksamkeit einzuleiten. Die Interaktanten agieren unter der Annahme, dass ihre Handlungsaktivitäten vom jeweiligen Ko-Partizipienten wahrgenommen werden („default-Annahme der Ko-Orientierung“). Aus dieser Konfiguration können episodische Handlungsparallelitäten entstehen, in denen die individuellen Handlungsvollzüge der Interaktanten – in unterschiedlich ausgeprägter sequentieller Entkopplung voneinander – das situative Handlungsgeschehen bestimmen.

Obgleich sich in dieser empirischen Beobachtung eine faktische Auswirkung des eingeschränkten “mutual monitoring” manifestiert, fiel auf, dass die Emergenz episodischer Handlungsparallelitäten im Korpus eine phänomenologische Randerscheinung bildet. Trotz Einschränkungen im “mutual monitoring” waren die Interaktanten größtenteils in der Lage, lokal anstehende Aufgaben interaktiv zu bearbeiten und das Handlungsgeschehen mit Blick auf die globale Aufgabenbearbeitung sequentiell voranzutreiben. Aus diesem Grund sind wir der Frage nachgegangen, mit welchen multimodalen Verfahren die Interaktanten in der AR-Bedingung den gesteigerten Koordinierungsanforderungen begegnen, wie diese dazu beitragen, Distorsions- und Störeffekte zu minimieren und welche Rückschlüsse sich daraus auf die Robustheit interaktiver Koordinierungsressourcen/-verfahren im Speziellen und interaktionaler Systeme im Allgemeinen ziehen lassen (Kapitel 3).

Die Analyseergebnisse haben zentrale Stabilitätsmechanismen der sozialen Interaktion offengelegt, die einerseits in den alltagspraktischen F2F-Verfahren selbst verankert sind (z.B. die “attention getter”-Funktion lokaldeiktischer Ausdrücke; die prospektive und retrospektive Orientierungsfunktion von Diskursmarkern), andererseits auf interaktiven Anpassungsstrategien der Teilnehmer beruhen, um das Interaktionssystem ‚funktionsfähig‘ zu halten (z.B. die “wait & see”-Strategie oder die Neuausrichtung von Relevanzsystemen). Dabei konnte gezeigt werden, dass die konversationsanalytische Grundannahme der holistischen Rezeption kommunikativer Gestalten (Goodwin 2000a, 2003a, 2009) keine notwendige Voraussetzung zum Orientierungserfolg und zur kooperativen Bearbeitung lokal anstehender Interaktionsaufgaben bildet. Dies ist insofern bemerkenswert, da sich Interaktanten zwar – sofern es die Interaktionsbedingungen zulassen – permanent beobachten, um das Verhalten des Gegenüber auf etwaige Handlungsimplicationen zu prüfen, jedoch nicht unbedingt auf diese Persistenzbedingung eines ganzheitlichen “mutual monitoring” angewiesen sind.

Das Potential, auch unter massiven Einschränkungen der Koordinierungsressource “mutual monitoring” kooperativ agieren zu können, deutet auf eine genuine Robustheit interaktionaler Systeme, für die sich in den Analysen unterschiedliche Ursachenfaktoren explizieren ließen. Grundlegend erscheint die empirische Feststellung, dass sich die einzelnen Ausdrucksmodi kommunikativer Orientierungsgestalten nicht nur im Sinne der multimodalen Bedeutungskonstitution gegenseitig elaborieren (z.B. Goodwin 2009), sondern in ihrem wechselseitigen Zusammenwirken auch gegenseitig stabilisieren, indem sie selbstreferentiell auf ihre ‚semiotischen Nachbarn‘ innerhalb der Gestaltform verweisen. Für die Interaktanten hat diese Selbstreferentialität handlungspraktische Konsequenzen, die sich in ihren ‚Suchmustern‘ äußern, eine fragmentarisch rezipierte Orientierungsgestalt hinsichtlich ihrer bedeutungstragenden Elemente zu vervollständigen. Besonders deutlich wurde dies in Fällen, in denen lokaldeiktische Ausdrücke eine Orientierungsverlagerung des Ko-Partizipienten initiieren und dieser versucht, den vektorialen Suchraum und das Zeigziel retrospektiv aus der Retraktionsphase der Zeigegeste zu erschließen.

Darüber hinaus ließ sich jedoch auch zeigen, dass Stabilität selbst von den Interaktanten interaktiv hergestellt und erzwungen wird. Bei eingeschränktem “mutual monitoring” agieren die Partizipienten unter einem gesteigerten Toleranzverhalten, in der die Präferenzorganisation offener gehandhabt wird als in der F2F-Bedingung. Dispräferierte Reaktionen werden nicht als Verletzung normativer Erwartungen behandelt, sondern akzeptiert und auf mögliche Anschlusshandlungen geprüft (“wait & see”). Insofern orientieren sich die Interaktanten an einem funktionalen Progressionsanspruch: Solange ein sinnhaftes Fortschreiten der interaktionalen Aufgabenbearbeitung beiderseitig erkennbar ist, besteht kein akuter Handlungsbedarf, dispräferierte Anschlussreaktionen zu bearbeiten. Neben der interaktiven Herstellung

von Normalität gründet die Robustheit interaktionaler Systeme auf dem Potential der Interaktanten, Relevanzsysteme und Bearbeitungspräferenzen neu auszuhandeln. Im kontrastiven Vergleich der Interaktionsbedingungen (F2F – AR-Obersee – AR-Exhibition) ließ sich zeigen, dass sich mit steigenden Koordinierungsanforderungen eine Relevanzverschiebung einstellt, in der die Koordinierung über materielle Umgebungsressourcen als “centers of coordination” (Suchman 1997) einen sukzessiv höheren Stellenwert in der Bearbeitung lokal anstehender Interaktionsaufgaben einnimmt. Die Einschränkung des “mutual monitoring” manifestiert sich dabei in einer Tendenz zur objektzentrierten Bearbeitung, in der die Interaktanten durch unterschiedliche Salienzkonfigurationen (Anheben, Neigen, Drehen etc.) wahrnehmbare Relevanzkategorien durch eine Erhöhung des Figur-Grund-Kontrastes erzeugen.

Betrachtet man die Analyseergebnisse im Hinblick auf mögliche Anwendungsfälle des AR-Systems, wie z.B. “medical visualization, maintenance and repair, annotation, robot path planning, entertainment, and military aircraft navigation and targeting” (Azuma 1997), ist auf unterschiedliche Auswirkungen hinzuweisen, die bei der Konzeption für die soziale Nutzung derartiger Systeme berücksichtigt werden sollten. Unsere Ergebnisse haben gezeigt, dass die Einschränkung des “mutual monitoring” zu einer ,entschleunigten Kommunikation‘ (“slowing-down-communication”) führt, in der die Interaktanten im Vergleich zur F2F-Bedingung nur vermindert Relationen zwischen unterschiedlichen Objekten im Handlungsraum herstellen. Mit steigenden Koordinierungsanforderungen gehen die Interaktanten zu einer objektzentrierten Bearbeitung über, in der raumreferentielle Bezüge weniger Bedeutung erlangen. Während in Anwendungsbereichen, in denen eine Fokussierung auf spezifische Raumentitäten vorteilhaft ist, sich dieser Effekt durchaus positiv nutzen lässt, stellt er in Anwendungsbereichen, in denen die Interaktanten ihre Aufgabe unter Abwägung verschiedener Raumkomponenten bewältigen müssen, ein potentielles Risiko dar. Ebenso ist darauf hinzuweisen, dass die situative Referenzierung auf Raumentitäten unter der Einschränkung des “mutual monitoring” oftmals auf Reparaturen angewiesen ist, wenn die *stroke*-Phase der Verweisgeste nicht erkannt oder inferentiell aus dem Handlungskontext ermittelt werden kann. Dabei erscheint die Reparatur von Verweishandlungen aus einer emischen Perspektive unproblematisch, da die intersubjektive Wahrnehmung einer fehlgeschlagenen Ko-Orientierung von den Interaktanten leicht behoben werden kann. Aus einer etischen Perspektive ist jedoch anzumerken, dass Reparaturen von Ko-Orientierung ,kostenintensiv‘ sind und sich auf die ,zeitliche Taktung‘ lokal anstehender Aufgabenbearbeitungen auswirken.

In diesem Zusammenhang ist auch auf einen Aspekt hinzuweisen, den es im Anwendungsfall unbedingt zu berücksichtigen gilt. In unseren Analysen wurde deutlich, dass in seltenen Fällen (“deviant cases”, 3.2.4) eine Verweishandlung nicht repariert werden kann, da die fehlgeschlagene Ko-Orientierung von den Interaktanten nicht erkannt wird. Ursächlich dafür war, dass der Rezipient die Desambiguierungshinweise der Orientierungsgestalt, wie in etwa die Blick- und Körperausrichtung auf das Referenzobjekt bei beidhändigen Objektmanipulationen, visuell nicht wahrnehmen konnte. In Fällen, in denen die *default*-Annahme der Ko-Orientierung der verweisgebenden Partei mit der sinnhaften Deutung der teilrezeptiv erfassten Angebotsstruktur des Ko-Partizipanten zusammenfällt, können Divergenzen zwischen der projizierten und der interpretierten Orientierungsgestalt entstehen, die u.U. folgenschwere Implikationen in den o.g. Anwendungsbereichen erkennen lassen. Die Problematik liegt hierbei darin begründet, dass beide Parteien davon ausgehen, dass ihr Handeln nicht nur wahrgenommen, sondern auch überprüft und ggf. korrigiert wird.

5.3 Anschlussoptionen für die Folgeforschung

Neben Implikationen für die Konzeption kollaborativer AR-Systeme in sozialen Anwendungsbereichen eröffnen die analytischen Ergebnisse dieser Arbeit verschiedene Anschlussoptionen für die multimodal ausgerichtete Interaktionsforschung.

Auf epistemischer Ebene lassen sich vor allem Anschlussoptionen im Rahmen der Analyseergebnisse zur konfligierenden Simultaneität erkennen. Aufgrund der enormen Variationsbreite, die sich sowohl in der Emergenz (den verschiedenen ,Diskontinuitätsbedingungen‘) als auch der interaktiven Bearbeitung konfligierender Simultanaktivitäten gezeigt hat, ist nicht davon auszugehen, dass die rekonstruierten Praktiken in den identifizierten Verfahrensgruppen der „gestischen Sequentialisierung“, „Sequentialisierung durch Direktiva“ und „Sequentialisierung durch Blickkoordination“ eine vollständige Beschreibung des

kommunikativen Inventars zur "overlap resolution" kinesischer Simultaneität geliefert haben. Vor diesem Hintergrund stellt sich insbesondere die Frage, inwieweit sich aus Fallkollektionen, die die Variationsbreite zur Bearbeitung konfligierender Simultanaktivitäten vollumfänglich abbilden, zugrundeliegende Regelmechanismen zum Expansionsverzicht (Handlungsretraktion oder -suspension) ableiten lassen. Eine analytische Herausforderung liegt hierbei darin, dass die Auflösung konfligierender Simultanaktivitäten aufgrund der multimodalen Ausdrucksdimensionalität der Körperlichkeit ungleich komplexer geregelt zu sein scheint als der für die selbstselektive Rederechtsübernahme dargelegte Mechanismus "first starter gets the turn" (Sacks et al. 1974). Auf Grundlage einer konkreten Fallalternation der „gestischen Sequentialisierung“ ließen sich erste potentiell relevante Prinzipien bestimmen (vgl. 2.3.7.1), die es in der Folgeforschung auf ihre empirische Validität zu überprüfen gilt, um zu einem präzisen Verständnis der konstitutiven Regelmechanismen des Expansionsverzichts zu gelangen.

Weitere Anschlussmöglichkeiten und Forschungsimplicationen ergeben sich aus den empirischen Befunden zur Teilrezeption kommunikativer Äußerungsgestalten. Unsere Analyseergebnisse haben gezeigt, dass die holistische Rezeption kommunikativer Gestalten kein dringliches Axiom bildet, da sich die fragmentarische Wahrnehmung multimodaler Äußerungsgestalten nicht per se in interaktionalen Problemen manifestiert. Neben den aufgezeigten Stabilitätsmechanismen der sozialen Interaktion (s.o.) ist diese interaktionale Robustheit vor allem auch auf ein zentrales Handlungsprinzip zurückzuführen, das sich systematisch quer durch die Analysen der F2F- und AR-Bedingung zieht: Das primäre Handlungsprinzip Interagierender scheint nicht in der Herstellung geordneter Strukturen zu liegen, sondern darin, ob die jeweiligen Handlungsaktivitäten der Teilnehmer Anschlussmöglichkeiten generieren oder der Bearbeitung von interaktionalen Fortsetzungsproblemen dienlich sind. Die „gelebte Geordnetheit“ (Garfinkel 1967) interaktionaler Prozesse beschreibt in diesem Sinne keine orientierungsrelevante Teilnehmerkategorie, sondern eine interaktionale Auswirkung des Progressionsanspruchs, an dem sich die Interaktanten orientieren (Geordnetheit als „Epiphänomen“).

Die Erkenntnisse zur Teilrezeption kommunikativer Gestalten und zur Primärorientierung am Progressionsanspruch werfen einige empirische Fragen auf, denen sich die multimodal ausgerichtete Konversationsanalyse künftig stellen müssen, da sie grundlegende Implikationen für das methodische Selbstverständnis der Konversationsanalyse erkennen lassen. Obgleich der holistische Ansatz einen konsequenten und notwendigen Perspektivenwechsel von logozentrischen auf multimodale Konzepte der Interaktion beschreibt, stellt sich die Frage, ob die Axiomatisierung des gestaltpsychologischen Paradigmas uneingeschränkt haltbar ist. Inwieweit semiotische Angebotsstrukturen holistisch oder nur teilrezeptiv erfasst werden, hängt einerseits von den Interaktionsbedingungen zum "mutual monitoring" ab, andererseits von der situativen Rezipientenorientierung zum Zeitpunkt der Äußerungsgestalt. Insofern rufen die analytischen Befunde dieser Arbeit dazu auf, soziale Interaktionssituationen nicht per se als "environment of mutual monitoring possibilities"⁸⁴ (Goffman 1964) voranzusetzen, sondern die Koordinierungsmöglichkeiten des "mutual monitoring" unter den jeweiligen kontextuellen Bedingungen der Interaktionssituation genau zu prüfen. Dabei stellt sich die methodische Herausforderung, die situativen Teilnehmerrezeptionen holistischer Äußerungsgestalten aus einer emischen Perspektive analytisch zugänglich zu machen. Betrachten wir die derzeitige technologische Entwicklung von Head-Mounted-Displays, die immer näher an das natürliche Sichtfeld der menschlichen Perzeption reichen, erschließen sich ganz neue Analysemethoden zur situativen Rekonstruktion der Teilnehmerbeobachtungen "from within" (Garfinkel 1967). Sie lassen das analytische Potential zur präziseren Bestimmung der kontextuellen Ursachenfaktoren erkennen, unter denen sich die fragmentarische Rezeption kommunikativer Äußerungsgestalten in divergenten Teilnehmerinterpretationen auswirkt oder unter der Primärorientierung am Progressionsanspruch interaktional folgenlos bleibt. Für die Erforschung des "common ground" (Clark & Brennan 1991) eröffnen sich hier fruchtbare Anschlussoptionen, da sich einerseits die Bedingungen rekonstruieren lassen, unter denen Interaktanten unterschiedliche Repräsentationen multimodal kommunizierter Sachverhalte aufbauen und andererseits analytisch erklärbar wird, warum diese Interpretationsunterschiede in der jeweiligen Situation zum Teil unbemerkt bleiben.

⁸⁴ "I would define a social situation as an environment of mutual monitoring possibilities, anywhere within which an individual will find himself accessible to the naked sense of all others who are 'present', and similarly find them accessible to him." (Goffman 1964: 135)

Literaturverzeichnis

- Akhtar, N., & Gernsbacher, M. A. (2007). Joint Attention and Vocabulary Development: A Critical Look. *Language and Linguistics Compass*, 1(3), 195-207.
- Argyle, M. & Cook, M. (1976). *Gaze and Mutual Gaze*. Cambridge: Cambridge University Press.
- Auer, P. (1986). Kontextualisierung. *Studium Linguistik*, 19, 22-47.
- Auer, P. (1992). Introduction: John Gumperz' Approach to Contextualization. In: P. Auer & A. Di Luzio (Eds.). *The Contextualization of Language* (pp. 1-38). Amsterdam: Benjamins.
- Auer, P. (1996). On the Prosody and Syntax of Turn-taking. In E. Couper-Kuhlen & M. Selting (Eds.). *Prosody and Conversation* (pp. 57-100). Cambridge: Cambridge University Press.
- Auer, P. (2005). Projection in Interaction and Projection in Grammar. *Text*, 25(1), 7-36.
- Auer, P. & Günther, S. (2003). Die Entstehung von Diskursmarkern im Deutschen – Ein Fall von Grammatikalisierung? *Interaction and Linguistic Structure*, 38, 1-30.
- Azuma, R. T. (1997). A survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments*, 6(4), 355-385.
- Bakeman, R. & Adamson, L. B. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Development*, 55, 1278-1289.
- Baron-Cohen, S. (1991): Precursors to a theory of mind: understanding attention in others. In: A. Whiten (Ed.). *Natural theories of mind: evolution, development and simulation of everyday mind-reading* (pp. 233-251). Oxford: Blackwell.
- Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind*. Massachusetts: MIT Press.
- Baron-Cohen, S. (1997). How to build a baby that can read minds: Cognitive mechanisms in mindreading. *The maladapted mind: Classic readings in evolutionary psychopathology*, 207-239.
- Bavelas, J. B., & Chovil, N. (2000). Visible acts of meaning. An integrated message model of language in face-to-face dialogue. *Journal of Language and Social Psychology*, 19, 163-194.
- Bavelas, J. B., Coates, L. & Johnson, T. (2002). Listener responses as a collaborative process: The role of gaze. *Journal of Communication*, 52, 566-580.
- Bergmann, J. R. (1981). Ethnomethodologische Konversationsanalyse. In: P. Schröder & H. Steger (Eds.). *Dialogforschung. Jahrbuch 1980 des Instituts für deutsche Sprache* (pp. 9-51). Düsseldorf: Schwann.
- Bergmann, J. R. (1985). Flüchtigkeit und methodische Fixierung sozialer Wirklichkeit. Aufzeichnungen als Daten der interpretativen Soziologie. In: W. Bonß & H. Hartmann (Eds.). *Entzauberte Wissenschaft: Zur Relativität und Geltung soziologischer Forschung* (pp. 299-320). Göttingen: Schwarz.
- Bergmann, J. R. (1994). Ethnomethodologische Konversationsanalyse. In: G. Fritz & F. Hundsnurscher (Eds.). *Handbuch der Dialoganalyse* (pp. 3-16). Tübingen: Max Niemeyer Verlag.
- Bergmann, J. R. (2004). Harold Garfinkel and Harvey Sacks. In: U. Flick, E. von Kardoff & I. Steinke (Eds.). *A Companion to Qualitative Research* (pp. 29-34). London: Sage.
- Bigelow, A. E. (2003). The development of joint attention in blind infants. *Development and Psychopathology*, 15, 259-275.
- Brennan, S. E. & Clark, H. H. (1996). Conceptual Pacts and Lexical Choice in Conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(6), 1482-1493.
- Brinton, L. J. (1996). *Pragmatik Markers in English: Grammaticalization and Discourse Functions*. Berlin: Mouton de Gruyter.
- Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, 8(6), 535-543.
- Bruner, J. (1985). The role of interaction formats in language acquisition. In: J. P. Forgas (Ed.). *Language and Situations*. New York: Springer.

- Brüning, B. - A., Schnier, C., Pitsch, K., & Wachsmuth, S. (2011). Automatic detection of motion sequences for motion analysis. In D. Heylen, P. Paggio, & M. Kipp (Eds.). *Proceedings of the ICMI Workshop on Multimodal Corpora for Machine Learning: Taking Stock and Roadmapping the Future*.
- Brüning, B. - A., Schnier, C., Pitsch, K., & Wachsmuth, S. (2012). PAMOCAT: Automatic retrieval of specified postures. In N. Calzolari, K. Choukri, T. Declerck, M. Uğur Doğan, B. Maegaard, J. Mariani, J. Odijk, et al. (Eds.). *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC 2012)* (pp. 4143–4148).
- Brünner, G. (2005). *Kommunikation in institutionellen Lehr-Lern-Prozessen. Diskursanalytische Untersuchungen zu Instruktionen in der betrieblichen Ausbildung*. Radolfzell: Verlag für Gesprächsforschung.
- Butterworth, G. (1998). What is special about pointing in babies? In: F. Simion & G. Butterworth (Eds.). *The development of sensory, motor and cognitive capacities in early infancy: From perception to cognition* (pp. 171-190). Hove: Psychology Press.
- Bühler, K. (1934/1982). *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Stuttgart/New York: Fischer.
- Camaioni, L. (1992): Mind knowledge in infancy: the emergence of intentional communication. *Early Development and Parenting*, 1(1), 15-22.
- Campos, J. J. & Stenberg, C. R. (1981). Perception, appraisal, and emotion: The onset of social referencing. In: M. E Lamb & L. R Sherrod (Eds.). *Infant Social Cognition: Empirical and Social Considerations* (pp. 273–314). NJ: Erlbaum.
- Charman, T., Swettenham, J., Baron-Cohen, S., Cox, A., Baird, G. & Drew, A. (1997). Infants with autism: an investigation of empathy, pretend play, joint attention, and imitation. *Developmental Psychology*, 33(5), 781-789.
- Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A. & A. Drew (2000): Testing joint attention, imitation, and play as infancy precursors to language and theory of mind. *Cognitive Development*, 15(4), 481-498.
- Clancy, P. M., Thompson, S. A., Ryoko Suzuki & Hongyin Tao (1996). The conversational use of reactive tokens in English, Japanese, and Mandarin. *Journal of pragmatics*, 26, 355-387.
- Clark, H. H. (1996). *Using Language*. Cambridge: Cambridge University Press.
- Clark, H. H., Schreuder, R. & Buttrick, S. (1983). Common Ground and the Understanding of Demonstrative Reference. *Journal of verbal learning and verbal behavior*, 22(2), 245-258.
- Clark, H. H. & Brennan, S. A. (1991). Grounding in Communication. In L. B. Resnick, J. M. Levine & S. D. Teasley (Eds.). *Perspectives on socially shared cognition* (pp. 127-149). Washington DC: APA Books.
- Clark, H. H. & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, 50, 62-81.
- Clark, H. H. (2005). Coordinating with each other in a material world. *Discourse Studies*, 7(4-5), 507-525.
- Couper-Kuhlen, B. (1996). Intonation and Clause Combining in Discourse: The Case of Because. *Pragmatics* 6(3), 389–426.
- Couper-Kuhlen, E. (2001). Interactional Prosody: High Onsets in Reason-for-the-Call Turns. *Language in Society*, 30, 29–53.
- Couper-Kuhlen, E. (2004). Prosody and Sequence Organization: The Case of New Beginnings. In E. Couper-Kuhlen & C. E. Ford (Eds.). *Sound Patterns in Interaction* (pp. 335–376). Amsterdam: Benjamins.
- Dausendschön-Gay, U. (2000). On-line-Hilfe für den Hörer. Verfahren zur Orientierung der Interpretationsstichtigkeit. In: B. Wehr & H. Thomaßen (Eds.). *Diskursanalyse. Untersuchungen zum gesprochenen Französisch* (pp. 17-55). Frankfurt/M: Lang.
- Dausendschön-Gay, U. & Krafft, U. (2002). Text und Körpergesten. Beobachtungen zur holistischen Organisation der Kommunikation. *Psychotherapie und Sozialwissenschaft*, 4(1), 30-60.
- Dausendschön-Gay, U. & Krafft, U. (2007). Prozesse interpersoneller Koordination. In R. Schmitt (Ed.). *Koordination. Analysen zur multimodalen Interaktion* (pp. 167-194). Tübingen: Narr.
- Dausendschön-Gay, U. & Krafft, U. (2009). Preparing Next Actions in Routine Activities. *Discourse Processes*, 46, 249-268.

- Deppermann, A. (2001). Aspekte einer konversationsanalytischen Untersuchung von Wortsemantik. In: H. Gruber & F. Menz (Eds.): *Interdisziplinarität in der Angewandten Sprachwissenschaft* (pp. 57-77). Frankfurt/M.: Peter Lang.
- Deppermann, A. (2010). Konversationsanalyse und diskursive Psychologie. In: G. Mey & K. Mruck (Eds.). *Handbuch Qualitative Forschung in der Psychologie* (pp. 643-661). Wiesbaden: VS Verlag für Sozialwissenschaften.
- 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.
- Dierker, A. (2012). *Computer-aided investigation of interaction mediated by an AR-enabled wearable interface*. Dissertation, Bielefeld University.
- Dierker, A., Bovermann, T., Hanheide, M., Hermann, T. & Sagerer, G. (2009). A multimodal augmented reality system for alignment research. *13th International Conf. on HCI* (pp. 422-426). NJ, Heidelberg: Springer.
- Egbert, M. (1997). Schisming. The collaborative transformation from a single conversation to multiple conversations. *Research on Language and Social Interaction*, 30(1), 1-51.
- Egbert, M. (2009). *Der Reparatur-Mechanismus in deutschen Gesprächen*. Mannheim: Verlag für Gesprächsforschung.
- Emery, N. J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews*, 24, 581-604.
- Fahrländer, F. (2013). Zur Syntax und Semantik der konzessiven Satzkonnectoren *obwohl* und *trotzdem*. In: H. Blühdorn, M. Elstermann & A. Klosa (Eds.): *Online publizierte Arbeiten zur Linguistik (OPAL)* (pp. 1-23), 2013, 1, Mannheim: Insitut für Deutsche Sprache.
- Fiehler, R., Kindt, W. & Schnieders, G. (1999). Kommunikationsprobleme in Reklamationsgesprächen. In: G. Brüner, R. Fiehler & W. Kindt (Eds.). *Angewandte Diskursforschung. Band 1: Grundlagen und Beispielanalysen* (pp. 120-154). Wiesbaden: Opladen.
- Ford, C.E., Fox, B.A. & Thompson, S.A. (1996). Practices in the Construction of Turns: The "TCU" revisited. *Pragmatics*, 6, 427-454.
- Ford, C.E. (2004). Contingency and units in interaction. *Discourse Studies*, 6(1), 27-52.
- Fox, B. A., Hayashi, M. & Jasperson, R. (1996). Resources and repair: A cross-linguistic study of syntax and repair. In: E. Ochs, E. A. Schegloff & S. A. Thompson (Eds.). *Interaction and Grammar* (pp. 185-237). Cambridge: Cambridge University Press.
- Fraser, M., Glover, T., Vaghi, I., Benford, S., Greenhalgh, C., Hindmarsh, J., & Heath, C. (2000). Revealing the realities of collaborative virtual reality. *Proceedings of the third international conference on Collaborative virtual environments* (pp. 29-37). New York: ACM.
- Frey, S. (1984). *Die nonverbale Kommunikation*. Stuttgart: SEL-Stiftung für technische und wirtschaftliche Kommunikationsforschung im Stifterverband für die Deutsche Wirtschaft, Stiftungsreihe 1.
- Fricke, E. (2007). *Origo, Geste und Raum. Lokaldeixis im Deutschen*. Berlin/New York: de Gruyter.
- Gardner, R. (2001). *When Listeners Talk: Response Tokens and Listener Stance*. Amsterdam: John Benjamins.
- Garfinkel, H. (1967). *Studies in Ethnomethodology*. Englewood Cliffs, N.J.: Prentice Hall.
- Gibson, J. J. (1982). *Wahrnehmung und Umwelt. Der ökologische Ansatz der visuellen Wahrnehmung*. München: Urban & Schwarzenberg.
- Goffman, E. (1963): *Behavior in public places. Notes on the social organization of gatherings*. New York: Free Press of Glencoe.
- Goffman, E. (1964). The neglected situation. In *American Anthropologist*, 66(6), part 2, 133-136.
- Goffman, E. (1974). *Frame Analysis: An Essay on the Organization of Experience*. Cambridge, Mass.: Harvard University Press.
- Goffman, E. (1981). *Forms of talk*. Philadelphia: University of Pennsylvania Press.
- Goffman, E. (1983a). The Interaction Order. *American Sociological Review*, 48, 1-17.
- Goffman, E. (1983b). Felicity's condition. *American Journal of Sociology*, 89(1), 1-53.

- Goodwin, C. (1980). Restarts, Pauses, and the Achievement of a State of Mutual Gaze at Turn-Beginning. *Sociological Inquiry*, 50(3-4), 272-302.
- Goodwin, C. (1981). *Conversational Organization: Interaction between Speakers and Hearers*. New York: Academic Press.
- Goodwin, C. (1986). Gestures as a resource for the organization of mutual orientation. *Semiotica*, 62, 29-49.
- Goodwin, C. (1995). Co-Constructing Meaning in Conversation with an Aphasic Man. *Research on Language and Social Interaction*, 28(3), 233-260.
- Goodwin, C. (2000a). Action and embodiment within situated human interaction. *Journal of Pragmatics*, 32, 1489-1522.
- Goodwin, C. (2000b). Pointing and the Collaborative Construction of Meaning in Aphasia. *Texas Linguistic Forum*, 43, 67-76. Proceedings of the seventh annual Symposium About Language and Society, Austin.
- Goodwin, C. (2002). Time in action. *Current Anthropology*, 43, 19-35.
- Goodwin, C. (2003a). The Body in Action. In J. Coupland & R. Gwyn (Eds.). *Discourse, the Body, and Identity* (pp. 19-42). Houndsmill Hampshire, New York: Palgrave/Macmillan.
- Goodwin, C. (2003b). Pointing as Situated Practice. In: S. Kita (Ed.) *Pointing: Where Language, Culture, and Cognition Meet* (pp. 217-241). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Goodwin, C. (2003c). Conversational Frameworks for the Accomplishment of Meaning in Aphasia. In C. Goodwin (Ed.). *Conversation and Brain Damage* (pp. 90-116). Oxford: Oxford University Press.
- Goodwin, C. (2006). Retrospective and prospective orientation in the construction of argumentative moves. *Text & Talk – An Interdisciplinary Journal of Language, Discourse Communication Studies*, 26(4-5), 443-461.
- Goodwin, C. (2007). Participation, stance and affect in the organization of activities. *Discourse & Society*, 18(1), 53-73.
- Goodwin, C. (2009). Video and the analysis of embodied human interaction. In: U. T. Kismmann (Ed.). *Video Interaction Analysis* (pp. 21-40). Frankfurt am Main: Peter Lang.
- Goodwin, C. (2010). Constructing Meaning Through Prosody in Aphasia. In: D. Barth-Weingarten, E. Reber & M. Selting (Eds.). *Prosody in Interaction* (pp. 373-394). Amsterdam: John Benjamins.
- Goodwin, C. (2011). Contextures of action. In: J. Streeck, C. Goodwin & C.D. Lebaron (Eds.). *Embodied Interaction: Language and Body in the Material World* (pp. 182-193). Cambridge: Cambridge University Press.
- Goodwin, C. (2013). The co-operative, transformative organization of human action and knowledge. *Journal of Pragmatics*, 46, 8-23.
- Goodwin, C. & Heritage, J. (1990). Conversation analysis. *Annual review of anthropology*, 19, 283-307.
- Goodwin, D., Pope, C., Mort, M. and Smith, A. (2005): Access, boundaries and their effects: Legitimate participation in anaesthesia. *Sociology of Health and Illness*, 27(6), 855-871.
- Goodwin, M. H. (1980). Processes of mutual monitoring implicated in the production of description sequences. *Sociological Inquiry*, 50, 303-317.
- Goodwin, M. H. & Goodwin, C. (2004). Participation. In: A. Duranti (Ed.). *A Companion to Linguistic Anthropology* (pp. 222-244). Oxford: Blackwell.
- Grice, P. (1975). Logic and conversation. In: P. Cole & J. L. Morgan (Eds.). *Syntax and Semantics, Vol 3: Speech acts* (pp. 41-58). New York: Academic Press.
- Grimm, J. & Grimm, W. (1889 [1984]): *Deutsches Wörterbuch*. Nachdruck der Erstausgabe. Leipzig: Hirzel.
- Gullberg, M. & Holmqvist, K. (2002). Visual Attention towards Gestures in Face-to-Face Interaction vs. on Screen. In: I. Wachsmuth & T. Sowa (Eds.). *Gesture and Sign Languages in Human-Computer Interaction: International Gesture Workshop, GW 2001, London* (pp. 206-214). Berlin: Springer.
- Gumperz, J. J. (1992). Contextualization revisited. In P. Auer & A. Di Luzio (Eds.). *The Contextualization of Language* (pp. 39-53). Amsterdam: Benjamins.
- Günthner, S. (2000). “wobei (.) es hat alles immer zwei seiten.” Zur Verwendung von *wobei* im gesprochenen Deutsch. *Interaction and Linguistic Structures (InLiSt)*, 18, 1-37.
- Hanks, W. F. (2005). Explorations in the Deictic Field. *Current Anthropology*, 46(2), 191-220.

- 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). Berlin/New York: de Gruyter
- Hayashi, M. (1999). Where Grammar and Interaction Meet: A Study of Co-Participant Completion in Japanese Conversation. *Human Studies*, 22, 475-499.
- Heath, C. and Luff, P. (1991). Disembodied conduct: communication through video in multimedia office environment. *Proceedings of CHI 1991*, 99-103, New Orleans: ACM Press.
- Heath, C. & Luff, P. (1992a). Collaboration and control: Crisis management and multimedia technology in London Underground line control rooms. *Computer-Supported Cooperative Work (CSCW)*, 1, 69-94.
- Heath, C., & Luff, P. (1992b). Media space and communicative asymmetries: Preliminary observations of video mediated interaction. *Human-Computer Interaction*, 7, 315-346.
- Heath, C., Jirotko, M., Luff, P. & Hindmarsh, J. (1995). The Individual and the Collaborative: the Interactional Organisation of Trading in a City Dealing Room. *Journal of Computer Supported Cooperative Work*, 3(1), 147-165.
- Heath, C., Svensson, M. S., Hindmarsh, J. & Luff, P. (2002). Configuring Awareness. *Computer-Supported Cooperative Work (CSCW)*, 11, 317-347.
- Heath, C., Luff, P. & M. S. Svensson (2003). Technology and medical practice. *Sociology of Health & Illness*, 25, 75-96.
- Heldner, M. & Edlund, J. (2010). Pauses, gaps and overlaps in conversations. *Journal of Phonetics*, 38, 555-568.
- Heritage, J. (1984). *Garfinkel and ethnomethodology*. Cambridge: Polity Press.
- Heritage, J. (1995). Conversation analysis: Methodological aspects. In U. M. Quasthoff (Ed): *Aspects of Oral Communication* (pp. 391-418). Berlin: de Gruyter.
- Herrmann, T. (1994). Psychologie ohne "Bedeutung"? Zur Wort-Konzept-Relation in der Psychologie. Heidelberg/Mannheim: *Arbeiten aus dem SFB 245 „Sprache und Situation“*, Bericht Nr. 75.
- Hindmarsh, J., Fraser, M., Heath, C., Benford, S. & Greenhalgh, C. (1998). Fragmented interaction: Establishing mutual orientation in virtual environments. *Proceedings of the Conference on Computer Supported Cooperative Work (CSCW 1998)* (pp. 217-226). New York: ACM.
- Hindmarsh, J. & Heath, C. (2000a). Embodied reference: A study of deixis in workplace interaction. *Journal of Pragmatics*, 32, 1855-1878.
- Hindmarsh, J., Fraser, M., Heath, C., Benford, S. & Greenhalgh, C. (2000b). Object-focused interaction in collaborative virtual environment. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 7(4), 477-509.
- Hindmarsh, J. & Pilnick, A. (2007). Knowing Bodies at Work: Embodiment and Ephemeral Teamwork in Anaesthesia. *Organization studies*, 28(9), 1395-1416.
- Hobson, P. (2002): *The Cradle of Thought. Exploring the Origins of Thinking*. London: MacMillan.
- Holling, C. (1973). Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics*, 4, (1-23).
- Holly, W. (2008). Rezension zu: Reinhold Schmitt (Hrsg.): Koordination. Analysen zur multimodalen Interaktion. Tübingen: Gunter Narr. In: *Zeitschrift für germanistische Linguistik*, 36, 277-283.
- Houtkoop-Steenstra, H. (1990). Accounting for proposals. *Journal of Pragmatics*, 14, 111-124.
- Houtkoop, H. & Mazeland, H. (1985). Turns and Discourse Units in Everyday Conversation. *Journal of Pragmatics*, 9, 595-619.
- Hutchby, I. & Wooffitt, R. (2008). *Conversation Analysis (2nd Edition)*. Cambridge: Polity Press.
- Jakobson, R. (1971). Shifters, verbal categories, and the Russian verb. In *Selected Writings II*, 130-147. The Hague: Mouton.
- Jefferson, G. (1978). Sequential Aspects of Storytelling in Conversation. In J. Schenkein (Ed.). *Studies in the Organization of Conversational Interaction* (pp. 219-248). New York: Academic Press.
- Jefferson, G. (1983). *Notes on some orderlinesses of overlap onset*. Tilburg Pap. Lang. Lit. 28 (whole vol.).

- Jefferson, G. (2004). A sketch of some orderly aspects of overlap in natural conversation. In: G. H. Lerner (Ed.). *Conversation Analysis. Studies from the First Generation* (pp. 43-59). Amsterdam/Philadelphia: John Benjamins.
- Jucker, A. H. & Ziv, Y. (1998). Discourse Markers: Introduction. In A. H. Jucker & Y. Ziv (Eds.). *Discourse Markers: Descriptions and Theory* (pp. 1-12). Amsterdam: John Benjamins.
- Kallmeyer, W. & Schütze, F. (1976). Konversationsanalyse. *Studium Linguistik*, 1, 1-28.
- Kallmeyer, W. & Schmitt, R. (1996). Forcieren oder: Die verschärfte Gangart. Zur Analyse von Kooperationsformen im Gespräch. In: W. Kallmeyer (Ed.). *Gesprächsrhetorik. Rhetorische Verfahren im Gesprächsprozess* (pp. 20-118). Tübingen: Gunter Narr.
- Kaplan, F. & Hafner, V. V. (2006). The challenges of joint attention. *Interaction Studies*, 7(2), 135-169.
- Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica*, 26, 22-63.
- Kendon, A. (1978). Looking in conversations and the regulation of turns at talk: A comment on the papers of G. Beattie and D. R. Rutter et al. *British Journal of Social & Clinical Psychology*, 17(1), 23-24.
- Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (Ed.). *The Relationship of Verbal and Nonverbal Communication* (pp. 207-227). The Hague: Mouton.
- Keseling, G. (1989). Die Partikel *ÄH*. Ein paraverbales Element im Sprachsystem? In H. Weydt (Ed.). *Sprechen mit Partikeln* (pp. 575-591). Berlin/New York: Walter de Gruyter.
- Kidwell, M. & Zimmerman, D. (2007). Joint Attention as action. *Journal of Pragmatics*, 39(3), 592-611.
- Kirk, D., Crabtree, A. & Rodden, T. (2005). Ways of the Hand. *Proceedings of the Ninth European Conference on Computer Supported Cooperative Work (ECSCW 2005)* (pp. 1-21). New York: Springer.
- Klein, W. (1990). Überall und nirgendwo. Subjektive und objektive Momente in der Raumreferenz. *Zeitschrift für Linguistik und Literaturwissenschaft*, 78, 9-42.
- Klein, W. (1991). Raumausdrücke. *Linguistische Berichte*, 132, 77-114.
- Klein, W. (1994). Keine Känguruhs zur Linken. Über die Variabilität von Raumvorstellungen und ihren Ausdruck in der Sprache. In H.-J. Kornadt, J. Grabowski & R. Mangold-Allwinn (Eds.). *Sprache und Kognition. Perspektiven moderner Sprachpsychologie* (pp. 163-182). Heidelberg: Spektrum.
- Klein, W. & Jungbluth, K. (2002). Einleitung. *Zeitschrift für Literaturwissenschaft und Linguistik*, 125, 5-9.
- Knoblauch, H. & Heath, C. (1999). Technologie, Interaktion und Organisation: Die Workplace Studies. *Swiss Journal of Sociology*, 25(2), 163-181.
- Kollenberg, T., Neumann, A., Schneider, D., Tews, T. - K., Hermann, T., Ritter, H., Dierker, A., et al. (2010). Visual search in the (un)real world: How head-mounted displays affect eye movements, head movements and target detection. *Proceedings of the 2013 Symposium on Eye Tracking Research & Applications* (pp. 121-124). ACM.
- Koschmann, T., LeBaron, C., Goodwin, C., and Feltovich, P. (2006): The mystery of the missing referent: objects, procedures, and the problem of the instruction follower. *Proceedings of the 2006 Conference on Computer-Supported Cooperative Work (CSCW 2006)* (pp. 373-382).
- Kranstedt, A., Lücking, A., Pfeiffer, T., Rieser, H. & Wachsmuth, I. (2006). Deictic Object Reference in Task-Oriented Dialogue. In: G. Rickheit & I. Wachsmuth (Eds.). *Situated Communication* (pp. 155-208). Berlin: Mouton de Gruyter.
- Krummheuer, A. (2010). *Interaktion mit virtuellen Agenten? Zur Aneignung eines ungewohnten Artefakts*. Stuttgart: Lucius & Lucius.
- Kurtic, E., Brown, G. J. & Wells, B. (2013). Resources for turn competition in overlapping talk. *Speech Communication*, 55, 1-23.
- Kuzuoka, H., Yamazaki, K., Yamazaki, A., Kosaka, J., Suga, Y. and Heath, C. (2004): Dual ecologies of robot as communication media: thoughts on coordinating orientations and projectability. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp.183-190). ACM.
- Labov, W. (1972). The Transformation of Experience in Narrative Syntax. In W. Labov (Ed.). *Language in the Inner City: Studies in the Black English Vernacular* (pp. 354-396). Philadelphia: University of Pennsylvania Press.

- Langton, S. R., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in cognitive sciences*, 4(2), 50-59.
- Lee, K., Eskritt, M., Symons, L. A. & Muir, D. (1998). Children's use of triadic eye gaze information for "mind reading". *Developmental psychology*, 34(3), 525-539.
- Lerner, G. H. (1991). On the syntax of sentences in progress. *Language in Society*, 20, 441-458.
- Lerner, G. H. (1996). On the 'semi-permeable' character of grammatical units in conversation: Conditional entry into the turn space of another speaker. In: E. Ochs, E. A. Schegloff & S. A. Thompson (Eds.). *Interaction and Grammar* (pp. 238-276). Cambridge: Cambridge University Press.
- Levinson, S. C. (1996). Language and Space. *Annual Review of Anthropology*, 25, 353-382.
- Lerner, G. H. (2003). Selecting Next Speaker: The context-sensitive Operation of a context-free organization. *Language in Society*, 32(2), 177-201.
- Luckmann, T. (1986). Grundformen der gesellschaftlichen Vermittlung des Wissens: Kommunikative Gattungen. In F. Neidhardt et al. (Eds.). *Kultur und Gesellschaft. Kölner Zeitschrift für Soziologie und Sozialpsychologie* (pp. 191-211). Sonderheft 27. Opladen.
- Luff, P., Heath, C., Kuzuoka, H., Hindmarsh, J., Yamazaki, K., & Oyama, S. (2003). Fractured ecologies: Creating environments for collaboration. *Human-Computer Interaction*, 18, 51-84.
- Luhmann, N. (1984). *Soziale Systeme. Grundriß einer allgemeinen Theorie*. Frankfurt: Suhrkamp.
- Maynard, D. W. & Clayman, S. E. (2003). Ethnomethodology and conversation analysis. In: L. T. Reynolds & N. J. Herman-Kinney (Eds.). *Handbook of symbolic interactionism* (pp. 173-202). Walnut Creek, CA: Rowman Altamira Press
- McNeill, D. (1992). *Hand and Mind. What Gestures Reveal about Thought*. Chicago: University of Chicago Press.
- Millis, K. K., Goding, J. M. & Barker, G. (1995). Causal connectives increase inference generation. *Discourse Processes*, 20(1), 29-49.
- Mondada, L. (2007a). Operating together through videoconference : Members' procedures for accomplishing a common space of action. In S. Hester & D. Francis (Eds.). *Orders of Ordinary Action* (pp. 51-67). Aldershot: Ashgate.
- Mondada, L. (2007b). Multimodal Resources for turn-taking: pointing and the emergence of possible next speakers. *Discourse Studies*, 9(2), 194-225.
- Mondada, L. (2007c). Interaktionsraum und Koordinierung. In: R. Schmitt (Ed.). *Koordination. Analysen zur multimodalen Interaktion* (pp. 55-94). Tübingen: Narr.
- Mondada, L. (2009). Emergent focused interactions in public places: A systematic analysis of the multimodal achievement of a common interactional space. *Journal of Pragmatics*, 41(10), 1977-1997.
- Mondada, L. (2011). The Organization of Concurrent Courses of Action in Surgical Demonstrations. In: J. Streeck, C. Goodwin & C.D. Lebaron (Eds.). *Embodied Interaction: Language and Body in the Material World* (pp. 207-225). Cambridge: Cambridge University Press.
- Mondada, L. (2012). Deixis: An integrated interactional multimodal analysis. In: P. Bergmann, J. Brenning, M. Pfeiffer & E. Reber (Eds.). *Prosody and embodiment in interactional grammar* (pp. 173-206). Berlin: de Gruyter.
- Mondada, L. & Oloff, F. (2011): Gesture in overlap. The situated establishment of speakership. In G. Stam & M. Ishino (Eds.). *Integrating Gestures: The interdisciplinary nature of gesture* (pp. 321-337). Amsterdam: John Benjamins.
- Moreira, T. (2004): Coordination and embodiment in the operating room. *Body & Society*, 10(1), 109-129.
- Müller, C. (2004). The palm-up-open-hand. A case of a gesture family? In: C. Müller & R. Posner (Eds.). *The semantics and pragmatics of everyday gestures* (pp. 233-256). Berlin: Weidler Verlag.
- Müllerová, O. & Nekvapil, J. (1990). Zur Untersuchung von Pausen in tschechischen gesprochenen Texten. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung*, 43(5), 701-709.
- Nardi, B. A., Schwartz, H., Kuchinsky, A., Lechner, R., Whitaker, S. and Scabassi, R. (1995). Turning away from talking heads: the use of video-as-data in neurosurgery. In S. J. Emmott (Ed.). *Information Superhighways: Multimedia Users and Futures*. London and San Diego: Academic Press.

- Neumann, A. (2011). *Design and implementation of multi-modal AR-based interaction for cooperative planning tasks*. Master Thesis, Bielefeld University.
- Neumann, A., & Hermann, T. (2013). A Verbal Interaction Measure Using Acoustic Signal Correlation for Dyadic Cooperation Support. In: A. van Berlo, K. Hallenborg, J. M. C. Rodriguez, D. I. Tapia, & P. Novais (Eds.). *Advances in Intelligent Systems and Computing: Vol. 219. Ambient Intelligence - Software and Applications* (pp. 71–78). Switzerland: Springer International Publishing.
- Neumann, A., Schnier, C., Hermann, T. & Pitsch, K. (2013). Interaction Analysis and Joint Attention Tracking in Augmented Reality. *Proceedings of the 15th ACM International Conference on Multimodal Interaction (ICMI)*, 165-172.
- Olbertz-Siitonen, M. (2007). Ähm vs. niinku – Verzögerungssignale in deutschen und finnischen Diskussionen. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 12(2), 1-22.
- Oloff, F. (2012): Withdrawal from turns in overlap and participation. In: P. Bergmann, J. Brenning, M. Pfeiffer & E. Reber (Eds.). *Prosody and Embodiment in Interactional Grammar* (pp. 207-236). Berlin/Boston: De Gruyter.
- Oloff, F. (2013): Embodied withdrawal after overlap resolution. *Journal of Pragmatics*, 46(1), 139-156.
- Parkhurst, D., Law, K. & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, 42, 107-123.
- Pitsch, K. (2006). *Sprache, Körper, Intermediäre Objekte: Zur Multimodalität der Interaktion im bilingualen Geschichtsunterricht*. Doctoral Thesis, Bielefeld University.
- Pitsch, K., Brüning, B. - A., Schnier, C., Dierker, H., & Wachsmuth, S. (2010). Linking Conversation Analysis and Motion Capturing: How to robustly track multiple participants? In M. Kipp, J.- C. Martin, P. Paggio, & D. Heylen (Eds.). *Proceedings of the LREC Workshop on Multimodal Corpora: Advances in Capturing, Coding and Analyzing Multimodality (MMC 2010)* (pp. 63–69).
- Pitsch, K., Neumann, A., Schnier, C. & Hermann, T. (2013). Augmented reality as a tool for linguistic research: Intercepting and manipulating multimodal interaction. *Multimodal Corpora: Beyond audio and vision* (IVA 2013 workshop).
- Pitsch, K., Vollmer, A.-L., Rohlfing, K., Fritsch, J. & Wrede, B. (2014). Tutoring in adult-child interaction: On the loop of the tutor's action modification and the recipient's gaze. *Interaction Studies*, 15(1), 55-98.
- Richardson, D. C. & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science*, 29, 1045-1060.
- Rieser, H. (2004). Pointing in dialogue. *Proceedings of the 8th Workshop on the Semantics and Pragmatics of Dialogue* (pp. 93-101).
- Rossano, F. (2005). *When It's Over is it Really Over? On the Effects of Sustained Gaze vs. Gaze Withdrawal at Sequence Possible Completion*. Paper presented at the International Pragmatic Association, Riva del Garda, July.
- Ryave, A. L. (1978). On the achievement of a series of stories. In J. Schenkein (Ed.). *Studies in the organization of conversational interaction* (pp. 113-132). New York: Academic Press.
- Sacks, H. (1984). Notes on methodology. In: J. M. Atkinson & J. Heritage (Eds.). *Structures of social action. Studies in conversation analysis* (pp. 21-27). Cambridge: Cambridge University Press.
- Sacks, H. (1992). *Lectures on Conversation*. Oxford: Blackwell.
- Sacks, H. (1974). An Analysis of a Joke's Telling in Conversation. In R. Bauman & J. Sherzer (Eds.). *Explorations in the Ethnography of Speaking* (pp. 337-353). Cambridge: Cambridge University Press.
- Sacks, H. & Schegloff, E. A. (2002). Home position. *Gesture*, 2(2), 133-146.
- Sacks, H., Schegloff, E. A. and Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, 50, 696-735.
- Scaife, M. & Bruner, J.S. (1975). The capacity for joint visual attention in the infant. *Nature*, 5489, 265– 266.
- Schegloff, E. A. (1968). Sequencing in conversational openings. *American Anthropologist*, 70, 1075-1095.
- Schegloff, E. A. (1980). Preliminaries to preliminaries: "Can I ask you a question?". *Sociological Inquiry*, 50(3-4), 104-152.

- 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.). *Georgetown University Roundtable on languages and linguistics*. Washington, D.C.: Georgetown University Press.
- Schegloff, E. A. (1984). On Some Gestures' Relation to Talk. In J. M. Atkinson and J. Heritage (Eds.). *Structures of Social Action* (pp. 266–98). Cambridge: Cambridge University Press.
- Schegloff, E. A. (1993). Reflections on quantification in the study of conversation. *Research on Language and Social Interaction*, 26(1), 99-128.
- Schegloff, E. A. (1996a). Turn Organization: One Intersection of Grammar and Interaction. In E. Ochs, E. A. Schegloff & S. A. Thompson (Eds.). *Interaction and Grammar* (pp. 52-133). Cambridge: Cambridge University Press.
- Schegloff, E. A. (1996b). Issues of Relevance for Discourse Analysis: Contingency in Action, Interaction and Co-Participant Context. In E. H. Hovy & D. R. Scott (Eds.). *Computational and Conversational Discourse: Burning Issues – An Interdisciplinary Account* (pp. 3-35). New York: Springer.
- Schegloff, E. A. (2000). Overlapping talk and the organization of turn-taking for conversation. *Language in Society*, 29, 1-63.
- Schegloff, E. A. (2001). Accounts of Conduct in Interaction. Interruption, Overlap, and Turn-Taking. In: J. H. Turner (Ed). *Handbook of Sociological Theory* (pp. 287-321). New York: Kluwer Academic.
- Schegloff, E. A. & Sacks, H. (1973). Opening Up Closings. *Semiotica*, 8, 289-327.
- Schegloff, E. A., Jefferson, G. & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, 53(2), 361-382.
- Schenkein, J. (1978). Sketch of an analytic mentality for the study of conversational interaction. In: J. Schenkein (Ed.). *Studies in the Organization of Conversational Interaction* (pp. 1-6). New York: Academic Press.
- Schmitt, R. (2004). Die Gesprächspause. Verbale ‚Auszeiten‘ aus multimodaler Perspektive. In: *Deutsche Sprache*, 32(1), 56-84.
- Schmitt, R. (2005). Zur multimodalen Struktur von *turn-taking*. *Gesprächsforschung*, 6, 17-61.
- Schmitt, R. & Deppermann, A. (2010). Die multimodale Konstitution eines imaginären Raumes als interaktive Problemlösung. In: A. Deppermann & A. Linke (Eds.). *Sprache intermedial. Stimme und Schrift, Bild und Ton* (pp. 199-241). Berlin/New York: de Gruyter.
- Schneider, W. L. (2005). *Grundlagen der soziologischen Theorie. Band 2: Garfinkel – RC – Habermas – Luhmann* (2. Auflage). Wiesbaden: VS Verlag.
- Schnier, C., Pitsch, K., Dierker, A., and Hermann, T. (2011a). Collaboration in Augmented Reality: How to establish coordination and joint attention? *12th ECSCW*, 405-416.
- Schnier, C., Pitsch, K., Dierker, A., and Hermann, T. (2011b). Adaptability of Communicative Resources in AR-based Cooperation. *Proceedings of the 2nd Workshop on Gesture and Speech in Interaction (GESPIN)*, Bielefeld.
- Schober, M. F. (1993). Spatial perspective-taking in conversation. *Cognition*, 47, 1–24.
- Schubert, C. (2006). Videographie im OP: Wie Videotechnik für technografische Studien im OP genutzt werden kann. In W. Rammert & C. Schubert (Eds.). *Technografie. Zur Mikrosoziologie der Technik* (pp. 223-248). Frankfurt/M. und New York: Campus Verlag.
- Schütz, A. (1960). *Der sinnhafte Aufbau der sozialen Welt. Eine Einleitung in die verstehende Soziologie*. Wien: Springer-Verlag.
- Scott, M. & Lyman, S. (1968). Accounts. *American Sociological Review*, 33, 46-62.
- Selting, M. (2000). The Construction of Units in Conversational Talk. *Language and Society*, 29, 477-517.
- Selting, M., Auer, P., Barden, B., Bergmann, J., Couper-Kuhlen, E., Günthner, S., Meier, C., Quasthoff, U., Schoblin, P. & Uhmann, S. (1998). Gesprächsanalytisches Transkriptionssystem (GAT). *Linguistische Berichte*, 173, 91-122.

- Selting, M., Auer, P., Barth-Weingarten, D., Bergmann, J., Bergmann, P., Brikner, K., Couper-Kuhlen, E., Deppermann, A., Gilles, P., Günthner, S., Hartung, M., Kern, F., Mertzlufft, C., Meyer, C., Morek, M., Oberzaucher, F., Peters, J., Quasthoff, U., Schütte, W., Stukenbrock, A. & Uhmann, S. (2009). Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion*, 10, 353-402.
- Sidnell, J. (2001). Conversational turn-taking in a Caribbean English creole. *Journal of Pragmatics*, 33(8), 1263-1290.
- Spiegel, C. & Spranz-Fogasy, T. (2001). Zur Methodologie der Handlungsstrukturanalyse von Gesprächen. In: Z. Iványi & A. Kertész (Eds.). *Gesprächsforschung. Tendenzen und Perspektiven* (pp. 243-257). Frankfurt: Peter Lang.
- Stenberg, G. (2003). Effects of maternal inattentiveness on infant social referencing. *Infant and Child Development*, 12(5), 399-419.
- Streeck, J. and Hartge, U. (1992). Gestures at the Transition Place. In P. Auer and A. di Luzio (Eds.). *The Contextualization of Language* (pp. 135-57). Amsterdam: John Benjamins.
- Streeck, J. (1996). How to Do Things with Things. *Human Studies*, 19, 365-384.
- Streeck, J. (2005). On projection. In E. Goody (Ed.). *Interaction and Social Intelligence* (pp. 84-110). Cambridge: Cambridge University Press.
- Streeck, J. and Jordan, J.S. (2009). Projection and Anticipation: The Forward-Looking Nature of Embodied Communication. *Discourse Processes*, 46(2-3), 93-102.
- Stukenbrock, A. (2008). “Wo ist der Hauptschmerz?” – Zeigen am eigenen Körper in der medizinischen Kommunikation. In: *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion*, 9, 1-33.
- Stukenbrock, A. (2012). *Deixis in der face-to-face-Interaktion*. Habilitationsschrift: Freiburg Institute for Advanced Studies (FRIAS), Albert-Ludwigs-Universität Freiburg.
- Suchman, L. (1997). Centers of Coordination: A case and some themes. In L. B. Resnick, R. Säljö, C. Pontecorvo & B. Burge (Eds.). *Discourse, Tools, and Reasoning: Essays on Situated Cognition* (pp. 41-62). Berlin: Springer.
- Svensson, M. S., Heath, C. and Luff, P. (2007). Instrumental action: the timely exchange of implements during surgical operations. *10th ECSCW* (pp. 41-60).
- ten Have, P. (2004). *Understanding Qualitative Research and Ethnomethodology*. London: Sage.
- ten Have, P. (2007). *Doing Conversation Analysis. A Practical Guide (Second Edition)*. London: Sage.
- Tomasello, M. (1995). Joint attention as social cognition. In: C. Moore & P. Dunham (Eds.). *Joint attention: its origins and role in development* (pp. 85-101). Hillsdale, NJ: LEA.
- Tomasello, M. (2000). The Social-Pragmatic Theory of World Learning. *Pragmatics*, 10(4), 401-413.
- Vorwerg, C. (2001). *Raumrelation in Wahrnehmung und Sprache: Kategorisierungsprozesse bei der Benennung visueller Richtungsrelationen*. Wiesbaden: Deutscher Universitäts-Verlag.
- Wald, B. (1976). *The Discourse Unit: A Study in the Segmentation and Form of Spoken Discourse*. Los Angeles: UCLA, Department of Linguistics (MS).
- Wald, B. (1978). Zur Einheitlichkeit und Einleitung von Diskurseinheiten. In U. Quasthoff (Ed.). *Sprachstruktur – Sozialstruktur. Zur linguistischen Theoriebildung* (pp. 128-150). Kronberg: Scriptor.
- Weber, M. (1984). *Soziologische Grundbegriffe* (pp. 1-30). Tübingen: Mohr.
- Weilhammer, K. & Rabold, S. (2003). Durational Aspects in Turn Taking. *Proceedings of the Interactional Conference of Phonetic Sciences*, Barcelona, Spain.
- Wolf, R. (2007). Rezension zu: Reinhold Schmitt (Hrsg.): Koordination. Analysen zur multimodalen Interaktion. Tübingen: Gunter Narr. *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion*, 8, 160-176.
- Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum.
- Yngve, V.H. (1970). On getting a word in edgewise. In: M. A. Campbell (Ed.). *Papers from the Sixth Regional Meeting, Chicago Linguistics Society*. Chicago: Department of Linguistics, University of Chicago.
- Yu, C., & Smith, L. B. (2013). Joint Attention without Gaze Following: Human Infants and Their Parents Coordinate Visual Attention to Objects through Eye-Hand Coordination. *PLoS ONE*, 8(11).

Abbildungsverzeichnis

Abbildung 1: Joint Attention & Ko-Orientierung	7
Abbildung 2: AR-System & Versuchsperson mit technischem Equipment	13
Abbildung 3: AR-System im Gebrauch mit dem Obersee II Szenario.....	14
Abbildung 4: Korpusstruktur	15
Abbildung 5: Attraktorfunktion der Gestenphase	24
Abbildung 6: Mediatorfunktion der Gestenphase.....	25
Abbildung 7: Mediatorfunktion auf einen spezifischen Referenten.....	33
Abbildung 8: Multimodaler Verdichtungsraum nach Stukenbrock (2012)	37
Abbildung 9: “Overlap”-Differenzierung nach Schegloff (2000).....	46
Abbildung 10: Emergenz eines „bifurkativ- multimodalen Verdichtungsraums“	56
Abbildung 11: Basisdifferenzierung zur Simultaneität nach Stukenbrock (2012).....	79
Abbildung 12: Funktionale Simultaneität	81
Abbildung 13: Konfligierende Simultaneität.....	83
Abbildung 14: Emergenz & Auflösung konkurrierender Aktivitäten (F2F vs. AR).....	98
Abbildung 15: Deiktische Dimensionen – Bühlers Konzeption der Origo (1934)	104
Abbildung 16: Deiktischer Gesten auf Objekte und Bereiche der Karte	114
Abbildung 17: Produktions- & Rezeptionsformat: „HIER auf der anderen seite-“.....	119
Abbildung 18: Salienzkonfigurationen: Anheben und Vorzeigen	139
Abbildung 19: Objektmanipulative Handlungen (F2F vs. AR)	140
Abbildung 20: AR-System im Gebrauch mit dem Exhibition-Design-Szenario (2010)	145
Abbildung 21: Instabilität der Referenzobjekte im Exhibition-Design-Szenario.....	146
Abbildung 22: Chronologie referentieller Verfahren auf Objekte.....	147
Abbildung 23: Referentielle Verfahren im kontrastiven Vergleich (F2F vs. AR)	153
Abbildung 24: Grundordnung der Koordinierung im Übergang interaktionaler Aufgaben	155
Abbildung 25: Sequenzstruktur der Verbalannotation (VP-Gr. 03, F2F)	164
Abbildung 26: Schematisierung der sequenzstrukturellen Auffälligkeiten der Verbalebene	165
Abbildung 27: Intervallverschiebung zur Bestimmung der maximalen Abweichung	168
Abbildung 28: Redeanteil der Versuchspersonen pro Minutenintervall (VP-Gr. 03, F2F).....	169
Abbildung 29: Objektmanipulative Handlungen der Versuchspersonen pro Minutenintervall	170
Abbildung 30: Deiktische Gesten (@objects+@map+@VP-A/B) pro Minutenintervall	172
Abbildung 31: Aktivitätspotentiale der Beteiligungsressourcen A/B-ver, -O-rel und -dG	174
Abbildung 32: Scatter Plots und Regressionsgrade (gesamt).....	175
Abbildung 33: Gesamtverteilung der deiktischen Gesten (VP-Gr. 03, F2F).....	176
Abbildung 34: Scatter Plots und Regressionsgrade (Deiktische Gesten)	177
Abbildung 35: Deiktische Gesten (@objects) pro Minutenintervall (VP-Gr. 03, F2F).....	177
Abbildung 36: Deiktische Gesten (@map) pro Minutenintervall (VP-Gr. 03, F2F).....	178
Abbildung 37: Initialkonfiguration der “chains of prior action” im Präkontext der Umschlagsdynamik..	186
Abbildung 38: Die manipulationsstärksten Objekte im Interaktionsverlauf (VP-Gr. 03, F2F).....	187

Anhang

Transkriptionskonventionen

Die Transkription der verbalsprachlichen Ebene folgt den von Selting et al. (1998; 2009) vorgeschlagenen Zeichenkonventionen des Gesprächsanalytischen Transkriptionssystems (GAT) und ist in Partiturschreibweise realisiert. Die einzelnen turn-by-turn Abfolgen im interpersonellen Dialog werden dabei nicht per se in einer vertikalen, sondern ebenso in einer horizontalen Zeitachse angeordnet, um einerseits den Expansionseffekt multimodaler Verschriftlichung zu minimieren und andererseits, die sequentielle Simultaneität verschiedener intra- und interpersoneller Aktivitäten in semiotischen Feldern visuell darstellbar zu machen.

quadbahn kann man sagen; genau; eine NATU:RBAHN	Überlappungen und simultane Ereignisse
=	schneller, unmittelbarer Anschluss neuer Turns oder Einheiten
(.)	Mikropause
(-), (--), (---), (1.2)	Pausen von ca. 0.25 bis 0.75 Sekunden, bzw. der angegebenen Zeit
und=äh	Verschleifungen innerhalb von Einheiten
:, ::, :::	Dehnung, Längung, je nach Dauer
äh, ähm, etc.	Hesitationssignale
hab ich doch gesa'	Abbruch durch Glottalverschluss
hehehe, hihhi, etc.	silbisches Lachen
ja, ne, ne=ne, hm=hm, ja=a, etc.	Rezeptionssignale
((lautes Räuspern))	para- und außersprachliche Handlungen und Ereignisse
(xxxxxx)	unverständliche Passage je nach Länge
(solche)	vermuteter Wortlaut
[...]	Auslassung im Transkript, ggf. mit Zeitangabe
naTÜRLich, naTÜRLICH	silbische Akzentuierung
auf !KEI!nen fall	extra starke silbische Akzentuierung
?	Tonhöhe steigt am Einheitenende
-	Tonhöhe bleibt gleich am Einheitenende
;	Tonhöhe fällt am Einheitenende
`SO	fallende Akzenttonhöhenbewegung
'SO	steigende Akzenttonhöhenbewegung
˘SO	gleichbleibende Akzenttonhöhenbewegung
^SO	steigend-fallende Akzenttonhöhenbewegung
<<p> >	=piano, leise
<<pp> >	=pianissimo, sehr leise
<<all> >	=allegro, schnell
<<acc> >	=accelerando, schneller werdend
°h, °hh, °hhh	Einatmen, je nach Dauer

Für die sichtbaren Ebenen der Interaktion (Blick, Gestik, Körperpositur, Objektmanipulationen, etc.) wurden eigene Transkriptionskonventionen entwickelt, die äquivalent zur verbalsprachlichen Ebene in Partiturschreibweise realisiert werden. Um eine möglichst präzise Vorstellung der ‚nonverbalen‘ Handlungen in ihren situativen Kontexten zu ermöglichen, werden aus dem Video extrahierte Standbilder verwendet und fortlaufend nummeriert (in der Form: *Nr.). Ihre zeitlich-sequentielle Verortung ist den entsprechenden Nummern unterhalb der einzelnen Transkriptpartituren zu entnehmen.

Wenn die zeitlich-sequentielle Emergenz für das im Analysefokus stehende Phänomen keine Rolle spielt, wird die Beschreibung para- und außersprachlicher Handlungen in der Form „((greift das Objekt XY und dreht es in der Handfläche))“ präferiert. Für pre-sequentiell eindeutig markierte Handlungsweisen (z.B. deiktische Gesten durch „dG“) gilt, dass nachfolgende Modifikation im weiteren Sequenzverlauf durch entsprechende Gestaltbeschreibungen erfasst werden (z.B. “tap” für ein deiktisches Tippen). Worauf sich die Gestaltbeschreibung bezieht, ergibt sich folglich aus der vorausgehend markierten Handlungsaktivität. Für die Lesbarkeit der multimodalen Transkripte ist dieses Prinzip entscheidend, denn gegenüber

verbalsprachlichen Transkriptionen, in denen prinzipiell an unterschiedliche Stellen im Transkript „gesprungen“ werden kann (ohne dass die Lesbarkeit des Transkripts leidet), machen es die dynamischen Veränderungen ‚nonverbaler‘ Handlungen erforderlich, aufeinander beziehbare Zeichenkonventionen zu entwickeln, die diese Dynamiken abbilden können. Für ein Verständnis der Transkripte ist es somit relevant, die transkribierten Zeichenkonventionen immer fortlaufend aufeinander zu beziehen und strikt dem sequentiellen Verlauf der Handlungsemergenz zu folgen.

A-ver:	Notation verbalsprachlicher Elemente von Interaktant A
A-gaz:	Notation der Blickrichtung von Interaktant A
>>>WPA	A blickt zum Objekt “WaterProtectionArea” (WPA)
@WPA	A fixiert das Objekt WPA
@WPA...	Sequentielle Dauer der Objektfixierung (im Verhältnis zur verbalsprachlichen Notation)
>>>B (face)	A blickt in B’s Gesicht (Klammerausdrücke spezifizieren die vorausgehende Notation)
A-act:	Notation nicht-verbalsprachlicher Aktivitäten von Interaktant A
ggf.: A-lH / A-rH	Notation für beidhändige Aktivitäten
dG>>>@WPA>>>HP	A zeigt auf das Objekt WPA (der <i>Onset</i> ist durch ‚dG‘ (deiktische Geste) markiert; der <i>Stroke</i> durch ein ‚@‘; die <i>Retraktion</i> endet in ‚HP‘ (home position))
dG>>>@WPA...	Sequentielle Dauer der <i>Strokephase</i> (im Verhältnis zur verbalsprachlichen Notation)
grasp>>>@WPA	A greift das Objekt WPA (äquivalent zur dG-Notation)
lift>>>	A hebt das Objekt an und bewegt es
place>>>@loc-1	A platziert das Objekt auf dem Lokalisationsort ‚loc-1‘
>>>	spezifische Bewegungen, die auf ein Ziel zulaufen
~~~	unspezifische Bewegungen, die auf kein Ziel zulaufen
hold	Aufrechterhalten (Einfrieren) einer Geste/Objektbewegung/Positur
frwd	Vorwärtsbewegung (egozentrischer Perspektive)
bwrđ	Rückwärtsbewegung (egozentrische Perspektive)

## **Eidesstattliche Erklärung**

Hiermit erkläre ich, dass mir die geltende Promotionsordnung der Fakultät für Linguistik und Literaturwissenschaft bekannt ist. Ich habe die vorliegende Arbeit selbst angefertigt, keine Textabschnitte von Dritten oder eigener Prüfungsarbeiten ohne Kennzeichnung übernommen und alle von mir benutzten Hilfsmittel und Quellen in meiner Arbeit angegeben. Dritte haben weder unmittelbar noch mittelbar geldwerte Leistungen von mir für Vermittlungstätigkeiten oder für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen. Des Weiteren habe ich diese Dissertation noch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht und ich habe keine gleiche, in wesentlichen Teilen Ähnliche oder eine andere Abhandlung bei einer anderen Hochschule als Dissertation eingereicht.

Bielefeld, 6. August 2015

---

Christian Schnier