

Nora Schönfelder

Multimodale Ressourcen beim Argumentieren im Gespräch

Interaktions- und erwerbsbezogene Analysen

Verlag für Gesprächsforschung

Heidelberg: Verlag für Gesprächsforschung 2025
<http://www.verlag-gespraechsforschung.de>
ISBN 978 - 3 - 936656 - 98 - 5

Inaugural-Dissertation zur Erlangung der Doktorwürde der Philosophie (Dr. phil.)
im Fach Germanistik in der Fakultät für Geistes- und Kulturwissenschaften
der Bergischen Universität Wuppertal

vorgelegt von Nora Schönfelder

Erstgutachterin: Prof.'in Dr. Vivien Heller
Zweitgutachterin: Prof.'in Dr. Uta Quasthoff

Alle Rechte vorbehalten.

© Verlag für Gesprächsforschung, Dr. Martin Hartung, Heidelberg 2025

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt.
Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne
Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigung,
Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in
elektronischen Systemen.

Danksagung

Die Freude an menschlichen Interaktionen und an ihren Feinheiten in der Bedeutungskonstitution scheinen in vielen Menschen zu resonieren – so auch in mir. Eine wissenschaftliche Auseinandersetzung mit ihren kleinen und kleinsten Strukturen wurde mir jedoch erst auf unterschiedlichen universitären Stationen zuteil: In dem BMBF-Forschungsprojekt *EfA-TT* von Prof.‘in Dr. Uta Quasthoff an der TU Dortmund wurde ich in die konversationsanalytische Arbeit mit Gesprächsdaten eingeführt. Uta Quasthoff weckte in mir auch die Begeisterungsfähigkeit für die kommunikative Gattung des Argumentierens und von ihr lernte ich nicht zuletzt, wie viel ‚Musik in Interaktionen‘ steckt! Ihr bin ich sehr dankbar dafür, dass sie mich bei meinen ersten Schritten in der Arbeit mit menschlichen Interaktionen begleitete und ich dabei so viel von ihr lernen durfte. Ihr verdanke ich auch die Datenbasis der vorliegenden Arbeit, denn die videografierten Eltern-Kind-Argumentationen stammen aus dem Vorgängerprojekt von *EfA-TT*, aus *FUnDuS*. Die zweite, für mich außergewöhnlich prägende Station bildet die Arbeit am Lehrstuhl von Prof.‘in Dr. Vivien Heller an der Bergischen Universität Wuppertal. Durch Vivien Heller lernte ich, das Interesse an Gesten und ihren Funktionen in konkrete Fragen zu gießen und ‚multimodal zu denken‘. In ihr hatte ich stets eine Mentorin, die die Faszination für die kleinsten Feinheiten in Interaktionen mit großer Begeisterung teilte und mit der ich lange über ressourcenreichen Transkripten gemeinsam brüten konnte. Vivien Heller gilt mein besonders großer Dank, denn in der Betreuung der Arbeit hat sie stets wahrgenommen, ob es gerade einer wissenschaftlichen oder eher einer emotionalen Begleitung bedarf, ob ein Ansporn notwendig ist oder aber eine Phase des Schreibens eher Ruhe benötigt. Vielen Dank für all das!

Doch in die Arbeit sind natürlich auch weitere Akteur*innen involviert: Meine Doktorgeschwister am Lehrstuhl waren (und sind!) mir meine wichtigsten Partnerinnen. Sie haben sich in die kleinteiligen Analysen mit mir gemeinsam vertieft, teilten meine Freude über neue Befunde und jedes weitere entstandene Kapitel. Denise, Noelle, Nicole und Nadine, Ihr habt auf unterschiedliche Weise großen Anteil am Entstehen der vorliegenden Arbeit – merci! Darüber hinaus möchte ich allen Teilnehmer*innen des Wuppertaler Forschungskolloquiums ‚Sprache – Interaktion – Lernen‘ für ihre Anregungen in den Datensitzungen danken.

Auch ohne die Unterstützung meiner Freund*innen außerhalb der Universität wäre der lange Entstehungsprozess der Dissertation kaum zu bewältigen gewesen. Ihnen allen gilt mein herzlicher Dank – insbesondere Valeria und Lydia.

Die Begeisterung für das Miteinander, für Sprache und für das Ergründen von Bedeutungen in der Interaktion ist in meiner Familie verankert. Dabei haben mich meine Eltern stets in dem Vorhaben, meiner Neugier zur Funktion von Gesten in der menschlichen Interaktion nachzugehen, mit großer Liebe bestärkt und meinen Weg beständig begleitet. Seit Langem geht dabei Jan an meiner Seite, der jeglicher Freude über neu Erlerntes mit großem Enthusiasmus begegnet, sodass Funken zu Feuer werden. Ich danke Euch, denn ohne Euch wäre alles nichts.

Der Abschluss der vorliegenden Arbeit ist dankenswerterweise durch ein Abschlussstipendium der Bergischen Universität Wuppertal ermöglicht worden. Die nachfolgende Fassung ist eine leicht angepasste Version meiner Dissertationsschrift, die ich im April 2024 an der Bergischen Universität Wuppertal eingereicht und im Juli 2024 verteidigt habe.

Inhaltsverzeichnis

Danksagung	3
1 Einleitung	7
2 Mündliches Argumentieren	12
2.1 Konversationsanalytischer Zugriff auf das Argumentieren	12
2.2 Mündliches Argumentieren als kommunikative Gattung & diskursive Praktik	16
2.2.1 Sequenzielle Organisation mündlicher Argumentationen	18
2.2.2 Pragmatische Mittel und Formen ihrer Realisierung	20
2.3 Zusammenfassung	26
3 Multimodale Ressourcen in der Interaktion	27
3.1 Eine konzeptionelle Einordnung: Untersuchung multimodaler Ressourcen in der Interaktion	27
3.1.1 Simultaneous Constructions (Adam Kendon)	28
3.1.2 Composite Utterances (Nick J. Enfield)	30
3.1.3 Contextual Configuration (Charles Goodwin)	31
3.1.4 Complex Multimodal Gestalts (Lorenza Mondada)	33
3.2 Gesten als visuelle Elemente multimodaler Gestalten	35
3.2.1 Gesten: Formen und Funktionen	36
3.2.2 Gesten mit pragmatischer Funktion: Interaktionsbezogene und diskurs- bzw. äußerungsstrukturierende Gesten	47
3.2.2.1 Kategorisierung pragmatischer Funktionen	47
3.2.2.2 Gestenfamilien	51
3.3 Funktionen pragmatischer Gesten für das Argumentieren	53
3.4 Zusammenfassung	58
4 Multimodale Ressourcen im Diskurserwerb	60
4.1 Erwerb mündlicher Diskurskompetenz	60
4.1.1 Modellierungen von Argumentationskompetenz: Argumentative (Teil-)Fähigkeiten und ihr Erwerb	61
4.1.2 Linien des Diskurserwerbs	68
4.1.3 Externe Erwerbsressourcen: Erwerbskontexte für das mündliche Argumentieren	70
4.2 Erwerbsbezogene Befunde zur Entwicklung und Funktionen multimodaler Ressourcen in diskursiven Praktiken	76
4.2.1 (Nach-)Erzählen und Erklären	76
4.2.2 Teilaspekte des Argumentierens	82
4.3 Zusammenfassung und Ausblick: Gesten als Teil multimodaler Gestalten beim Argumentieren und im Diskurserwerb	84
5 Forschungsdesiderate und Fragestellungen	86
6 Daten	89
6.1 Datengrundlage	89
6.2 Datenerhebung	91
6.3 Datenaufbereitung	93

7	Methode.....	96
7.1	Forschungsmethodische Einordnung.....	96
7.2	Vorgehen in der Datenanalyse.....	99
8	Interaktionsbezogene Analysen der multimodalen Gestalten zur Realisierung argumentativer Handlungen für das Argumentieren.....	102
8.1	Position beziehen.....	103
8.1.1	Hervorheben von Optionen beim Beziehen von Positionen.....	103
8.1.2	Position beziehen durch Ausschließen einer Option	108
8.1.3	Zusammenfassung.....	111
8.2	Begründen	112
8.2.1	Präsentieren während des Begründens	112
8.2.1.1	Präsentieren von (komplexeren) Begründungen	113
8.2.1.2	Kurzzeitiges Präsentieren und Hervorheben relevanter Bedeutungselemente der Begründung	118
8.2.2	Präsenhalten eines verhandelten Gesprächsgegenstandes innerhalb von Begründungen	121
8.2.2.1	Etablieren und Präsenhalten eines verhandelten Gesprächsgegenstandes	122
8.2.2.2	Etablieren und Präsenhalten eines Gesprächsgegenstandes bei konkurrierender Turnübernahme	128
8.2.2.3	Referenzieren und Präsenhalten eines verhandelten Gesprächsgegenstandes als Verstehensdokumentation (rezipient*innenseitig)	132
8.2.3	Multimodales Einführen einer Turnfortsetzung	135
8.2.4	Multimodale Kohäsion und Kohärenz	137
8.2.5	Verwerfen in Begründungen	143
8.2.6	Zusammenfassung.....	146
8.3	Zustimmen.....	147
8.3.1	Nachdrückliches Zustimmung.....	147
8.3.2	Schwaches Zustimmung in Konzessionen	151
8.3.3	Schwaches Zustimmung zu evidenten Begründungen & Positionen	153
8.3.4	Zusammenfassung.....	158
8.4	Einwenden.....	159
8.4.1	Präsentieren während des Einwendens	160
8.4.2	Sichtbares Explizieren von Negationen in Einwänden	164
8.4.3	Anhalten einer vorangegangenen Opposition in Konzessionen	167
8.4.4	Mahnendes Einwenden	169
8.4.5	Zusammenfassung.....	171
8.5	Einfordern von Rückmeldungen.....	172
8.5.1	Einfordern von Rückmeldungen im Listenformat.....	173
8.5.2	Präsenhalten des Gesprächsgegenstandes in eingeforderten Rückmeldungen	177
8.5.3	Zusammenfassung.....	180
8.6	Markieren epistemischer Modalisierung	181
8.6.1	Markieren von Evidenz	182
8.6.1.1	Markieren von Evidenz durch Fallenlassen der Hand	182
8.6.1.2	Markieren von Evidenz durch Palm Up Gesten mit lateraler Bewegung	184
8.6.1.3	Markieren von Evidenz durch Hand Shrug	186
8.6.2	Markieren des ablehnenden Fallenlassens eines Gesprächsgegenstandes.....	188
8.6.3	Anzeige von epistemischer Unsicherheit	192
8.6.4	Markieren von Unbestimmtheit	196
8.6.5	Zusammenfassung.....	198
9	Erträge für eine multimodale Gattungsforschung	203

10	Längsschnittliche Fallstudien zu erwerbsbezogenen Veränderungen der multimodalen Gestalten für das Begründen	206
10.1	Diskursstrukturelle Profile	206
10.2	Sprachlich-diskursive Entwicklung der Fokuskinder	209
10.2.1	Fokuskind I – Exploratives Position-Beziehen und Begründen	209
10.2.1.1	Kontextualisierung	209
10.2.1.2	Vertextung	212
10.2.1.3	Epistemische Modalisierung	216
10.2.1.4	Zusammenfassende Beschreibung der sprachlich-diskursiven Entwicklungslinie	218
10.2.2	Fokuskind II – Zustimmungendes Position-Beziehen und Bewerten	219
10.2.2.1	Kontextualisierung	219
10.2.2.2	Vertextung	223
10.2.2.3	Epistemische Modalisierung	228
10.2.2.4	Zusammenfassende Beschreibung der sprachlich-diskursiven Entwicklungslinie	230
10.2.3	Interindividueller Vergleich	232
10.3	Entwicklung multimodaler Gestalten für das Begründen	233
10.3.1	Fokuskind I – Begründen	234
10.3.1.1	Kontextualisierung	234
10.3.1.2	Vertextung	240
10.3.1.3	Epistemische Modalisierung	248
10.3.1.4	Zusammenfassende Beschreibung der multimodalen Entwicklungslinie	254
10.3.2	Fokuskind II – Begründen	258
10.3.2.1	Kontextualisierung	258
10.3.2.2	Vertextung	262
10.3.2.3	Epistemische Modalisierung	270
10.3.2.4	Zusammenfassende Beschreibung der multimodalen Entwicklungslinie	272
10.3.3	Interindividueller Vergleich	276
10.4	Zusammenhänge zwischen multimodalen und sprachlich-diskursiven Entwicklungslinien der Kinder in der argumentativen Handlung des Begründens	280
10.5	Zusammenfassung und Ausblick: Hypothesen zum Erwerb	285
11	Diskussion: Funktionalität multimodaler Ressourcen für die kommunikative Gattung des Argumentierens und den Erwerb argumentativer Diskurskompetenz	289
12	Ausblick	296
13	Literatur	302
14	Anhang	324
14.1	Abbildungsverzeichnis	324
14.2	Tabellenverzeichnis	325
14.3	Transkriptionskonventionen	327

1 Einleitung

It makes no sense to speak of ‘verbal communication’ and ‘nonverbal communication.’
There is only communication [...]. (Kendon 1972a: 443)

Über die Frage, was eigentlich mündliches Argumentieren ist bzw. was mündliche Argumentationen ausmacht, lässt sich selbst vortrefflich diskutieren. Scheinen sie doch in ihren Formen und Funktionen so vielgestaltig zu sein, dass Deppermann (2003: 23, Hervorh. N. S.) von „Gattungen des Argumentierens“ spricht. Einig scheint sich die Forschung aber darin zu sein, dass das Argumentieren als eine komplexe sprachliche Aktivität zu verstehen ist (vgl. Klein 1980; Spranz-Fogasy 2003; Heller 2012), in der Interagierende einen im Gespräch aufgetretenen Dissens bzw. ein Problem bearbeiten und dabei u. a. Argumente gegeneinander abwägen, Perspektiven anderer Interagierender einnehmen oder die Argumente anderer entkräften (vgl. u. a. Grundler 2011; Heller 2012). Im Wesentlichen kann die Aktivität von Interagierenden im Gespräch nur *gemeinsam* und in wechselseitiger Bezugnahme aufeinander bearbeitet werden. Der von Deppermann (2003) gewählte Begriff der Gattungen verweist nun darauf, dass sich in der Bearbeitung kommunikativer Probleme strukturelle Regelhaftigkeiten verfestigt haben, an denen sich Mitglieder einer Sprachgemeinschaft unbewusst orientieren: So zeigt sich, dass Argumentationen durch einen Dissens oder ein Problem im laufenden Gespräch ausgelöst werden, in Bezug auf den/das die Interagierenden sich gegenseitig einen weiteren Aushandlungsbedarf anzeigen. Durch das Einbringen von Positionen, ihre Plausibilisierung durch Begründungen oder auch das Infragestellen der Beiträge anderer Sprecher*innen kann das Problem bzw. der Dissens dann bearbeitet werden, bevor die Interagierenden anschließend ihr Gespräch fortsetzen oder sich anderen Diskursaktivitäten zuwenden können (vgl. Spranz-Fogasy 2003; Heller 2012; vgl. zum Gattungsbegriff Günthner und Knoblauch 1996). Bislang fokussiert die hier vorgenommene Beschreibung des mündlichen Argumentierens allerdings auf seine Funktion sowie *eine* mögliche Form seiner Bearbeitung – durch Sprache. Gehen wir jedoch davon aus, dass die oben sichtbar werdenden Aushandlungsprozesse in gemeinsamen Gesprächssituationen in Kopräsenz der Interagierenden entstehen, stellt sich auch die Frage danach, welche Rolle eigentlich *körperliche* Ressourcen, wie Hand- und Kopfgesten oder auch Gesichtsausdrücke, für das Argumentieren übernehmen.

Unlängst ist bekannt, dass im Rahmen mündlicher Argumentationen multimodale Ressourcen eingesetzt werden, um einen Dissens sichtbar zu machen, wie z. B. Clift (2021) für das Augenrollen rekonstruiert, mit dem Interagierende, ohne ihr Gegenüber zu unterbrechen, eine abweichende Meinung anzeigen können. In der Bearbeitung von Dissensen oder Problemen treten dann immer wieder besonders *gestische* Ressourcen in Erscheinung – v. a. Gesten mit *pragmatischer* Funktion (vgl. Kendon 2004: 158f.), die nicht semantisch mit der Äußerung verbunden sind, sondern anderen Beteiligten beispielsweise eine interpretative Rahmung für die eingebrachte Äußerung liefern (vgl. u. a. Jacquin 2015, 2017; Schönfelder und Heller 2019; Marrese et al. 2021; Schönfelder 2024). So können Interagierende z. B. einen Gesprächsgegenstand wie eine Begründung oder einen Einwand als abstraktes diskursives Objekt auf der geöffneten Handfläche dem Gegenüber zur Betrachtung anbieten (vgl. Müller 2004) und, wird die Geste von mehreren Sprecher*innen nacheinander eingesetzt, das Geben-und-Nehmen der eingebrachten Argumente visuell wahrnehmbar anzeigen (vgl. Schönfelder und Heller 2019) – die gestische Ressource trägt somit im o. g. Sinn zur interpretativen Rahmung der Gesprächssituation bei.

Dieser erste Einblick in die Funktionen von multimodalen Ressourcen für das Argumentieren zeigt bereits, dass dem Körper für die soziale Interaktion eine wichtige Rolle zukommt. Denn durch ihn können Menschen nicht nur in eine gemeinsame Interaktion eintreten (vgl. Goffman 1963) und sich gegenseitig z. B. durch ihre (körperliche) Orientierung aufeinander anzeigen, dass sie Teil der gemeinsamen, sozialen Interaktion sind – oder sich aus dieser herauslösen (vgl. Goodwin 1981) –, sondern körperlich-sichtbare Ressourcen können eben auch weitere kommunikative Funktionen für die Interaktion erfüllen.¹ Dabei werden Ressourcen wie die Gestik, Mimik oder auch die körperliche Ausrichtung nicht ‚willkürlich‘ von den Interagierenden eingesetzt, sondern sie tragen ‚regelmäßig‘ mit sprachlichen Ressourcen zur Bedeutungskonstitution in der Interaktion bei (vgl. u. a. Goodwin 2000; Mondada 2014b; Deppermann 2018). Seither wurde in zahlreichen Arbeiten die Breite des Funktionsspektrums multimodaler Ressourcen beleuchtet. Es konnte dabei rekonstruiert werden, dass (relativ) stabile Verbünde von Ressourcen *wiederkehrend* an der Herstellung sozialer Interaktion beteiligt sind (vgl. Mondada 2014b, 2016, 2018) und auch in Argumentationen interaktive Funktionen erfüllen. So legt beispielsweise die Studie von Jacquin (2018) dar, dass Interagierende zum Anzeigen von Opposition sowohl in öffentlichen Debatten als auch in Arbeitsplatzbesprechungen systematisch Gesten und den Blick zum*zur Adressat*in nutzen, wobei zugleich eine explizit-sprachliche Referenz auf den*die Proponent*in vermieden und somit der Grad der Konfrontation reduziert wird. Die körperlich-sichtbaren Ressourcen, d. h. die deiktische Geste sowie die Blickadressierung, werden in den untersuchten Kontexten der Arbeitsplatzbesprechungen und öffentlichen Debatten wiederholt zur Referenzherstellung in Oppositionen eingesetzt und erfüllen dabei konstant die Funktion, die Konfrontation abzumildern. Diese multimodalen Ressourcenverbünde bilden dabei in ihrer Struktur komplexe multimodale Gestalten (vgl. Mondada 2014b: 137). Aus einem soziologisch geprägten konversationsanalytischen Verständnis heraus werden die körperlich-sichtbaren und sprachlichen Ressourcen somit als Ausdrucksressourcen aufgefasst, die in bestimmten Verbindungen regelmäßig eingesetzt werden und somit zur Ordnung der sozialen Interaktion beitragen.

Trotz der Erkenntnis, dass multimodale Ressourcen also relevante Funktionen für das mündliche Argumentieren übernehmen, steht eine systematische Untersuchung bislang noch aus. Es zeigt sich außerdem, dass die existierenden Studien ihre Befunde nur selten hinsichtlich der untersuchten Diskursaktivität – im Fall der vorliegenden Arbeit: dem mündlichen Argumentieren – einordnen. Daher bleibt bislang größtenteils verborgen, wie multimodale Ressourcen funktional für das mündliche Argumentieren eingesetzt werden. Das Beispiel aus der Studie von Jacquin (2018) gibt einen ersten Einblick in die Beantwortung der Frage, denn die Funktion der körperlich-sichtbaren Ressourcen (Geste und Blick) wird hier im Rahmen der eingenommenen *Opposition* untersucht. Erst indem die argumentative Handlung des Opponierens bzw. Einwendens berücksichtigt wird, kann die Funktion der interpersonellen deiktischen Referenz erfasst werden, mit der die Sprecher*innen den Grad ihrer Konfrontation reduzieren. Bliebe diese kontextuelle Information unberücksichtigt, könnte die Funktion der gestischen Ressource beispielsweise auch als Rederechtszuweisung verstanden werden, mit der ein*e nächste*r Sprecher*in adressiert wird. Die Relevanz des Entstehungskontextes ist für die Rekonstruktion der Funktion multimodaler Ressourcen unverkennbar. Umso erstaunlicher ist es, dass bislang noch keine systematische Untersuchung unter Berücksichtigung der Diskursaktivität des Argumentierens vorliegt.

¹ In der Unterscheidung zwischen sprachlichen und körperlich-sichtbaren Ressourcen folgt die Arbeit Stivers und Sidnell (2005: 2), die zwischen „vocal/aural and visuospatial modalities“ differenzieren. In der vorliegenden Arbeit ist mit dem Begriff der sprachlichen Ressource die oberflächensprachliche Ebene samt ihren prosodischen Merkmalen gemeint, wohingegen die körperlich-sichtbaren Ressourcen Hand-, Kopf- und faziale Gesten, Blickadressierungen etc. umfassen. Sobald in der vorliegenden Arbeit dann von *multimodalen* Ressourcen gesprochen wird, sind alle Ausdrucksressourcen gemeinsam adressiert.

Dieses Forschungsdesiderat adressiert die vorliegende Arbeit und untersucht in einem ersten Schritt die folgende Forschungsfrage I: *Welche Funktionen übernehmen multimodale Ressourcen für das mündliche Argumentieren?*

Die vorliegende Studie basiert auf einem Datenkorpus von insgesamt 20 Eltern-Kind-Argumentationen, die im Rahmen einer gemeinsamen Entscheidungsfindung erhoben worden sind. Es handelt sich um ein längsschnittliches Datenkorpus, das über drei Erhebungszeitpunkte, d. h. im Alter der Kinder von 12, 13 und 15 Jahren, mündliche Argumentationen von den Dyaden elizitiert hat. Als Grundlage für die Untersuchung der o. g. Fragestellung dienen die Daten des dritten Erhebungszeitpunktes, zu dem die jugendlichen Sprecher*innen bereits argumentationserfahren(er) sind. Wenn nun multimodale Ressourcen innerhalb des Argumentierens untersucht werden und ihr interaktiver Entstehungskontext mit berücksichtigt werden soll, muss freigelegt werden, vermittels welchen *argumentativen Handlungen*, wie z. B. dem Beziehen einer Position, dem Begründen oder dem Einwenden, die Interagierenden die Entscheidungsfindung bearbeiten (vgl. u. a. Gohl 2006; Heller 2012; Grundler et al. 2020). Dabei wird die sequenzielle Entfaltung der Argumentation Redezug-um-Redezug nachvollzogen, in der die Interagierenden z. B. durch Einwände auf vorangegangene Positionen reagieren oder eine selbst eingebrachte Position durch Begründungen stützen. Es wird also konversationsanalytisch rekonstruiert, wie die Interagierenden auf Vorgängeräußerungen reagieren und zugleich durch eigene argumentative Handlungen Anschlussstellen für den Fortgang der Interaktion herstellen. Erst mit Nachvollzug der interaktiven Entfaltung der Argumentation wird es dann möglich, die innerhalb der argumentativen Handlungen eingesetzten multimodalen Ressourcen hinsichtlich ihrer Funktionalität für die jeweilige argumentative Handlung zu untersuchen (vgl. u. a. Mondada 2014b). Das Augenmerk in der vorliegenden Arbeit wird auf gestische Ressourcen als visuell-wahrnehmbare Ausdrucksressourcen der Interaktion gelegt. Da Gesten aber i. d. R. mit weiteren Ressourcen, wie dem Blick oder der Körperpositur, koordiniert werden und gemeinsam Ressourcenverbünde bilden, werden diese konsequent in der Untersuchung berücksichtigt (vgl. Kap. 3). Das Ziel der ersten empirischen Fragestellung besteht nun darin, auf Grundlage des Datenkorpus multimodale Ressourcenverbünde und ihre pragmatischen Funktionen für das Argumentieren zu rekonstruieren. Indem die Studie dabei die Diskursaktivität einbezieht, kann sie bisherige Forschungsbefunde um grundlegende, systematische Einsichten hinsichtlich der Funktion multimodaler Ressourcen für das Argumentieren ergänzen.

Ausgehend von dem dargelegten Verständnis, dass multimodale Ressourcen in (festen) Verbünden wiederkehrende kommunikative Funktion in der Interaktion übernehmen, stellt sich nun die Frage nach dem *Erwerb* dieser routinisiert verwendeten Ressourcenverbünde – i. S. eines multimodalen Diskurserwerbs. Aus Studien zum sprachlich-diskursiven Argumentationserwerb geht hervor, dass heranwachsende Sprecher*innen mit ansteigenden diskursiven Fähigkeiten ihre Beiträge in Argumentationen zunehmend selbstständig einbringen; die Sprecher*innen erkennen, dass eingebrachte Positionen einer Begründung bedürfen und es gelingt ihnen ebenfalls vermehrt, ihre Beiträge zu größeren Äußerungspaketen auszubauen (vgl. u. a. Wild et al. 2012; Quasthoff et al. 2021a: 16). Dazu nutzen die Heranwachsenden allerdings nicht nur sprachlich-diskursive, sondern auch körperlich-sichtbare Ressourcen in kommunikativer Funktion. Beispielsweise werden gestische Ressourcen verwendet, um einer Beschwerde Nachdruck zu verleihen (vgl. Ladewig und Hotze 2023) oder aber sich gegenseitig das Geben-und-Nehmen von Argumenten anzuzeigen (vgl. Schönfelder und Heller 2019). Es ist heranwachsenden Sprecher*innen also durchaus möglich, multimodale Ressourcen in mündliche Argumentationen einzubringen, sodass sich die Frage stellt, wie sich der Einsatz der körperlich-sichtbaren Ressourcen im Zusammenhang mit ihren sprachlich-diskursiven Fähigkeiten entwickelt. Die vorliegende Arbeit geht i. d. S. von einem *multimodalen* Diskurserwerb argumentativer Fähigkeiten aus. Diese Annahme stützt sich auch auf Studien zum multimodalen Diskurserwerb narrativer und explanativer Fähigkeiten von Heranwachsenden im Alter

von 4 bis 11 Jahren, in denen sichtbar wird, dass sich mit dem fortschreitenden Erwerb der Fähigkeit, eine Erzählung bzw. Erklärung aufzubauen, vormals häufig eingesetzte abbildende Gesten reduzieren, während Gesten mit pragmatischer Funktion zunehmen (vgl. u. a. Colletta und Pellenq 2009; Colletta et al. 2010; Reig Alamillo et al. 2012; Levy und McNeill 2013, 2015; Graziano 2014b; Kern et al. 2024). Befunde bezüglich des *multimodalen Argumentationserwerbs* liegen derzeit noch nicht vor.

Dieses Forschungsdesiderat greift die vorliegende Arbeit nun in einem zweiten Schritt auf und ist dabei an der nachfolgenden Forschungsfrage II interessiert: *Welche Veränderungen zeigen sich in einer längsschnittlich angelegten Studie zweier Fokuskinder im Alter von 12, 13 und 15 Jahren hinsichtlich der sprachlich-diskursiven Fähigkeiten sowie des Einsatzes multimodaler Ressourcen?*

Die Studie fokussiert somit auf den multimodalen Diskurserwerb einer *dialogischen* Diskurspraktik und nimmt dabei eine Alterskohorte in den Blick, die in der bisherigen Forschung des multimodalen Diskurserwerbs noch nicht berücksichtigt worden ist, obwohl sich auch in dieser Zeitspanne komplexe Diskurspraktiken wie die des Erklärens und Argumentierens weiterentwickeln (vgl. u. a. Wild et al. 2012; Quasthoff et al. 2021a: 16; Kinalzik 2024). Auf Grundlage des oben benannten Datenkorpus werden zwei Fallstudien in der Altersspanne der Heranwachsenden von 12 bis 15 Jahren durchgeführt. Die Untersuchung fokussiert auf zwei Kinder mit unterschiedlichen diskursstrukturellen Fähigkeiten und ermittelt vergleichend, welche Veränderungen über die Zeitspanne der Erhebungen sowohl sprachlich-diskursiv als auch hinsichtlich der verwendeten körperlich-sichtbaren Ressourcen zu beobachten sind. Dieses Vorgehen ermöglicht sowohl einen *intra-* als auch *interindividuellen* Vergleich der rekonstruierten sprachlich-diskursiven sowie multimodalen Fähigkeiten der Fokuskinder und ihrer Erwerbsverläufe. Aus diesem Vergleich resultierend können dann Zusammenhänge zwischen den sprachlich-diskursiven sowie multimodalen Entwicklungen begründet abgeleitet und diskutiert werden. Ziel der erwerbsbezogenen, explorativen Fallstudie ist es, erste Einsichten in den multimodalen Diskurserwerb zu gewinnen, auf deren Grundlage Erwerbs-hypothesen abgeleitet werden können.

Der Aufbau der Arbeit folgt dabei der folgenden Struktur: Es werden in einem ersten Schritt die relevanten theoretischen Fundierungen vorgenommen, auf denen die empirischen Untersuchungen aufbauen. Dazu wird eingangs aus einer konversationsanalytischen Perspektive vorgestellt, welchem Verständnis von mündlichem Argumentieren die vorliegende Studie folgt und wie die interaktive Entfaltung mündlicher Argumentationen analytisch nachvollzogen werden kann (vgl. Kap. 2). Den Blick nun auf multimodale Ressourcen lenkend wird daran anschließend das Verhältnis zwischen Sprache und körperlich-sichtbaren Ressourcen beleuchtet. Kendons eingangs benannter Prämisse zur Gleichberechtigung sprachlicher und körperlich-sichtbarer Ressourcen folgend (vgl. hierzu das Eingangszitat Kendon 1972a: 443) wird eine Konzeptualisierung dieses Verhältnisses vorgestellt, die multimodale Ressourcen, d. h. sprachliche sowie körperlich-sichtbare Ressourcen als fein aufeinander abgestimmte Ressourcen begreift, die gemeinsam holistische multimodale Gestalten aufbauen können. Diese Konzeptualisierung bildet die Grundlage der empirischen Untersuchung in Kapitel 8. Anschließend fokussiert das Kapitel auf Funktionen *gestischer* Ressourcen und stellt diese – mit Zuschnitt auf das mündliche Argumentieren – vor (vgl. Kap. 3). Schließlich werden in Vorbereitung auf die Fallstudien in Kapitel 10 bisherige Befunde zum Erwerb argumentativer, d. h. *sprachlich-diskursiver* Fähigkeiten sowie zur Verwendung *multimodaler* Ressourcen im Diskurserwerb beleuchtet (vgl. Kap. 4). Die sich aus dem Forschungsstand ergebenden Forschungsdesiderate und daraus abgeleiteten Forschungsfragen werden in Kapitel 5 dargelegt. Es folgen erläuternde Ausführung zur Datenerhebung und -aufbereitung (vgl. Kap. 6) sowie zum forschungsmethodischen Zugriff (vgl. Kap. 7), bevor dann der empirische Teil der Arbeit anschließt. Entlang der Forschungsfrage I erfolgt in Kapitel 8 eine empirische Untersuchung sowie systematische

Beschreibung der Funktionen multimodaler Ressourcen für das mündliche Argumentieren. Der Schwerpunkt der Untersuchung liegt darin, multimodale Gestalten, die an der Realisierung argumentativer Handlungen beteiligt sind, zu rekonstruieren und sie in ihrer Funktion zu untersuchen. Die dabei ermittelten Befunde werden hinsichtlich einer multimodalen Gattungsforschung diskutiert (vgl. Kap. 9) und fließen sodann in die erwerbsorientierte Fallstudie ein, die die Forschungsfrage II adressiert. In der Fallstudie werden die multimodalen Entwicklungslinien zweier Fokuskinder untersucht (vgl. Kap. 10). Im dritten Teil werden dann die Forschungsbefunde diskutierend zusammengeführt (vgl. Kap. 11) und es wird in einem Ausblick beleuchtet, welche weiterführenden u. a. sprachdidaktischen Forschungsdesiderate sich aus der vorliegenden Studie ergeben (vgl. Kap. 12).

2 Mündliches Argumentieren

Mit Blick auf das Forschungsinteresse der vorliegenden Studie soll zu Beginn der theoretischen Fundierung das mündliche Argumentieren in den Fokus gerückt werden. Dabei erweist sich das Erforschen mündlicher Argumentation als interaktives und nicht primär logisches oder sprecher*innenbezogenes Phänomen gemessen an der Geschichte des Gegenstandsbereichs als ein recht junges Forschungsfeld (vgl. zsf. Grundler 2011: Kap. 2.1). Lange Zeit widmen sich Arbeiten der Argumentationsforschung vornehmlich formal-logischen, rhetorischen oder (pragma-)dialektischen Kriterien zur Untersuchung von Argumentationen. Dabei werden zum einen die aufgestellten Prämissen, die zu einem Argument zusammengeführt werden, in der Logik ihrer Beziehungen zueinander, aber ohne Einbezug eines*einer Sprecher*in, untersucht (*logische Perspektive*). Zum anderen werden die Argumente hinsichtlich ihrer Überzeugungskraft, die sich in der Persuasion des Gegenübers manifestiert, beschrieben (*rhetorische Perspektive nach der rhetorischen Wende der 1950er Jahre*; vgl. u. a. Perelman und Olbrechts-Tyteca 1958/2004; Toulmin 1958). Nicht zuletzt wird auch der Aufbau eines prototypischen, idealen Argumentationsverlaufs dargelegt (*pragma-dialektische Perspektive*; vgl. van Eemeren et al. 1987; van Eemeren und Grootendorst 2004). Erst ab den 1980er Jahren entstehen dann erste empirische Arbeiten in der Argumentationsforschung, die anhand von Gesprächsdaten mündliche Argumentationen deskriptiv zu beschreiben suchen und dabei auf Struktur der Argumentationen fokussieren (vgl. u. a. Klein 1980; Miller 1980, 1984). Abgrenzend von diesen *strukturbezogenen* Forschungszugängen, wie Grundler (2011) sie fasst, wird im Folgenden eine zusammenfassende Einordnung des *prozessbezogenen* konversationsanalytischen Zugangs zum mündlichen Argumentieren gegeben (vgl. Kap. 2.1), der die Rekonstruktion der *interaktiven* Entfaltung von mündlichen Argumentationen ermöglicht und dem die vorliegende Studie folgt. Anschließend wird dargelegt, inwiefern es sich bei mündlichen Argumentationen um eine kommunikative Gattung handelt (vgl. Kap. 2.2), wie mündliche Argumentationen sequenziell organisiert sind (vgl. Kap. 2.2.1) und mithilfe welcher pragmatischen Mittel bzw. welcher argumentativen Handlungen und sprachlichen Ressourcen die Interagierenden den argumentationsauslösenden Diszens interaktiv bearbeiten (vgl. Kap. 2.2.2). Insgesamt fungiert das vorliegende Kapitel also als Herleitung und Plausibilisierung des konversationsanalytischen Zugriffs, mit dem die Studie die zugrundeliegenden Gesprächsdaten der mündlichen Argumentationen untersucht.

2.1 Konversationsanalytischer Zugriff auf das Argumentieren

Bis in die 1980er Jahre hinein wird das Argumentieren u. a. aus der rhetorischen und pragma-dialektischen Perspektive v. a. theoretisch modelliert und konzeptualisiert. Erst in den 1980er Jahren entstehen im Rahmen der Argumentationsforschung einzelne empirische Arbeiten auf Grundlage von Gesprächsdaten. Diese ermöglichen, sowohl die interaktive Funktion als auch den Verlauf mündlicher Argumentationen und die Beziehung einzelner Äußerungen, die sich zu Argumenten entwickeln können, sichtbar und nachvollziehbar zu machen (vgl. Klein 1980; Miller 1980, 1984). Klein (1980: 19) stellt heraus, dass die Funktion von Argumentationen darin besteht, etwas Strittiges bzw. kollektiv Fragliches (die *Quaestio*) in kollektiv Geltendes zu überführen. Dabei offenbart die Arbeit mit realen Gesprächsdaten hinsichtlich des interaktiven Verlaufs, anders als in theoretischen Modellierungen von Argumentationen, dass die Interagierenden ihre argumentativen Äußerungen erst nach und nach „in scheinbar anarchischer Weise“ (Klein 1980: 16) zu einer Argumentation verknüpfen. In dem Prozess der Argumentationen können dann Argumente unkommentiert fallengelassen werden, es kann zu Verschiebungen der *Quaestio* kommen und es wird sichtbar, dass sich auch die interaktiven Rollen von Proponent*in und Opponent*in verändern können. Mündliche Argumentationen weisen in ihrer Struktur sowie in ihrem

Verlauf also einen äußerst hohen Grad an Komplexität auf (vgl. Klein 1980: 16), den erst deskriptive Untersuchungen auf Grundlage empirisch erhobener Daten aufdecken können.

Wie allerdings mündliche Argumentationen tatsächlich in sozialen Interaktionen *entstehen*, wie das (sprachliche) Handeln der Beteiligten sequenziell, d. h. Redezug-um-Redezug, *organisiert* ist, um sich gegenseitig ein kommunikatives Problem anzuzeigen und es zu bearbeiten, diese Fragen fokussieren v. a. soziologisch geprägte konversationsanalytische Arbeiten (vgl. u. a. Eisenberg und Garvey 1981; Jacobs und Jackson 1982; Maynard 1985; Goodwin und Goodwin 1987; Goodwin 1990).² Sie rekonstruieren die schrittweise Entfaltung von Argumentationen in (Alltags-)Interaktionen (vgl. Kap. 2.2.1) mit dem Ziel,

die konstitutiven Prinzipien und Mechanismen zu bestimmen, die im situativen Vollzug und Nacheinander des Handelns die sinnhafte Strukturierung und Ordnung eines ablaufenden Geschehens und der Aktivitäten [hier: der Argumentation], die dieses Geschehen ausmachen, erzeugen (Bergmann 2001: 919, Anm.: N. S.; vgl. auch Goodwin und Heritage 1990).

Es zeigt sich, dass Argumentationen in (Alltags-)Interaktionen entstehen können, wenn Äußerungsinhalte, Geltungsansprüche oder Wissensbestände problematisiert und zur weiteren Aushandlung (erneut) in das Gespräch eingebracht werden (vgl. Eisenberg und Garvey 1981; Spranz-Fogasy 2003, 2005a; Heller 2012). Durch die Anzeige eines kommunikativen Problems (vgl. Günthner und Knoblauch 1996), wie beispielsweise die Strittigkeit eines Gesprächsgegenstandes, wird der Fortgang der aktuellen Aktivität zum Erliegen gebracht. Bevor die Aktivität fortgesetzt werden kann, ist es erforderlich, dass die Interagierenden zuerst den infrage stehenden Gesprächsgegenstand bearbeiten (vgl. Eisenberg und Garvey 1981: 151). Interagierende können beispielsweise eigene Positionen einbringen sowie auch Begründungen oder Einwände (vgl. Kap. 2.2.2) und dadurch eine *gemeinsame* Bearbeitung des Strittigen ermöglichen – Sprecher*innen können sich den Begründungshandlungen aber auch (vorerst) entziehen und ihre jeweiligen Positionen ausschließlich durch opponierende Äußerungen in einer Art wechselseitigem „Aufschlagwechsel“ (Knoblauch 1995: 121) verteidigen. In diesem Fall wird eine gemeinsame Arbeit an dem kommunikativen Problem verhindert und die Interaktion kann in einem Streit münden (vgl. Knoblauch 1995: 124; vgl. diesbezüglich auch Arendt 2019: 328f.).³ Wie die Interagierenden das kommunikative Problem bearbeiten, kann dementsprechend auch einen Einfluss auf den Modus der Argumentation nehmen: Dieser kann sich zwischen einer persuasiven, eher konfrontativen und einer kollaborativ-explorierenden Aushandlung bewegen (vgl. Mercer 2009; Ehlich 2014). Konstitutiv für mündliche Argumentationen ist nun nicht nur, dass ein Dissens in der Interaktion *angezeigt* wird, sondern dass die eingebrachten Positionen durch Begründungshandlungen für die anderen Beteiligten *nachvollziehbar gemacht* werden und somit auch weiter *verhandelbar* sind (vgl. zur sequenziellen Organisation

² Mit der aufkommenden funktionalen Pragmatik entstehen auch weitere Ansätze zur Untersuchung von Argumentationen, die u. a. in Ablaufdiagrammen sichtbar machen, welche Handlungsalternativen Interagierenden zu entsprechenden Zeitpunkten in der Interaktion zur Verfügung stehen (vgl. Ehlich und Rehbein 1986; vgl. auch Trautmann 2004). Aufgrund des konversationsanalytischen Zugriffs der vorliegenden Arbeit, der an der konkreten Entfaltung argumentativer Interaktionen interessiert ist, wird der Ansatz der funktionalen Pragmatik an dieser Stelle nicht weiter elaboriert.

³ Grundsätzlich kann für die Definition eines Streits nach Knoblauch (1995: 124) angeführt werden, dass dieser droht, „wenn beide Parteien im Hin und Her des Dissens bekunden, daß sie keine Begründung liefern“ und somit ein gemeinsames Arbeiten an dem kommunikativen Problem ablehnen. Morek (2015: 39) zeigt für opponierende Sequenzen in argumentativen Interaktionen von Jugendlichen, dass der Aufschlagwechsel von den Beteiligten auch fallengelassen werden kann und anschließend „versendet“. Ein Streit, wie bei Knoblauch (1995) beschrieben, ist daher keine obligatorische Konsequenz eines Aufschlagwechsels. Insgesamt zeigt sich eine Vielfalt verwendeter Begriffe in der Forschung (wie *Diskussion*, *Disput*, *Streit*, *Argumentation* oder *Konflikt(-gespräch)*), die z. T. als verschiedene Spielarten von Argumentationen verstanden werden können (vgl. auch Deppermann 2003).

Kap. 2.2.1). Dabei handeln die Beteiligten die Gültigkeit⁴ der eingebrachten Positionen und Begründungen *interaktiv* und schrittweise, d. h. *turn-by-turn*, aus.

Im Kontrast zu theoretisch konzeptualisierten Argumentationsmodellen wird durch konversationsanalytische Studien nicht nur rekonstruierbar, wie mündliche Argumentationen in Alltagsinteraktionen interaktiv entstehen. Es wird auch sichtbar, dass als obligatorisch angenommene Merkmale, wie das Einbringen von Beispielen, um eine Behauptung zu stützen, in mündlichen Argumentationen genauso ausbleiben kann wie explizit-sprachliche Markierungen, die eine Äußerung in ihrer Funktion als Begründung erkennbar werden lassen (vgl. Deppermann 2003; Morek 2015). Denn

Argumentationsindikatoren verhalten sich zur Argumentationsstruktur insofern wie Kohäsionsmarkierungen zur Kohärenz: Sie sind nur selten obligatorische, meist optionale, vielfach kontextabhängige und unterbestimmte sprachliche Oberflächenmarkierungen, die den Ausgangspunkt für Inferenzprozesse zur Rekonstruktion der Argumentationsstruktur bilden können (Deppermann 2003: 17).

D. h., dass in den argumentationstheoretischen Modellen (idealiter) angenommene Vorkommen bestimmter Strukturelemente sowie ihre eindeutige Identifizierbarkeit in mündlichen Argumentationen nur selten gegeben ist. Diese Vagheit und die daraus resultierenden Inferenzprozesse sind typisch für mündliche Interaktionen.

Konversationsanalytische Arbeiten rekonstruieren nicht nur die Struktur mündlicher Argumentationen im Gespräch, sondern sie zeigen auch auf, dass sich diese je nach *Kontext* in ihren strukturellen Eigenschaften sowie in ihrer Funktion und in ihrer sprachlich-diskursiven Ausgestaltung verändern können. Dabei stellen die Interagierenden reflexiv, d. h. *durch* ihre Handlungen selbst eine kontextuelle Rahmung her, in der ihr sprachliches Handeln zu verstehen ist, und zugleich wird *in* ihrem Handeln wiederum selbst eine „Kontextorientierung“ sichtbar (Bergmann 2001: 921; vgl. Garfinkel 1967). Um die o. g. kontextuellen Unterschiede mündlicher Argumentationen zu veranschaulichen, seien hier zwei von ihnen exemplarisch und kontrastierend gegenübergestellt: Argumentationen in Peer-Interaktionen sowie im Unterricht. Während Argumentationen in Peer-Interaktionen häufig von den Interagierenden genutzt werden, um soziale Rollen und Gruppenidentitäten auszuhandeln (vgl. u. a. Goodwin 1990; Morek 2015), werden mündliche Argumentationen im Unterricht für Wissenskonstruktionsprozesse relevant (vgl. Quasthoff et al. 2021a). Die kontextuellen Unterschiede erstrecken sich dann auch auf die Ebene der sprachlich-diskursiven Ausgestaltung. Denn während in Peer-Interaktionen beispielsweise eine explizitsprachlich markierte Opposition durch eine emphatisch aufgeladene Negation der Bezugsäußerung ausgedrückt werden kann und diese Praktik in dem Kontext als angemessen gilt (vgl. u. a. Goodwin und Goodwin 1987; Morek 2015), würde diese Form der nicht-abgemilderten Opposition von Jugendlichen im Unterrichtsgespräch gegenüber der Lehrkraft sicherlich als unangemessen behandelt werden. Hier werden Heranwachsende vielmehr zur Plausibilisierung von Positionen explizit aufgefordert (vgl. Heller und Morek 2015; vgl. zum Argumentieren als Unterrichtsgegenstand Mundwiler et al. 2017; Morek 2017), die es erforderlich machen, dass Sprecher*innen einzelne Äußerungen zu einer (satzübergreifenden) Begründung aufbauen. Diese Aufforderung zu einer Begründung können direkt mit einem globalen

⁴ Der Begriff der ‚Gültigkeit‘ entspringt nicht genuin einem konversationsanalytischen Verständnis. Denn ob ein Argument vom Gegenüber akzeptiert wird oder eine Position noch weiterer Begründungen bzw. Elaborierung bedarf, zeigen sich die Interagierenden im Gespräch gegenseitig an (vgl. Heller 2012). Die Begrifflichkeit wird an dieser Stelle jedoch gewählt, um sie mit Zugängen zu kontrastieren, in denen die Geltung von Argumenten u. a. vom Publikum (vgl. Perelman und Olbrechts-Tyteca 1958/2004) oder dem schematischen Aufbau von Argumenten, i. S. des *data*, *warrant*, *backing* und *rebuttal* (vgl. Toulmin 1958), durch eine von außen angelegte Norm abgeleitet wird.

Zugzwang expliziert werden, wie ein Beispiel aus Heller und Morek (2015: 13) zeigt, in dem die Lehrkraft fragt „welche LÖ:sung (1.4) erscheint (--) uns !SINN!voll; (1.5)“ und direkt anschließt „↑MIT begründung;“. Exemplarisch verweisen die angeführten Kontexte auf grundsätzliche Befunde konversationsanalytischer Studien zum mündlichen Argumentieren: Alters- und Beziehungsstrukturen der Beteiligten (Peers vs. Schüler*innen und Lehrkräfte) sowie der kontextuelle Rahmen (Clique vs. Unterricht) sind bedeutsam für die Vielfalt der Ausgestaltung von Argumentationen in Alltagsinteraktionen. In diesen verhandeln Interagierenden nicht nur divergierende Geltungsansprüche, Problemstellungen sowie konfligierende Wissensbestände, sondern z. T. auch Querliegendes wie soziale Identitäten und (familiäre) Rollen (vgl. u. a. Goodwin und Goodwin 1987; Goodwin 1990; Goodwin und Kyratzis 2007; Blum-Kulka 2008; Morek 2015; Goodwin und Cekaite 2018), epistemische Autoritäten und Verantwortungen (vgl. u. a. Morek 2015; Heller 2018a) sowie Rangordnungen (vgl. u. a. Maynard 1985; Komor 2010).

Zusammenfassend kann an dieser Stelle festgehalten werden, dass die Rekonstruktion mündlicher Argumentationen in Alltagsinteraktionen anhand eines konversationsanalytischen Zugriffs ermöglicht, ihre Entstehung *in situ*, ihren Verlauf sowie die interne Struktur und die eingesetzten sprachlichen Mittel abhängig vom jeweiligen Kontext aufzudecken. Eine Untersuchung anhand eines rein theoretisch fundierten Argumentationsmodells würde „Argumentationen in natürlichen Gesprächen i.Allg. als regelverletzend, unvollständig und defizitär erscheinen [lassen]“ (Morek et al. 2017: 15), da beispielsweise stützende Belege (vgl. Toulmin 1958), die aus argumentationstheoretischer Perspektive zur Absicherung eines Argumentes notwendig sind, in mündlichen Argumentationen entweder nicht realisiert werden oder auf sprachlicher Ebene implizit verbleiben. Zum anderen beruhen rhetorische sowie dialektische argumentationstheoretische Modelle

auf problematischen sprechakttheoretischen Annahmen und restringieren die möglichen Züge der Dialogteilnehmer, entkontextualisieren und postulieren Argumentationsvoraussetzungen [...], die empirisch nicht anzutreffen sind (Deppermann 2003: 12).

Dies meint, dass neben der Annahme eines ‚idealen‘ Argumentationsverlaufs auch die konkret vorliegende Interaktionssituation bzw. der Kontext ausgeklammert werden. Aus konversationsanalytischer Forschungsperspektive ist aber die Berücksichtigung des Kontextes, beispielsweise ob eine Argumentation unter Peers oder im schulischen Rahmen zwischen Lehrkräften und Schüler*innen oder auch zwischen Eltern und Kindern stattfindet, notwendig, da dieser u. a. Einfluss auf die Beteiligungsstrukturen der Sprecher*innen, die Aufrechterhaltung eines Dissenses aufgrund sozialer, familialer oder institutioneller Rollen etc. nehmen kann (vgl. u. a. Heller 2012; Morek 2015; vgl. oben). In der vorliegenden Studie wird daher der familiäre Kontext, in dem die mündlichen Argumentationen stattfinden, in der Interpretation der empirischen Forschungsbefunde mit berücksichtigt. Anknüpfend an die in diesem Kapitel vorgenommene globale Einordnung mündlicher Argumentationen folgt nun eine eingehende Zusammenschau konversationsanalytischer, rekonstruktiv erarbeiteter Forschungsbefunde, die u. a. die sequenzielle Entfaltung mündlicher Argumentationen sowie die dabei eingesetzten sprachlichen Mittel beleuchtet.

2.2 Mündliches Argumentieren als kommunikative Gattung & diskursive Praktik

Im vorangegangenen Kapitel 2.1 wurde einführend die Relevanz eines empirischen Zugriffs auf mündliche Argumentationen sowie die daraus resultierende Notwendigkeit der Arbeit mit realen Gesprächsdaten dargelegt. Das folgende Kapitel 2.2 dient nun der Klärung, inwiefern das mündliche Argumentieren als eine kommunikative Gattung bzw. diskursive Praktik verstanden werden kann. Anschließend wird dargelegt, wie mündliche Argumentationen sequenziell organisiert sind (vgl. Kap. 2.2.1) und welche pragmatischen *Mittel* Sprecher*innen zur Bearbeitung der jeweiligen Gesprächsaufgaben (*Jobs*) einsetzen sowie durch welche Ressourcen diese realisiert werden können (vgl. Kap. 2.2.2). Insgesamt kommt diesem Kapitel die Funktion zu, ein theoretisches Fundament für die zwei anschließenden Kapitel zur Körperlichkeit des Argumentierens zu legen, in denen gezeigt wird, dass Interagierende nicht nur sprachliche, prosodische und lexikalische, sondern auch körperliche Ressourcen für die Realisierung semantisch-pragmatischer Mittel einsetzen (vgl. Kap. 3). Die konversationsanalytische Einordnung von Argumentationen als kommunikative Gattung soll nachvollziehbar machen, welche komplexen diskursiven Anforderungen das Argumentieren an die Interagierenden stellt, wenn sie ein interaktives Problem argumentativ bearbeiten. Somit werden die Erkenntnisse dieses Kapitels auch relevant für die Modellierung argumentativer Diskurskompetenz (vgl. Kap. 4).

Kommunikative Gattung

Das Argumentieren wird als eine kommunikative Gattung verstanden, die einen gewissen Grad an interaktiver und struktureller Komplexität aufweist (vgl. Bergmann 2001; Deppermann et al. 2016). Ursprünglich wurde der Begriff der *Gattungen* lange Zeit vornehmlich der Schriftlichkeit und hier vor allem der Literaturwissenschaft zugeordnet, in denen Gattungen „Typen literarischer Texte wie ‚Novelle‘, ‚Briefroman‘, ‚Ballade‘, etc. [bezeichnen]“ (Auer 2013: 182). William Labov überträgt ihn dann in den 1960er Jahren auf mündliche Erzählungen, wodurch der Gattungsbegriff nun auch für mündliche ‚Texte‘ fruchtbar gemacht (vgl. Auer 2013: 182) und dann durch Luckmann (1986: 196) aus einer sensussoziologischen Perspektive in die Terminologie der „kommunikativen Gattungen“ überführt wird. Konstituierend für kommunikative Gattungen ist, dass sie sich im Unterschied zu „spontanen“ kommunikativen Handlungen“ an „vorgefertigten Mustern“ orientieren (Luckmann 1986: 201)⁵ und „mehr oder minder wirksame und verbindliche ‚Lösungen‘ von spezifisch kommunikativen ‚Problemen‘ [sind]“ (Luckmann 1986: 202; vgl. auch Bergmann und Luckmann 1995; Günthner und Knoblauch 1996). Ein interaktiv manifest gewordenes kommunikatives Problem wird also mithilfe gesellschaftlich verfestigter Problemlöseverfahren, die wiederkehrende Musterhaftigkeiten aufweisen, von den Interagierenden gemeinsam bearbeitet. In dieser Weise können beispielsweise auftretende Wissensasymmetrien mithilfe von *Erklärungen* (vgl. u. a. Morek 2012; Kinalzik 2024), erzählwürdige Ereignisse mit *Erzählungen* (vgl. u. a. Quasthoff 1980; Ohlhus 2014) und auftretende Dissense mithilfe von *Argumentationen* (vgl. u. a. Heller 2012; Kreuz 2021) bearbeitet werden:

Was immer wir kommunikativ tun, wir verwenden dabei in der Regel bestimmte Formen, die wir nicht selbst erfunden haben, sondern die uns als Teil unseres Wissensvorrats zur Verfügung stehen. Damit die Handelnden von den Routineproblemen der Kommunikation ‚entlastet‘ werden, stehen ihnen vorgeprägte kommunikative Formen zur Verfügung. Sind diese Formen fest vorgeprägt und weisen einen bestimmten Komplexitätsgrad auf, so reden wir von *Gattungen* (Günthner und Knoblauch 1996: 36, Hervorh. i. O.).

⁵ Vgl. zu einer umfassenden Auseinandersetzung mit der Terminologie der *Muster* den Sammelband *Sprachliche und kommunikative Praktiken* von Deppermann et al. (2016).

In kommunikativen Gattungen, die in ihrer sequenziellen Struktur einen gewissen Grad an Komplexität bzw. Globalität aufweisen (vgl. Knoblauch 2021; Mazeland 2021), werden also routinisierte, (mehr oder weniger) verfestigte Strukturen sprachlicher Handlungen in Form von „*Ensembles* konventionell situierter, obligatorischer konversationeller Handlungszüge“ (Deppermann et al. 2016: 12f., Hervorh. N. S.) erkennbar.⁶ Auch mündliche Argumentationen, durch die die Interagierenden ein in der Interaktion manifest werdendes Problem oder ein Dissens mithilfe gesellschaftlicher verfestigter Strukturen bearbeiten, gelten aufgrund ihrer Komplexität als kommunikative Gattung. Aus den Untersuchungen zu kommunikativen Gattungen geht nun hervor, dass diese in der Alltagsinteraktion nicht unverbunden *nebeneinander* existieren, sondern sie in der Interaktion durchaus *füreinander* funktional eingesetzt werden können (vgl. Goodwin und Goodwin 1987). So können beispielsweise Erzählungen argumentativ gerahmt sein, wie eine Arbeit von Lucius-Hoene und Deppermann (2002) zeigt, oder *Belegerzählungen* funktional in Begründungssequenzen mündlicher Argumentationen eingesetzt werden (vgl. Quasthoff 1980; Heller 2012: 86).

Untersucht werden können die Gattungen hinsichtlich der von den Interagierenden eingesetzten lexiko-semanticen, morpho-syntaktischen, prosodischen, mimischen sowie weiteren Ressourcen (*Binnenstruktur*), mit denen einzelne, kommunikativ verfestigte Handlungen aufgebaut werden (*strukturelle Zwischenebene*). Diese wiederum formen die jeweilige Gattung (*Außenstruktur*), wie Günthner und Knoblauch (1996: 41ff.) mit Verweis auf Luckmann (u. a. 1986) darlegen. Die analytische Dreiteilung von Binnen- und Außenstruktur sowie der strukturellen Zwischenebene ähnelt u. a. dem Beschreibungsinstrument von Hausendorf und Quasthoff (1996/2005), mit dem kommunikative Gattungen im interaktiven Vollzug untersucht werden können und das in der vorliegenden Arbeit in Kapitel 2.2.1 vorgestellt wird.

Aus der Wissenssoziologie für die Konversationsanalyse fruchtbar gemacht, fokussiert der vorgestellte Gattungsbegriff also insgesamt auf die „soziale Verfestigung“ (Quasthoff et al. 2021b: 39) der gesellschaftlich zur Verfügung stehenden Problemlösungsverfahren. Mit der Einführung des Begriffs der *diskursiven Praktik* soll im Nachfolgenden der Fokus leicht verschoben und beleuchtet werden, inwiefern sich das mündliche Argumentieren als globale Aktivität *interaktiv* vollzieht (vgl. Quasthoff et al. 2021b: 39) und somit das kommunikative Problem divergierender Geltungsansprüche von den Interagierenden durch Rückgriff auf routinisierte Handlungen interaktiv gelöst werden kann (vgl. u. a. Heller 2012; Quasthoff et al. 2017, 2021a). Dazu werden, aus konversationsanalytischer Perspektive, ein lokales Praktikenverständnis von einem globalen abgegrenzt.

Diskursive Praktik

Ein besonderes Erkenntnisinteresse der Konversationsanalyse liegt in der Frage, wie eigentlich die Ordnung in der sozialen Interaktion entsteht und welche Mittel die Beteiligten einer Interaktion zur Herstellung dieser Ordnung einsetzen. Dabei erweisen sich die von den Interagierenden eingesetzten „*Gesprächspraktiken* [...] als Grundbausteine der sozialen Interaktion“ (Deppermann et al. 2016: 1, Hervorh. N. S.). Mittlerweile haben sich zwei voneinander abzugrenzende Praktikenbegriffe herausgebildet, die sich hinsichtlich ihrer Lokalität bzw. Globalität unterscheiden:

⁶ Vgl. zur Problematisierung des Begriffs der *Verfestigung* auch Knoblauch (2021: 20), der die Frage aufwirft, wie damit umzugehen ist, „wenn sich die Verfestigung lediglich auf bestimmte Aspekte kommunikativer Abläufe bezieht“, und anschließend für weniger stark verfestigte kommunikative Formen den Begriff der *Muster* vorschlägt (vgl. Knoblauch 2021: 24).

- Mikrostrukturelle, *lokal* organisierte Praktiken bewegen sich strukturell auf den Ebenen der Einzeläußerungen und Paarsequenzen. Es handelt sich um jene routinisierten sprachlichen *Praktiken*, mit denen Interagierende Redezug-um-Redezug *Handlungen* (*actions*, vgl. Schegloff 1997) und somit soziale Interaktion aufbauen.⁷ So kann beispielsweise die formelhafte Phrase *Guten Tag* die Handlung des Begrüßens vollziehen, auf die im Rahmen einer Paarsequenz, d. h. einer Sequenz, in der erwartbare Äußerungsteile aufeinander folgen, nun ein Gegengruß z. B. durch Wiederholung der Phrase (*Guten Tag*) durch das Gegenüber erfolgen.
- Von den lokal organisierten Praktiken kann ein makrostrukturelles Begriffsverständnis abgegrenzt werden, das globalere, d. h. äüßerungs- und paarsequenzübergreifende Praktiken fokussiert. Der Begriff der *diskursiven Praktiken* umfasst also „semiotische, interaktive Großformen des Sprachgebrauchs“ (Deppermann et al. 2016: 12), wie beispielsweise das Erzählen oder Argumentieren, die jeweils eigene Formen und Funktionen aufweisen. So werden Erzählungen von einem*einer primären Sprecher*in um ein erzählwürdiges Ereignis herum aufgebaut, auf das sich die Erzählung hinbewegt (vgl. Quasthoff 1980; Ohlhus 2014). In Argumentationen wird hingegen beispielsweise ein strittiger Gesprächsgegenstand durch wechselseitig eingebrachte Positionen, Begründungen und Einwände von mehreren Beteiligten bearbeitet (vgl. Heller 2012). Beide Diskurspraktiken unterscheiden sich dabei in Form und Funktion. Abgrenzend von den oben beschriebenen lokalen Praktiken handelt es sich i. d. S. also um ein Begriffsverständnis, das sich auf die Makrostruktur einer Praktik bezieht. In jüngsten Arbeiten weisen Quasthoff et al. (2021b: 39) darauf hin, dass die Begriffe von *diskursiver Praktik* und *Gattung* „synonym verwendet werden [können], wobei ersterer auf den interaktiven Vollzug abhebt und letzterer stärker die soziale Verfestigung betont“.

Die vorliegende Arbeit nutzt das Verständnis der global organisierten, gattungstheoretischen Konzeptualisierung und wird den Begriff der *diskursiven Praktiken* „als Oberbegriff für interaktiv vollzogene sprachliche Aktivitäten [verwenden], die global organisiert sind [...] und die oftmals den Status verfestigter kommunikativer Gattungen bzw. Genres haben“ (Quasthoff et al. 2021a: 16). Dabei greifen Interagierende auf „diskursive[], gattungsförmig verfestigte[] Praktiken“ (Quasthoff et al. 2021a: 13), wie das Argumentieren, zurück, um globale argumentative Zusammenhänge zu versprachlichen (vgl. Morek und Quasthoff 2018).

Der terminologischen Klärung der Begriffe der kommunikativen Gattung sowie der diskursiven Praktik folgt nun eine Einordnung, wie mündliche Argumentationen sequenziell organisiert sind (vgl. Kap. 2.2.1) und welche pragmatischen Mittel Sprecher*innen dabei (wiederkehrend) einsetzen (vgl. Kap. 2.2.2).

2.2.1 Sequenzielle Organisation mündlicher Argumentationen

In mündlichen Argumentationen verhandeln Interagierende gemeinsam divergierende Geltungsansprüche oder manifest gewordene Probleme. In der Untersuchung dieser Aushandlungsprozesse erweist es sich nun als analytisch notwendig, die *makrostrukturellen*, d. h. die globalen kommunikativen Aufgaben, die die Interagierenden *gemeinsam* bearbeiten, sowie die Anforderungen an die *einzelnen* Sprecher*innen, i. S. einzusetzender pragmatischer und semantischer *Mittel*, die durch sprachliche, prosodische und paraverbale *Formen* realisiert werden, getrennt voneinander zu beschreiben. Für diese analytische Trennung der global organisierten Aufgaben sowie der dafür eingesetzten Mittel und Formen

⁷ Der Begriff der Praktiken lehnt sich an den vergleichbaren Ethno-Methoden der Ethnomethodologie an (vgl. Garfinkel 1967).

greift die vorliegende Untersuchung auf das von Hausendorf und Quasthoff (1996/2005) für den Diskurserwerb des Erzählens entwickelte Modell GLOBE⁸ zurück, das von Heller (2012) für das Argumentieren weiterentwickelt worden ist und dessen drei Analyseebenen der *Jobs*, *Mittel* und *Formen* in den Kapitel 2.2.1 und 2.2.2 vorgestellt werden.

In der vorliegenden Arbeit werden mündliche Argumentationen in ihrer globalen, makrostrukturellen Organisation als *Diskurseinheiten* (vgl. Wald 1978; vgl. auch Quasthoff et al. 2017: 86) verstanden, die aus dem *turn-by-turn talk* heraus entstehen. Gemeint sind zusammengehörige, paarsequenzübersteigende Äußerungen, die „are produced and recognized as coherent building blocks in conversation because participants themselves mark off their beginning and ending from the surrounding turn-by-turn talk“ (Quasthoff et al. 2017: 86). Diese treten beispielsweise auf, um ein im *turn-by-turn talk* entstandenes kommunikatives Problem wie einen Dissens zu lösen. Dazu verlassen die Sprecher*innen den *turn-by-turn talk* zwar (kurzfristig), jedoch wird der Sprecher*innenwechsel hier nicht, wie etwa beim Erzählen (vgl. Hausendorf und Quasthoff 1996/2005) oder Erklären (vgl. Morek 2012; Kinalzik 2024), ausgesetzt, weil einem*einer Beteiligten die Rolle des*der primären Sprecher*in zukommt, sondern der Sprecher*innenwechsel wird aufrechterhalten und die Beteiligten nehmen stetig aushandelnd die Rollen von Proponent*in und Opponent*in bzw. Problematisierer*in und Problemlöser*in ein (vgl. Heller 2012). Um den entstandenen Dissens bzw. das Problem, der/das begründungspflichtig gemacht worden ist, zu bearbeiten, setzen die Interagierenden selbst die sequenziellen gesprächsstrukturellen Aufgaben relevant. D. h. sie machen einander begründungspflichtig und fordern z. B. globale, d. h. einen Redezug übersteigende, argumentative Äußerungspakete wie Begründungen voneinander ein oder bringen lokal operierende Beiträge wie Widersprüche ein (vgl. Quasthoff et al. 2021a: 17). Dabei sind die sich entfaltenden Argumentationen in „übergeordnete Handlungsaufgaben“ (Spranz-Fogasy 2003: 31) bzw. in den *turn-by-turn talk* eingebettet – sie entstehen aus ihm und „[müssen] in diesen wieder eingegliedert werden“ (Heller 2012: 67). Heller (2012) rekonstruiert in ihrer konversationsanalytischen Untersuchung mündlicher Argumentationen in familialen Tischgesprächen die folgenden kontextfreien gesprächsstrukturellen Aufgaben, die von den Interagierenden in der Diskursaktivität des Argumentierens bearbeitet werden:

- (1) Dissens herstellen bzw. Problematisieren,
- (2) Etablieren von Begründungspflicht,
- (3) Kernjob Begründen,
- (4) Abschließen und
- (5) Überleiten.

Innerhalb des *Herstellens des Dissenses* bzw. *Problematisierens* werden divergente Geltungsansprüche sichtbar bzw. ein Gesprächsgegenstand als ‚problematisch‘ angezeigt (vgl. Heller 2012: 68). Ob sich aus der Problematisierung bzw. dem Dissens überhaupt eine argumentative Bearbeitung ergibt, zeigt sich erst, wenn die Sprecher*innen eine *Begründungspflicht etablieren* (vgl. Heller 2012) oder selbst in den Kernjob des *Begründens* einsteigen. Interagierende können auch sich selbst dazu verpflichten, eine Begründung für eine Position oder eine Behauptung zu liefern, ohne dass das Gegenüber eine Begründungspflicht etabliert hat (vgl. Heller 2012; Arendt 2019). Mit dem Etablieren der Begründungspflicht

⁸ Das Analyseinstrument GLOBE (**G**lobalität und **L**okalität in der **O**rganisation **b**eidseitig-konstruierter **E**inheiten) wurde von Hausendorf und Quasthoff (1996/2005) im Rahmen einer Erwerbsstudie zum Erzählen entwickelt. Das Modell ermöglicht aufgrund der beschriebenen analytischen Trennung der Makro- und Mikrostruktur, die „Analyse von Gesprächskompetenz als Anteil *eines* der Beteiligten an der *gemeinsamen* Herstellung von Sinn im Rahmen der jeweiligen Gesprächsaktivität“ (Quasthoff 2009: 86, Hervorh. N. S.). Das Beschreibungsinstrument konnte bereits erfolgreich für das Erklären (vgl. Morek 2012; Kinalzik 2024) sowie das Argumentieren (vgl. Heller 2012; Kreuz 2021) adaptiert werden.

wird der *turn-by-turn talk* endgültig verlassen. Im Kernjob des Begründens wird dann der Dissens oder das aufgetretene Problem verhandelt, indem die Interagierenden beispielsweise Belegerzählungen für ihre Position begründend anführen, ein Argument widerlegen oder sich gegenseitig beipflichten (vgl. Heller 2012: Kap. 5.3; vgl. Kap. 2.2.2). Wenn die Interagierenden sich dann gegenseitig ratifizierend anzeigen, dass sie „die Begründungspflicht als erfüllt ansehen oder aus anderen Gründen weitere argumentative Behandlungen abbrechen wollen“ (Heller 2012: 95), kommt es zum *Abschließen* des Kernjobs. Dabei weist das Modell in den Jobs des Etablierens der Begründungspflicht und des Abschließens eine Komplementarität auf (vgl. Heller 2012: 68; vgl. zum Erzählen Hausendorf und Quasthoff 1996/2005: 128), denn diese Jobs leiten als sogenannte Randjobs (vgl. Morek et al. 2017) in den Kernjob des Begründens ein beziehungsweise aus diesem wieder heraus. Sowohl der Kernjob Begründen als auch die gesamte Diskurseinheit der Argumentation werden durch das Abschließen beendet und die Interagierenden kehren dann mit dem Job des *Überleitens* wieder in den *turn-by-turn talk* zurück bzw. treten in andere Aktivitäten ein (vgl. Heller 2012: 68).

Das Bearbeiten der gesprächsstrukturellen Aufgaben verläuft dabei gemäß einer „hochgradig erwartbaren Regularität“, wie Hausendorf und Quasthoff (1996/2005: 127) für das Erzählen konstatieren (vgl. zum Argumentieren Heller 2012 sowie zum Erklären Morek 2012). Die globalen und sequenziellen Erwartungen, wie das Begründen einer eingebrachten Position, werden von den Sprecher*innen selbst etabliert und im Weiteren (gemeinsam) bearbeitet. Kommen Interagierende diesen etablierten „globalen sequenziellen Verpflichtungen“ beziehungsweise Erwartungen nicht nach und verweigern beispielsweise das Erbringen einer eingeforderten Begründung, „muss diese Verweigerung begründet“ (Quasthoff et al. 2019: 195), d. h. *accountable* (vgl. Garfinkel 1967) und für das Gegenüber nachvollziehbar gemacht werden. Dabei kann das Erfüllen bzw. Verweigern einer sequenziellen Erwartung auch weitere, u. a. soziale Funktionen erfüllen (vgl. Kap. 2.1), d. h.

[d]ie Art und Weise, in der die gattungsspezifischen Jobs jeweils bearbeitet werden, ist dabei nicht immer gleich, sondern verlangt den Zuschnitt der genutzten Mittel und Formen auf die kommunikativen Anforderungen eines bestimmten sozialen und kommunikativen Kontextes (Heller 2012: 17).

Insgesamt ermöglicht das Beschreibungs- bzw. Analyseinstrument GLOBE i. S. eines *tertium comparationis* vergleichende Untersuchungen von mündlichen Argumentationen unterschiedlicher Sprecher*innen hinsichtlich der Bearbeitung der Jobs. Es folgt nun in Kapitel 2.2.2 die Vorstellung einer Auswahl der für die Arbeit relevanten pragmatischen Mittel bzw. argumentativen Handlungen und ihrer Formen, die die Sprecher*innen einsetzen, um die beschriebenen gesprächsstrukturellen Aufgaben hervorzubringen beziehungsweise zu bearbeiten.

2.2.2 Pragmatische Mittel und Formen ihrer Realisierung

Auf der Ebene der pragmatischen Mittel kann innerhalb des GLOBE-Modells untersucht werden, wie die Beteiligten die oben beschriebenen gesprächsstrukturellen Aufgaben innerhalb der Diskursaktivität des Argumentierens bearbeiten. Die Sprecher*innen realisieren dabei, wie von Heller (2012) rekonstruiert werden konnte, unterschiedliche genretypische Handlungen: So kann das Widersprechen oder Aufwerfen einer Frage einen *Dissens herstellen* oder eine Äußerung *problematisieren*; eine Ergänzungsfrage nach einer Begründung oder eine Frage nach der Wissensquelle kann *Begründungspflicht etablieren*; das Anführen von Belegerzählungen oder das Widerlegen kann im Kernjob *Begründen* eingesetzt werden; das Formulieren eines Fazits kann anzeigen, dass der infrage stehende Gesprächsgegenstand ausreichend begründet worden ist und das Begründen somit *abschließen*; das Schweigen oder eine thematische Rückbindung an die übergeordnete Interaktion kann in den *turn-by-turn talk überleiten*. Dabei

tragen die rekonstruierten pragmatischen Mittel zwar zur Konstituierung der Teilaufgaben bei, sie sind jedoch nicht notwendigerweise an einzelne Jobs gebunden.⁹ Die sprachlichen Handlungen selbst werden wiederum aufgrund der eingesetzten sprachlichen, parasprachlichen sowie aufgrund weiterer Ressourcen in ihrer jeweiligen Funktion erkennbar.¹⁰ Die pragmatischen Mittel entfalten sich dabei einerseits auf lokaler Ebene Redezug um Redezug und i. d. S. sequenziell, zugleich bauen sie aber auch gemeinsam Teilaufgaben der Diskurseinheit des Argumentierens auf (vgl. Kap. 2.2.1) und operieren somit auch auf globaler Ebene. Hausendorf und Quasthoff (1996/2005: 123) sprechen daher auch von einer „doppelseitigen Struktur“ der pragmatischen Mittel.

Die vorliegende Untersuchung legt, wie oben skizziert, die vorgestellte Dreiteilung des GLOBE-Modells (*Jobs, Mittel, Formen*) zugrunde. Dabei werden die pragmatischen Mittel in der vorliegenden Arbeit allerdings allgemeiner, d. h. mit einem größeren Grad der Spezifik, als sogenannte *argumentative Handlungen* konzeptualisiert. Diese methodische Entscheidung ermöglicht sowohl innerhalb der Analysen als auch in der Ergebnispräsentation eine bessere Nachvollziehbarkeit der rekonstruierten multimodalen Musterhaftigkeiten. Analog zu den pragmatischen Mitteln verstehe ich argumentative Handlungen als notwendig, um die gesprächsstrukturellen Aufgaben zu bearbeiten, sodass die argumentativen Handlungen innerhalb des GLOBE-Modells unterhalb der Jobebene liegen. In den argumentativen Entscheidungsfindungen der vorliegenden Datengrundlage (vgl. Kap. 1) erweisen sich nun folgende argumentative Handlungen als relevant: das *Beziehen einer Position*, *Begründen*, *Zustimmen*, *Einwenden* sowie *Einfordern von Rückmeldungen*. Sie setzen sich jeweils aus einzelnen pragmatischen Mitteln zusammen. So kann beispielsweise die argumentative Handlung des Begründens durch das pragmatische Mittel des *Illustrierens* anhand eines hypothetischen Szenarios erfolgen (vgl. Heller 2012; Morek 2020) und ggf. mit weiteren pragmatischen Mitteln kombiniert werden (vgl. z. B. zur Kombination von Belegerzählungen und Berufen auf Augenzeugenschaft Heller 2012: 86f.). Wie aus den genannten argumentativen Handlungen hervorgeht, kommt es nun innerhalb des GLOBE-Modells zu einer terminologischen Überschneidung zwischen dem Kernjob *Begründen* und dem *Begründen* als argumentative Handlung. Diese Überschneidung gilt es an dieser Stelle aufzulösen: Wie Heller (2012: Kap. 5.3) rekonstruiert, können innerhalb des Kernjobs unterschiedliche Mittel zum Begründen eingesetzt werden – sowohl proponierende Handlungen wie das Anführen von Belegerzählungen als auch opponierende Handlungen wie das Widerlegen (vgl. Heller 2012: 85). Das Begründen als *argumentative Handlung*, wie es in der vorliegenden Arbeit aufgefasst wird, umfasst hingegen das Stützen bzw. Plausibilisieren einer eigenen Position; opponierende Handlungen werden dementsprechend als Einwände aufgefasst (vgl. unten, Begründen, Einwenden).

Im Folgenden werden nun, basierend auf Studien zu den pragmatischen Mitteln, Befunde zu den ausgewählten argumentativen Handlungen des Beziehens einer Position, Begründens, Zustimmens, Einwendens und Einforderns einer Rückmeldung sowie die an ihrem Aufbau beteiligten sprachlichen und parasprachlichen Formen vorgestellt. Es handelt sich dabei um argumentative Handlungen, die in der empirischen Untersuchung der vorliegenden Arbeit (vgl. Kap. 8) frequent auftreten.

⁹ So kann beispielsweise das pragmatische Mittel des Widersprechens sowohl einen Dissens markieren als auch im Kernjob Begründen „für das Zurückweisen von Begründungen eingesetzt werden“ (Heller 2012: 91).

¹⁰ Bereits Deppermann und Lucius-Hoene (2003: 142, Fußnote 13) weisen darauf hin, „dass im geeigneten Kontext sogar mit hinweisenden Gesten argumentiert werden kann“ (vgl. dazu Kap. 3). Allerdings, so Knoblauch (2021: 24), seien in der Untersuchung kommunikativer Gattungen bislang „[i]hre Materialität und Dinglichkeit [...] ebenso wenig konzeptionell berücksichtigt [worden] wie die Körperlichkeit ihres Vollzugs“.

Position beziehen

Das Beziehen einer Position meint innerhalb einer argumentativen Entscheidungsfindung (vgl. Quasthoff und Kluger 2021) das Einbringen einer lokalen Präferenz gegenüber einer der zur Auswahl stehenden Optionen. Diese Präferenzäußerung wird in der vorliegenden Studie von den Interagierenden regelmäßig eingebracht.^{11,12} Grundler et al. (2020) untersuchen angelehnt an Gätje et al. (2012) sogenannte Positionierungshandlungen, die auch implizite Positionierungen, wie beispielsweise Einwände, einschließen. Im Unterschied zu diesem Vorgehen fasst die vorliegende Studie das Beziehen einer Position als das sprachlich markierte (bewertende) Auswählen oder Ablehnen einer der zur Wahl stehenden Optionen. Oberflächensprachlich kann das Beziehen einer Position unterschiedlich realisiert werden: So können Sprecher*innen Stellungnahmen beispielsweise durch Äußerungen wie *ich bin für/gegen X* (vgl. Gätje et al. 2012: 129) oder wie *ich würde X nehmen* einbringen. Sie können aber auch einzelne Optionen evaluieren, z. B. in der Form *Ich finde X* + (NEG) Adj. (*schön/(nicht so) gut/blöd*). Dabei können mithilfe von Bewertungen subjektive Positionen eingebracht werden, die dadurch „[...] nicht wie Tatsachenbehauptungen wahr oder falsch, sondern höchstens mehr oder weniger plausibel bzw. für bestimmte Menschen und Gruppen mehr oder weniger akzeptabel sein [können]“ (Bayer 1982: 16, zit. n. Hrmcal 2018: 238; vgl. auch Hrmcal 2020). Unabhängig davon, ob eine Stellungnahme (z. B. *ich bin für/gegen X*) oder eine Bewertung (z. B. *ich finde X schön*) eingebracht wird, weisen die damit geäußerten Positionen (vorerst) eine fehlende Plausibilisierung und somit für die anderen Interagierenden eine geringe(re) Nachvollziehbarkeit in Bezug auf die Ursachen und Gründe der Position auf (vgl. dazu in schriftlichen Texten auch Augst et al. 2007: Kap. 5).

Wie die o. g. Beispiele bereits illustrieren, kann das Beziehen einer Position mit unterschiedlichen epistemischer Modalisierungen eingebracht werden: Während die Verwendung des Konjunktivs II in Kombination mit modalisierenden Satzadverbien (*ich würde vielleicht X nehmen*) die Position des*der Sprecher*in abschwächt und sie grundsätzlich als potenziell weiter verhandlungswürdig markiert, können Positionen im Indikativ (*wir nehmen X*) epistemische Sicherheit ausdrücken und das Gegenüber über die Wahl eher informieren als es zur weiteren Aushandlung einzuladen (vgl. zu Vorschlägen u. a. Stevanovic und Peräkyä 2012; Stevanovic 2013; Couper-Kuhlen 2014; Stivers und Sidnell 2016; vgl. zu *informings* Heritage 1984). Das Einbringen und Markieren der eigenen Präferenz in Bezug auf die Auswahl einer Option kann also auch relevante Hinweise über die potenzielle Verhandelbarkeit einer bezogenen Position beinhalten.

Sequenziell können an das Beziehen einer Position als eine vorerst lokale Form der Bearbeitung der Entscheidungsaufgabe dann beispielsweise Begründungen (vgl. unten, Begründen; vgl. auch Grundler et al. 2020: 104) anschließen, die die Position stützen (vgl. Kyratzis et al. 2010; Kreuz 2021). Bleiben diese Begründungen zur Plausibilisierung der Position aus, können sie von anderen Beteiligten beispielsweise durch Ergänzungsfragen eingefordert werden (vgl. unten, Einfordern von Rückmeldungen;

¹¹ Das Einbringen einer Position in Form einer Präferenzäußerung ist auf das Setting der Datenerhebung zurückzuführen: Da die Interagierenden die Aufgabe erhielten, eine der vier Optionen begründet auszuwählen, erscheint eine explizite Benennung favorisierter bzw. abgelehnter Optionen hochgradig erwartbar.

¹² Der vorliegende Begriff des *Beziehens einer Position* ist terminologisch von Begriffen der *Positionierung* bzw. des *Positionierens* abzugrenzen. Anhand der dahinterliegenden Begriffskonzepte wird u. a. untersucht, wie sich Zugehörigkeiten zu sozialen Gruppen sowie auch epistemische Haltungen manifestieren (vgl. zsf. zu den Konzepten des Positionierens Weiser-Zurmühlen 2021: Kap. 3). Auch ist der hier verwendete Begriff des *Beziehens einer Position* von den von Gätje et al. (2012) beschriebenen Positionierungsprozeduren abzugrenzen, die i. S. des von Feilke (2012) beschriebenen Textprozeduren auf Musterhaftigkeiten in *schriftlich* verfassten argumentativen Texten fokussieren. Die Verfasserin der vorliegenden Arbeit hingegen begreift mit dem Terminus *Position beziehen* eine argumentative Handlung, die, wie Quasthoff und Kluger (2021) darlegen, von den Interagierenden als eine lokale Präferenzäußerung innerhalb des Entscheidungsdiskurses eingebracht werden kann.

vgl. Heller 2012). Dadurch zeigen sich die Interagierenden gegenseitig nicht nur das Fehlen einer Plausibilisierung an, sondern durch das Einfordern einer Begründung übernimmt auch das Gegenüber eine epistemische Verantwortung für die gemeinsame Entscheidungsfindung (vgl. Stivers et al. 2011). Das Beziehen einer Position kann sequenziell aber auch eine vorangegangene Begründung, i. S. einer resümierenden Konklusion, abschließen (vgl. Kreuz 2021; vgl. zu Textrahmungselementen in schriftlichen Argumentationen auch Augst et al. 2007: Kap. 5.4). Die argumentative Handlung wird daher sowohl für den Kernjob des Begründens relevant, wenn Interagierende eine Position initiieren, diese durch Begründungen plausibilisieren (z. B. *ich würde X nehmen, weil*) oder eine eingebrachte Begründung lokal resümierend abschließen (z. B. *deswegen würde ich X nehmen*) und somit zur Disposition stellen (vgl. Heller 2012; Kreuz 2021). Das Beziehen einer Position kann aber auch zum Job des Abschließens beitragen, indem Interagierende eine gemeinsame Position wie *dann entscheiden wir uns für X* als eine finale Entscheidung präsentieren (vgl. Heller 2012).

Begründen

In mündlichen Argumentationen ist die argumentative Handlung des Begründens notwendig, um gemeinsam das kommunikative Problem zu bearbeiten. Das der Arbeit zugrundeliegende Verständnis des Begründens erfasst Äußerungen, in denen Interagierende intersubjektiv nachvollziehbare Gründe für ihre eigene Position (vgl. oben, Position beziehen) oder eine Proposition, i. S. einer Behauptung oder eines Statements (vgl. Goetz und Shatz 1999; Kyratzis et al. 2010), nennen. Der Begriff der Proposition umfasst satzförmige Behauptungen, wie *wir nehmen X, weil es für alle das Beste ist*. Im Rahmen der Begründung können die Interagierenden zur Plausibilisierung von Positionen und Proposition dann auf geteiltes Wissen rekurren, neues Wissen etablieren oder vorhandenes Wissen neu strukturieren (vgl. zusammenfassend Kreuz 2021: 24ff.; vgl. auch Trautmann 2004; Ehlich 2014; Scarpaglieri 2017). Die Plausibilisierung der Position durch das Anführen von Gründen kann dabei beispielsweise durch Beleg Erzählungen, argumentatives Erklären, Beispielnennungen oder auch (hypothetische) Szenarios eingebracht werden (vgl. u. a. Spiegel 2003; Heller 2012; Morek 2020). Insgesamt versuchen Interagierende dadurch in mündlichen Argumentationen, „etwas kollektiv Fragliches in etwas kollektiv Geltendes zu überführen“ (Klein 1980: 19). Das Begründen als argumentative Handlung nimmt dabei häufig die Form eines größeren, d. h. satzübergreifenden Äußerungspakets an und kann oberflächensprachlich kausale Konjunktionen wie *weil* oder *deshalb*, finale Konjunktionen wie *dass* oder konditionale Konjunktionen wie *wenn* enthalten, die die Art der Begründungszusammenhänge markieren (vgl. u. a. Gohl 2006). Die Äußerungspakete können dann lokal mit der argumentativen Handlung einer resümierenden Position (vgl. oben, Position beziehen) abschließen.

In den Daten zeigt sich ebenfalls, dass Interagierende Propositionen i. S. von Statements einbringen, die sie z. B. durch kausale Konjunktionen als Begründungen markieren (z. B. *(wir X nehmen X,) denn es ist ja von der Schule her*), ohne sie weiter zu elaborieren und somit für ihr Gegenüber zu plausibilisieren. Diese werden im Folgenden als nicht-elaborierte Begründungen in die Untersuchung aufgenommen und sind den Begründungen beigeordnet. Anders als in einer funktional-pragmatischen Definition von Begründungsversuchen von Ehlich und Rehbein (1986: 94) ist für die Unterscheidung zwischen Begründung und nicht-elaborierter Begründung in der vorliegenden gesprächsanalytischen Studie nicht relevant, ob mit der Begründung eine „Veränderung im Wissen des Hörers [...] erreicht [worden] ist“, sondern inwiefern eine stützende Begründung zur intersubjektiven Plausibilisierung erfolgt oder ausbleibt.

Obwohl das Begründen selbst als der Kern des Argumentierens gilt, erweist sich das Erfassen und Beschreiben von Begründungshandlungen in mündlichen Interaktionen als durchaus schwierig. So stellt Kreuz (2021: 39) fest, dass es weder an seiner Struktur, noch an spezifischen sprachlichen Markierungen zu erkennen sein muss, da

Begründungen [...] im Mündlichen nicht zwangsläufig einen formal-logischen, vollständigen bzw. explizit verbalsprachlichen Teil eines Arguments dar[stellen], sondern [...] sich durch Unvollständigkeit, Implizitheit oder fehlende sprachliche Markierung auszeichnen [können] (vgl. hierzu auch Deppermann 2003).

Zur Rekonstruktion von Begründungen muss also der sequenzielle Kontext der Handlung sowie ihre Funktion und die (anschließende) interaktive Konsequenz, also wie das Gegenüber die Äußerung weiter be- und verhandelt, berücksichtigt werden.

Zustimmen

In den vorliegenden gemeinsamen Entscheidungsfindungen können sich Interagierende gegenseitig ihren Konsens gegenüber individuellen Positionen (z. B. in Form von Vorschlägen, vgl. Stevanovic 2013; Stivers und Sidnell 2016) oder eingebrachten Begründungen zustimmend anzeigen. Gerade im Rahmen gemeinsamer Entscheidungsfindungen sind ratifizierende Zustimmungen durchaus erwartbar, denn es besteht, wie Stevanovic (2013: 520, Hervorh. i. O.) feststellt, eine „strong normative expectation that it should be *more than one* participant who has a word to say on the decisions to be made“. Durch das gegenseitige Zustimmen können Interagierende sowohl lokalen als auch globalen Konsens anzeigen. Während durch den lokalen Konsens beispielsweise das Aushandeln einer Option für den Moment abgeschlossen wird (vgl. Kreuz 2021), kann er auf globaler Ebene auch die kommunikative Aufgabe des Begründens beenden (vgl. Heller 2012: 91) und somit in das Abschließen der gesamten Diskurseinheit steuern. Bleiben Zustimmungen zu eingebrachten Position oder Begründung hingegen aus, kann dies den*die bisherige*n Sprecher*in zu einer Elaboration seiner*ihrer ursprünglichen Position oder Begründung veranlassen (vgl. Houtkoop 1987/2019). Sprachlich werden Zustimmungen u. a. durch Affirmations- bzw. zustimmenden Negationspartikeln¹³ umgesetzt sowie durch ausgebaute Turnkonstruktionseinheiten, in denen z. B. Teile der ersten Äußerung wiederholt oder Paraphrasen angeschlossen werden (wie *ja, das sehe ich auch so*) (vgl. Pomerantz 1984; Houtkoop 1987/2019; Heritage und Raymond 2005).

Eingebettet in größere Strukturen wie Konzessionen kann sich die affirmierende Funktion von Zustimmungen auch abschwächen. Als sogenannte „weak agreement[s]“ (Pomerantz 1984: 68) projizieren sie in der von Couper-Kuhlen und Thompson (2000) beschriebenen musterhaften sequenziellen Abfolge der sogenannten *Cardinal Concessive* nachfolgende Einwände: Die Interagierenden stimmen dabei einer Position (X) partiell zu (X'), bevor sie anschließend gegen die Position einwenden (Y) (vgl. auch Kotthoff 1993; Houtkoop 1987/2019; Couper-Kuhlen und Thompson 2000; Barth-Weingarten 2006; Rezat 2007; Heller 2012; vgl. zu konzessiven Strukturen innerhalb des *Turns* eines*einer Sprecher*in Antaki und Wetherell 1999; Pfänder 2016). In einer Variante der *Cardinal Concessive* (X, X', Y') kann die schwache Zustimmung die Konzession auch abschließen (X, Y, X'). Dadurch wird einerseits der vorangegangene Einwand abgemildert und zugleich die schwache Zustimmung durch ihre äußerungsinterne Position prominent gesetzt (vgl. Couper-Kuhlen und Thompson 2000: 393).

¹³ Wenn die vorangegangene Äußerung eine Negation enthält, können in anschließenden Äußerungen Negationsmorpheme Zustimmung ausdrücken.

Einwenden

Beim Einwenden handelt es sich um eine reaktive, auf die Vorgängeräußerung reagierende und diese problematisierende opponierende Handlung (vgl. Spranz-Fogasy 2005b; Komor 2010: 125). Anders als in Alltagsinteraktionen liegt in der Diskursaktivität des Argumentierens eine *Präferenz* zu opponierenden argumentativen Handlungen vor, wie bereits in frühen Arbeiten von Goodwin (1990) sowie Kott-hoff (1993) deutlich wird. So werden Einwände im Rahmen dieser Diskursaktivität nicht notwendigerweise durch vorlaufende Zögerungspartikeln o. Ä. abgeschwächt (vgl. zur Markierung von Dispräferenz Pomerantz 1984; Sacks 1987), sondern sie können auch unmittelbar an die Vorgängeräußerung angeschlossen werden. Sowohl für die Interagierenden als auch für Forscher*innen ist dabei die sequenzielle Position der argumentativen Handlung für die Entscheidung, ob es sich um einen reagierenden, gegen eine Vorgängeräußerung wendenden Beitrag handelt, relevant (vgl. Spranz-Fogasy 2005b: 13). So wie auch das Begründen (vgl. oben, Begründen) kann die argumentative Handlung des Einwendens unterschiedlich realisiert werden: Die Sprecher*innen können einer vorangegangenen Äußerung widersprechen, eine Gegenbehauptung einbringen (vgl. Spranz-Fogasy 2005b; Heller 2012) oder die Vorgängeräußerung begründet widerlegen (vgl. Heller 2012). Dabei können die Einwände oberflächensprachlich durch Negationen wie *nicht*, *kein* sowie prosodische Kontrastierungen (z. B. *dann werden ALLE seekrank / KEIner wird dann seekrank*; vgl. u. a. Goodwin und Goodwin 1987; Goodwin 1990) oder adversative Junktoren wie *aber* markiert und abgeschwächt oder verstärkt modalisiert werden (vgl. u. a. Grundler 2011: 350ff.; Hauser und Luginbühl 2017; Morek 2020). Innerhalb der globalen Struktur von mündlichen Argumentationen können Interagierende durch einwendende Handlungen einen Dissens herstellen oder aber im Kernjob des Begründens eine Position (vgl. oben, Position beziehen) oder Begründung problematisieren bzw. zurückweisen (vgl. Heller 2012). In interaktiver Konsequenz kann das Einwenden dazu führen, dass Vorgängeräußerungen plausibilisiert oder vorangegangene Begründungen weiter elaboriert werden.

Einwände können auch in größere Strukturen wie Konzessionen eingebettet sein (vgl. oben Zustimmung). Dabei folgt in einer dreiteiligen Struktur der Äußerung (X) eines*einer ersten Sprecher*in eine schwache Zustimmung (X') sowie der Einwand (Y) eines*einer weiteren Sprecher*in (vgl. Couper-Kuhlen und Thompson 2000). Durch die Struktur X, X', Y wird die Opposition durch ihre Position innerhalb des *Turns* besonders hervorgehoben und interaktiv relevant gemacht.

Einfordern von Rückmeldungen

In mündlichen Argumentationen kann das Einfordern von Rückmeldungen durch unterschiedliche Fragetypen (vgl. Tracy und Robles 2009), wie beispielsweise Rück- und Ergänzungsfragen, Entscheidungs- sowie Alternativfragen, umgesetzt werden, die je nach sequenziellem Kontext zur Bearbeitung unterschiedlicher kommunikativer Aufgaben (vgl. Kap. 2) beitragen.¹⁴ So können Rück- und Ergänzungsfragen, die einen Bezug zu einem vorangegangenen Beitrag herstellen und u. a. mit dem Interrogativpronomen *warum* realisiert werden, zur Elaborierung einer Begründung oder Plausibilierung einer Position führen (vgl. Heller 2012; vgl. Hausendorf und Quasthoff 1996/2005). Auf diese Weise kann sowohl eine Begründungspflicht etabliert als auch innerhalb des Kernjob des Begründens eine Vertiefung einer be-

¹⁴ Die Benennung der argumentativen Handlung als *Einfordern von Rückmeldungen* ist angelehnt an den Artikel von Stivers und Rossano (2012), in dem die Autor*innen untersuchen, wie die Verpflichtung zu antworten interaktiv angezeigt wird. Aus der Studie geht hervor, dass die graduelle Stärke einer Antwortverpflichtung von der sozialen Handlung (*action*), ihrer sequenziellen Position sowie von den (interrogativen) Merkmalen des Turndesigns (lexikalisch, morpho-syntaktisch, prosodisch sowie der Blickadressierung durch den*die Sprecher*in) abhängig ist.

reits erbrachten Begründung oder Position eingefordert werden. Je nach Turndesign kann in der Rückfrage auch eine Bewertungsdivergenz sichtbar werden. So können *denn*-Fragen (vgl. Deppermann 2009; Heller 2012) sowie prosodische Merkmale, beispielsweise in Form kontrastierender Fokusakzentuierungen sowie in Form von Tonhöhenprüngen (vgl. Heller 2012: 79, 140f.), eine deutliche Nicht-Übereinstimmung des*der aktuellen Sprecher*in mit dem vorangegangenen Beitrag anzeigen.

Insbesondere in argumentativen Entscheidungsfindungen zeigt sich, dass die Beteiligten lokale Zugzwänge einleitend im Rahmen des Entscheidungsdiskurses (vgl. Quasthoff und Kluger 2021) als auch resümierend und als gemeinsames, die Argumentation abschließendes Fazit stellen. Die lokalen Zugzwänge können dann als Präferenzabfragen (z. B. *Wofür bist du?*) oder als Entscheidungs- bzw. Ja-/Nein-Fragen (z. B. *Also entscheiden wir uns für X?*) eingebracht werden (vgl. auch Koshik 2002; Raymond 2010; Heritage und Raymond 2012). Interagierende können diese dann entweder lokal bearbeiten, indem sie ausschließlich ihre Position i. S. einer Präferenz benennen, oder aber bereits mit Begründungshandlungen verknüpfen und somit ein größeres Äußerungspaket aufbauen (vgl. u. a. in Peer-Interaktionen Hauser und Luginbühl 2017). Letztlich ermöglichen Ja-/Nein-Fragen (vgl. Koshik 2002) auch – beispielsweise nach längeren Begründungshandlungen – konklusiv die Präferenz eines*einer Sprecher*in lokal abzusichern (z. B. *Also bist du für X?*), sodass i. S. des *Common Grounds* (vgl. Clark 1996) eine gemeinsame Wissensbasis für die Fortsetzung des Kernjob des Begründens geschaffen oder durch gegenseitiges Beipflichten (vgl. Heller 2012) in das Abschließen der Argumentation hineingesteuert werden kann.

Im Anschluss an die vorgestellten argumentativen Handlungen, die im Rahmen der Studie in argumentativen Entscheidungsfindungen relevant werden, schließt das Kapitel mit einer Zusammenfassung und Überleitung zu Kapitel 3.

2.3 Zusammenfassung

In dem vorliegenden Kapitel wurde einführend (vgl. Kap. 2.1) und in Abgrenzung zu argumentationstheoretischen Modellen hergeleitet, dass sich für das Verständnis der *interaktiven* und *sequenziellen* Entfaltung mündlicher Argumentationen ein konversationsanalytischer Zugang als sinnvoll erweist. Dieser ermöglicht Forscher*innen, mündliche Argumentationen nicht als defizitäre Formen einer ‚idealen‘ Argumentation aufzufassen, sondern sie unter Berücksichtigung des Kontextes als eine *kommunikative Gattung* bzw. *diskursive Praktik* zu rekonstruieren, in denen Interagierende auf verfestigte Muster bzw. Problemlöseverfahren zurückgreifen, um einen Dissens zu lösen (vgl. Kap. 2.2). Dabei ermöglicht das analytische Beschreibungsinstrument GLOBE, sowohl die makrostrukturelle Organisation (vgl. Kap. 2.2.1) i. S. der Jobs zu erfassen und somit die Beteiligungsstruktur der Sprecher*innen sichtbar zu machen – die in Kapitel 10 unter erwerbsorientierten Fragestellungen in den Fokus der Untersuchung rücken – als auch die zur Bearbeitung der Teilaufgaben eingesetzten pragmatischen Mittel bzw. argumentativen Handlungen herauszupräparieren, die nach Hausendorf und Quasthoff (1996/2005) v. a. mit sprachlichen und parasprachlichen Ressourcen (Formen) realisiert werden (vgl. Kap. 2.2.2).

Da in mündlichen Argumentationen Interagierende nicht nur verbalsprachliche und parasprachliche Ressourcen nutzen, sondern, wie in der Einleitung (vgl. Kap. 1) skizziert, auch körperlich-sichtbare Ressourcen zur Realisierung argumentativer Handlungen eingesetzt werden, wird das nachfolgende Kapitel 3 vorstellen, wie das Verhältnis zwischen körperlich-sichtbaren und sprachlichen Ressourcen aus unterschiedlichen Forschungsrichtungen konzeptualisiert wird und welche Funktionen multimodale Ressourcen in der Interaktion bzw. für das mündliche Argumentieren übernehmen.

3 Multimodale Ressourcen in der Interaktion

Im Hinblick auf die erste Forschungsfrage der vorliegenden Arbeit, *welche Funktionen multimodale Ressourcen für das mündliche Argumentieren übernehmen* (vgl. Kap. 1; vgl. ausführlich Kap. 5), sollen zuerst vier ausgewählte Konzeptualisierungen des Zusammenspiels multimodaler Ressourcen und ihrer Funktion für die Interaktion vorgestellt werden (Kap. 3.1). Anschließend (Kap. 3.2) werden sie im Hinblick auf ihre pragmatische Funktion – mit besonderem Fokus auf Gesten in multimodalen Ressourcenverbünden – und unter Berücksichtigung der kommunikativen Gattung des Argumentierens eingeordnet (Kap. 3.3). Das Kapitel schließt mit einer Zusammenfassung (vgl. Kap. 3.4) der Befunde und ihrem Ertrag für die o. g. übergeordnete Forschungsfrage. Analog zu Kapitel 2 wird in dem vorliegenden Kapitel 3 also die forschungsmethodische sowie terminologische Grundlage für die Untersuchung multimodaler Ressourcen in mündlichen Argumentationen geschaffen.

3.1 Eine konzeptionelle Einordnung: Untersuchung multimodaler Ressourcen in der Interaktion

Das wissenschaftliche Interesse unterschiedlicher Forschungsbereiche an der Untersuchung multimodaler Ressourcen im Gespräch erfährt im Laufe des 20. Jahrhunderts einen deutlichen Anstieg. Es entstehen *kulturvergleichende* (vgl. Efron 1941/1972) und *psychologische* Studien (vgl. u. a. Condon und Ogston 1966, 1971; Ekman und Friesen 1972; McNeill 1985, 1986, 1992; Bavelas et al. 1992, 1995) sowie Studien der sich in den 1970er Jahren gerade entwickelnden *Gestenforschung* (vgl. Kendon 1970, 1972b, 1990a, 1990b). Und auch an der *ethnomethodologischen Konversationsanalyse* orientierte Arbeiten (vgl. u. a. Goodwin 1979, 1981; Goodwin 1980) verweisen bereits auf die Relevanz körperlicher Ressourcen in der Interaktion. Eine zentrale Frage der Arbeiten ist, wie sich das Verhältnis v. a. zwischen sprachlichen und körperlich-sichtbaren Ressourcen gestaltet. Die Frage wird jeweils unterschiedlich von den Forschenden beantwortet. Im Folgenden werden daher vier ausgewählte Konzeptualisierungen vorgestellt, deren Modellierungen neben dem Verhältnis der sprachlichen und körperlich-sichtbaren Ressourcen auch räumliche sowie materielle Bedingungen einschließen und dabei untersuchen, wie diese Ressourcen zur Bedeutungskonstitution in der Interaktion sowie zur Markierung eines kommunikativen Zwecks beitragen bzw. eine kommunikative Handlung aufbauen. Die Auswahl der vier Konzeptualisierungen orientiert sich an ihrem grundlegenden Verständnis von körperlich-sichtbaren Ressourcen als *bedeutungstragende* bzw. *bedeutungskonstituierende* Ressourcen, dem die vorliegende Studie folgt. Zugleich kontrastieren sie u. a. in ihrer Auffassung, welche Ressourcen sie jeweils in ihrer Konzeptualisierung berücksichtigen; auf welche Beteiligte sie fokussieren (z. B. nur Sprecher*innen oder alle Beteiligte) und ob beispielsweise auch materielle und räumliche Bedingungen, die Konsequenzen für die stattfindende Interaktion haben können (z. B. wenn Interagierende in ihren Beiträgen konkrete Bezüge zu Gegenständen der Umgebung herstellen), in die Konzeptualisierungen eingehen. Die Vorstellung der teils divergenten und teils einander ergänzenden Konzeptionalisierungen ermöglicht schließlich begründet abzuleiten, dass bzw. warum die vorliegende Arbeit v. a. Mondadas (2014b) Verständnis von multimodalen Gestalten folgt (vgl. Kap. 3.1.4).

3.1.1 Simultaneous Constructions (Adam Kendon)

Adam Kendon, der die Gestenforschung in den 1970er Jahren begründet, legt in seinen Studien vornehmlich einen Fokus auf das Verhältnis von Sprache und der körperlich-sichtbaren Ressource der Geste. Dabei begreift er die Ressourcen von Sprache und Geste in ihrer Gleichzeitigkeit (*simultaneity*) als gleichberechtigte Elemente einer Äußerung, in der „[g]esture is *partner* with speech in the utterance as finally constructed“ (Kendon 2004: 111, Hervorh. i. O.).¹⁵ Im Aufbau (*construction*) der Äußerung werden dabei Gesten und Sprache als ein ‚gesture-speech ensemble‘ stetig *aufeinander* sowie auf den *kommunikativen Zweck* (vgl. Kendon 2013: 21f.) abgestimmt. Kendon (2004) nimmt dabei anders als kognitionspsychologische Ansätze (vgl. McNeill 1992: 295) an, dass Gesten und Sprache zwei *unterschiedliche* Zeichensysteme bilden, die aber in ihrem *Ensemble* eng miteinander verbunden sind (vgl. auch Streeck 2009b: 22). Darin bringen beide Ressourcen *gemeinsam* eine komplexe Bedeutung hervor, in der sich die Ressourcen reziprok aufeinander beziehen und in ihrer Bedeutung verstärken oder ergänzen können:

We shall see that speakers create ensembles of gesture and speech, by means of which a semantic coherence between the two modalities is attained. This is not to say that speech and gesture express the same meanings. They are often different. Nevertheless, the meanings expressed by these two components interact in the utterance and, through a reciprocal process, a more complex unit of meaning is the result. (Kendon 2004: 108f.)

Insgesamt tragen die Ressourcen gemeinsam zur *kommunikativen Handlung* bei, in denen Gesten als visuell sichtbare Formen der Handlung – gemeinsam mit verbalen Ressourcen oder autark – „show that one is asking a question, pleading for an answer, is in disagreement, and a host of other things, specific to the current linguistically managed interchange“ (Kendon 2013: 8; vgl. auch Kap. 3.2).

Auch wenn in Kendons späteren Arbeiten ein starker Fokus auf die Ressourcen der Gesten und Sprache hervortritt (vgl. u. a. Kendon 1980, 1997, 2000, 2004), der sich u. a. in Begriffen wie dem *gesture-speech ensemble* niederschlägt, sind auch weitere Ressourcen wie der Blick (vgl. Kendon 1967), Bewegungen des Kopfes wie das Kopfschütteln (vgl. Kendon 2002, 2004) oder auch das (körperliche) Eintreten in die Interaktion (vgl. Kendon 1970, 1990b)¹⁶ Gegenstände seiner Untersuchungen. Im Gegensatz zum Kopfschütteln bilden allerdings die Blickadressierung sowie die körperlichen Orientierungen der Interagierenden in Kendons Verständnis keine Gesten im Sinne der ‚visible bodily actions‘, denn „they are done in a way that subordinates them to actions that must be done merely to maintain whatever spatial and orientational arrangement the participants of a given conversation may be using“ (Kendon 2004: 9). D. h., dass diese multimodalen Ressourcen nicht an der kommunikativen Handlung selbst beteiligt sind, sondern sie vielmehr als Wegbereiter aufgefasst werden, um eine kommunikative Handlung ausführen zu können (vgl. auch Kendon 2000: 49).

¹⁵ In Bezug auf die Grammatikalisierung von ‘simultaneous constructions’ von Sprache und Gesten verweist Kendon (2013: 21) darauf, dass die Konzeptualisierung von Sprache und „utterance visible action has not, in any community of speakers, become stabilized and formalized as shared practice to the point that it must be considered as a part of the formal grammar of any spoken language“. Die vorliegende Arbeit nimmt allerdings nicht im Sinne der Grammatikalisierung Bezug auf den Begriff der ‚simultaneous constructions‘; vielmehr soll dieser zu dem Verständnis beitragen, wie Kendon die Beziehung zwischen sprachlichen und körperlichen Ressourcen modelliert und dadurch zu einem innovativen Begriffsverständnis sowie zu einer ebensolchen analytischen Haltung beigetragen hat (vgl. auch Pekarek Doehler et al. 2022).

¹⁶ Goffman (1963) würde dies aus soziologischer Perspektive das *Eintreten in eine fokussierte Interaktion* nennen. Kendon (1990b: 220) berücksichtigt diesen Ansatz in seiner Untersuchung.

Indem Kendon insgesamt eine Beschreibung der v. a. gestischen und sprachlichen Ressourcen als gemeinsam hervorgebrachte Äußerung forciert, wirkt er einer Hierarchisierung der Ressourcen entgegen: Gesten sind in diesem Verständnis keine additiv hinzutretenden, sondern gleichberechtigte Ressourcen, die gemeinsam mit der Sprache eine Gesamtäußerung hervorbringen. Kendons Verständnis der nicht-hierarchischen Beziehung zwischen gestischen und sprachlichen Ressourcen öffnet den Blick dafür, Gesten nicht ausschließlich als eine Form der *Repräsentation* der sprachlichen Äußerung, sondern als eine *Komponente* der gemeinsamen Äußerung in ihrer *Funktion* sowie ihrem *semantischen Gehalt* zu untersuchen (vgl. Kendon 2000, 2004, 2017).¹⁷ Dass es Funktionsunterschiede gibt und auch geben muss, führt Kendon (2000: 61) auf die jeweils zugrundeliegende Medialität zurück, denn während sprachliche Ressourcen an die Linearität und Flüchtigkeit der Sprachproduktion gebunden sind, können Gesten diese durchbrechen und so beispielsweise Aspekte über die sprachliche Äußerung der Sprecher*innen hinaus präsent halten.

Kendons Arbeiten sind auch wegweisend in ihrem analytischen Zugriff, denn durch seine Art der Notation macht er die Entfaltung von Sprache *und* Geste sichtbar (vgl. Kendon 2004; vgl. auch Kendon 1980). Er nimmt dabei an, dass gestische Ressourcen, ebenso wie sprachliche Ressourcen, in einzelne Abschnitte unterteilt werden können (vgl. Kendon 2004: 110). Diese nennt er Gestenphrasen (*Gesture Phrases*) und Gesteneinheiten (*Gesture Units*). Eine Gesteneinheit dauert vom Beginn der Vorbereitung (*Preparation*) bis zur endgültigen Rückkehr der Geste (*Recovery*) – z. B. in eine Ruheposition (vgl. Abb. 1). Die Gesteneinheit kann dabei aus mehreren, aufeinander folgenden Gestenphrasen bestehen, deren jeweiliger Kern aus *einem* Gestenhöhepunkt (*Stroke*) besteht und die (vorerst) nicht in eine Ruheposition zurückgeführt werden. Der Gestenhöhepunkt bildet nun das, was allgemein als ‚Geste‘ bezeichnet wird (vgl. Kendon 2004: 111f.):

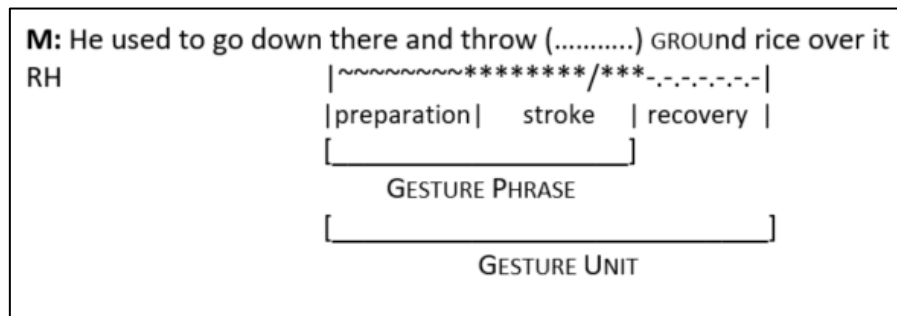


Abbildung 1: Gestenphasierung nach Kendon (2004: 114)

Die dezidierte Beschreibung der Entfaltung der Gesten in ihren Einheiten sowie die Notation innerhalb des Verbaltranskripts ermöglicht es, Geste(n) und sprachliche Ressourcen gemeinsam zu untersuchen. Trotz der Berücksichtigung der verbalsprachlichen Äußerung zeigen Kendons Untersuchungen eine starke Fokussierung auf die gestische Ressource – eine tiefergehende Beschreibung und Analyse sprachlicher (z. B. lexikalischer oder syntaktischer) oder prosodischer Phänomene bleibt hier aus. Ferner liegt der Schwerpunkt von Kendons Arbeiten auf Interaktionsbeteiligten, die als *Sprecher*innen* in Erscheinung treten. Inwiefern auch *Rezipient*innen* am Aufbau (multimodaler) kommunikativer Handlungen beteiligt sein können, wie konversationsanalytische Arbeiten, u. a. von Mondada (2006, 2007, 2015b;

¹⁷ Kendon (2004) verweist darauf, dass in der Reziprozität der Gesten sowie der Sprache diese immer auch eine semantische Kohärenz aufweisen. Daher spricht er auch von einem „semantically coherent gesture-speech ensemble“ (Kendon 2004: 127, Hervoh. i. O.).

vgl. Kap. 3.1.4), zeigen, geht aus Kendons Studien nicht hervor. Somit bleibt die Rolle der Rezipient*innen als *aktiv* an der Interaktion Beteiligte in seinen Untersuchungen unbeleuchtet.

3.1.2 Composite Utterances (Nick J. Enfield)

Nick J. Enfield modelliert aus einem anthropologischen linguistischen Verständnis heraus das Verhältnis zwischen Körperlichkeit, Räumlichkeit sowie Zeitlichkeit und Sprache als eine Beziehung diverser *Zeichen* (vgl. Enfield 2009), die kognitive Prozesse, i. S. von Gedanken, und kommunikative Absichten der Sprecher*innen sichtbar machen (vgl. Enfield 2009: 2, 2013: 690). Diese können im Verbund ein zusammengesetztes Zeichen, ein *composite sign*, bilden, das Informationen trägt. Anders als Kendon (2004) lehnt sich Enfield (2009, 2013) in der Beschreibung der Multimodalität von Äußerungen an ein *semiotisches* Verständnis an und schließt darin auch jegliche Form der Multidimensionalität sowie Multimedialität ein.¹⁸ In Abgrenzung zu einem starren Saussureschen Zeichenverständnis, in dem angenommen wird, dass das Wort *X* die Bedeutung *Y* trägt, legt Enfield (2013: 690) ein dynamisches Verständnis von Zeichen und ihrer Bedeutungskonstitution zugrunde. Dieses berücksichtigt sowohl die Sprecher*innen, die Rezipient*innen als Interpret*innen der Äußerung, wie Enfield (2009: 1) sie nennt, als auch den Äußerungskontext: „people who utter the word *X* are normatively taken by others to intend *Y* across a range of contexts“. Die eingesetzten einzelnen Zeichen können dann wiederum in ihrem Grad der Konventionalisierung variieren, sodass eine Äußerung sowohl aus „part conventional, part non-conventional [signs]“ bestehen kann (Enfield 2009: 6). In dem Zusammenschluss der Zeichen zu einer Äußerung, einer *composite utterance*, trägt diese Äußerung dann eine Bedeutung: „we see speaker’s overall utterance meaning as a **unified** product of multiple resources of information“ (Enfield 2009: 6, Hervorh. N. S.). Die *composite utterance* beinhaltet Informationen über die kommunikative Absicht der Sprecher*innen, deren Bedeutung durch die Rezipient*innen bzw. Interpret*innen im Rahmen der kommunikativen Handlung¹⁹ sowie des jeweiligen Kontexts als Ganzes dechiffriert wird:

The gesture, gaze, and speech components of the utterance are taken together as a unified whole. As interpreters, we effortlessly integrate them as relating to one overall idea. [...] [S]igns co-occurring with other signs, acquiring unified meaning through being interpreted as co-relevant parts of a single whole. (Enfield 2009: 7)

Enfields (2009) Konzeption der *composite utterance* als ein Zusammenschluss von Zeichen, die weder in ihrer Modalität noch ihrer Medialität bereits einschränkend definiert werden, ermöglicht es, jegliche Zeichen, die gemeinsam Bedeutung tragen und durch die Rezipient*innen dechiffriert bzw. enkodiert werden, in der sozialen Interaktion zu untersuchen (vgl. Enfield 2009: 7, 2013). Mit diesem weiten forschungsmethodischen Zugriff können u. a. räumliche Gegebenheiten, die u. U. von den Akteur*innen in ihrem kommunikativen Handeln relevant gemacht werden, analytisch mit einbezogen werden.²⁰ Ferner wird in dem grundlegenden Verständnis unterschiedlich beschaffener Zeichen sichtbar, dass, in Abgrenzung zu Kendon (2004), die sprachliche Ressource eingehender in Untersuchungen der *composite utterance* einbezogen werden kann, indem u. a. syntaktische, lexikalische und auch prosodische Phäno-

¹⁸ Enfield (2009: 3ff.) untersucht u. a. ein Foto von Willy Brandts Kniefall im Hinblick auf die Zeichen, die gemeinsam eine Bedeutung konstituieren. Um die Bedeutung des Fotos verstehen zu können, muss der Kontext (Akteur*innen, Ort, Zeit etc.) berücksichtigt werden.

¹⁹ Enfield (2009: 9) spricht hier, in Abgrenzung zu linguistischen Einheiten wie der Intonationseinheit oder wie dem *Turn*, von einem sogenannten „communicative move“. Dies ermöglicht, auch solche Handlungen forschungsmethodisch mit zu berücksichtigen, die nicht sprachlich verbalisiert werden, sondern ausschließlich körperlich stattfinden (vgl. zu einer kritischen Einordnung des Begriffs ‚move‘ Streeck 2010: 461f.).

²⁰ Vgl. hierzu Streeck (2010: 463), der kritisch feststellt, dass Enfield (2009) selbst v. a. Interaktionen untersucht, in denen Sprecher*innen keine weiteren Materialien der Umgebung in die Interaktion einbinden.

mene berücksichtigt werden (vgl. Enfield 2009: 108f.). Als logische Konsequenz der Konzeptualisierung Enfields, die das Verhältnis von Sprache und anderen Zeichen als *einen* Zeichenverbund annimmt, der *eine* Bedeutung trägt, zeigt sich nun allerdings, dass die Notation keine Informationen über die zeitliche Entfaltung der multimodalen Ressourcen, die mit Gegebenheiten des Raumes etc. koordiniert werden, enthält. Die Transkripte bilden, anders als beispielsweise bei Kendons (2004; vgl. Kap. 3.1.1) Notationssystem, ausschließlich das Verbaltranskript ab, das dann um Beschreibungen der Ressourcen in der Analyse selbst sowie Standbilder ergänzt wird. Dieses Vorgehen birgt zwei Schwierigkeiten: Erstens, indem die Transkripte ausschließlich Informationen über die sprachlichen Zeichen beinhalten und weitere Ressourcen über Standbilder ergänzt werden, kann nicht intersubjektiv nachvollzogen werden, *wie* sich die Ressourcen entfalten. Daraus resultiert, zweitens, dass verborgen bleibt, ob die zeitliche Entfaltung der Ressourcen möglicherweise Auswirkungen auf die jeweilige Bedeutungskonstitution hat.

3.1.3 Contextual Configuration (Charles Goodwin)

Auch Charles Goodwin untersucht durch einen linguistischen anthropologischen sowie ethnomethodologisch geprägten Forschungszugang, wie Menschen in der Interaktion ihr (gemeinsames) Handeln organisieren und dabei unterschiedliche Ressourcen, wie Sprache, Körper, Objekte der Umgebung sowie auch räumliche Gegebenheiten, relevant setzen. Ähnlich wie Enfield (2009; vgl. auch Kap. 3.1.2), konzeptualisiert Goodwin (2000, 2018) die sich entfaltenden Handlungen als jeweils durch semiotische Ressourcen bzw. *Zeichen* konfiguriert. In dem Prozess der Entfaltung der Handlung verändern sich dabei kontinuierlich sowohl die Handlung als auch die beteiligten Ressourcen, da zu unterschiedlichen Zeitpunkten der Handlung andere Ressourcen von den Interagierenden relevant gesetzt werden können (vgl. Goodwin 2000, 2018: Kap. 12). Mit seinem Ansatz wendet Goodwin (2000: 1490) sich dezidiert gegen eine *isolierte* Betrachtung der *sprachlichen* Ressource. Vielmehr konzeptualisiert er Interaktionen als bestehend aus übergeordneten Aktivitäten (*activities*), die sich aus Handlungen (*actions*) der Sprecher*innen zusammensetzen, an denen wiederum diverse semiotische Ressourcen (*semiotic resources*) beteiligt sind.²¹ Abhängig von ihrer Medialität (z. B. Sprache, Körper, Objekt der Umgebung), sind die semiotischen Ressourcen in sogenannte *semiotische Felder* (*semiotic fields*) eingebettet (vgl. Goodwin 2000: 1494, 2018: 169f.).

Die lokale Anordnung der semiotischen Felder, die aus unterschiedlichen Zeichen bestehen (vgl. Goodwin 2000: 1490, 2011: 183) und die die sich entfaltende Handlung aufbauen, bezeichnet Goodwin (2000: 1490) als *contextual configuration*. Durch die *Orientierung* der Interagierenden wird erkennbar, *welche* semiotischen Felder für die gemeinsame Handlung relevant sind (vgl. Goodwin 2000: 1490). So beobachtet Goodwin in Studien, die er an archäologischen Ausgrabungsstätten durchführt, wie die Interagierenden z. B. Farbtafeln verwenden, um die Farbgebung der Erdschichten zu dokumentieren (vgl. Goodwin 2000: 1509). Die Farbtafel wird somit zu einer wichtigen Ressource in der gemeinsamen Interaktion – solange, wie Interagierende sie für ihr Handeln relevant setzen. In der Interaktion können dann (weitere) semiotische Felder neu hinzutreten oder wegfallen, „with the effect that the contextual configurations which frame, make visible, and constitute the actions of the moment undergo a continuous process of change“ (Goodwin 2000: 1490). Wenden sich die Archäolog*innen beispielsweise in ihrem Arbeitsprozess nach der Kategorisierung einer ersten Fläche einer anderen, noch zu bestimmenden Fläche zu, tritt der bereits kategorisierten Bereich in den Hintergrund, während ein anderer als semiotische Ressourcen für das aktuelle Handeln relevant gesetzt wird. Innerhalb dieser kontinuierlichen

²¹ Goodwin (2000: 1490) kritisiert, dass in der isolierten Betrachtung von Sprache körperliche und räumliche Ressourcen lediglich unter dem Begriff des Kontextes subsumiert werden und durch diesen Zugriff eine Dichotomie zwischen Sprache und Kontext forciert wird.

Veränderung besteht aber eine „continuing contextual gestalt“ (Goodwin 2018: 184) fort. In dem genannten Beispiel kann darunter die Tätigkeit im archäologischen Feld verstanden werden. Sie umspannt gewissermaßen rahmenbildend die Handlungen in der Interaktion. Insgesamt ist die kontextuelle Konfiguration der Handlung also geprägt durch die von den Interagierenden eingesetzten multiplen semiotischen Ressourcen, die, verbunden durch ihre sequenzielle, zeitlich dynamische Entfaltung, *gemeinsam* semiotische Felder bilden und sich dabei gegenseitig elaborieren (vgl. Goodwin 2000: 1490ff.).

Die forschungsmethodische Implikation, die mit dem Verständnis von Handlungen als *contextual configuration* einhergeht, ist die Folgende:

When action is investigated in terms of contextual configurations, domains of phenomena that are usually treated as so distinct, that they are the subject matter of entirely separate academic disciplines, e.g., language and material structure in the environment, can be analyzed as integrated components of a common process for the social production of meaning and action. [...]

A theory of action must come to terms with both the details of language use and the way in which the social, cultural, material and sequential structure of the environment where action occurs, figure into its organization. (Goodwin 2000: 1490f.)

Goodwins Ansatz ermöglicht somit, auch Materialien der Umgebung (wie beispielsweise der Boden archäologischer Ausgrabungen und die zur Bestimmung eingesetzten Farbtafeln (vgl. Goodwin 2000, 2018) in der Analyse der menschlichen Handlung sowie in ihrer Bedeutung für die Interaktion und Gesprächsorganisation zu berücksichtigen (vgl. Goodwin 2000: 1491). Er nutzt den Begriff der semiotischen Felder, um die einzelnen Medialitäten und Materialitäten analytisch voneinander zu trennen. So bildet die Sprache mit ihren Zeichen ein semiotisches Feld oder auch Subsystem – und Gesten jeweils ein anderes. Die analytische Trennung bedeutet aber keine tatsächliche Trennung der Ressourcen: Die unterschiedlichen Ressourcen, die gleichzeitig geäußert werden und sich gegenseitig elaborieren, „create a whole that is different from, and greater than, any of its constituent parts“ (Goodwin 2007: 55). Dieser analytische Rahmen ermöglicht eine Beschreibung der *einzelnen* Ressourcen, die aber in ihrer schrittweisen sequenziellen, zeitlichen Entfaltung *gemeinsam* eine *contextual configuration* bilden. Um diese Entfaltung nachvollziehen zu können, bilden Goodwins Analysen, im Gegensatz zu beispielsweise Enfield (2009), teilweise in Form ikonischer Abbildungen in Alignierung mit dem Verbaltranskript (vgl. Goodwin 2018: 172) oder in Form verbaler Beschreibungen aligniert mit dem Verbaltranskript (vgl. Goodwin 2018: 162ff.) ab. Welche der einzelnen Ressourcen nun in der *contextual configuration* relevant werden, kann nur anhand der Handlungen der Beteiligten rekonstruiert werden. Es werden die Ressourcen für die *contextual configuration* relevant, die die *Interagierenden* reflexiv in ihre Handlung einbeziehen (vgl. Goodwin 2000: 1504, 1509). Im Unterschied zu Kendon und Enfield (vgl. Kap. 3.1.1 & 3.1.2) trägt Goodwin (2000, 2018) in seinen Arbeiten dem Umstand Rechnung, dass auch die *Adressat*innen* wichtige Beteiligte der Interaktion darstellen, deren Handeln dynamische Veränderungen im aktuellen Geschehen bewirken (können). Beispielsweise kann sich ein*e Rezipient*in einer sprecher*innenseitigen Adressierung entziehen, indem er*sie den Blick von dem*der Sprecher*in abgewandt hält. Wird dies von dem*der Sprecher*in wahrgenommen, können darauf interaktive Konsequenzen folgen, um die Aufmerksamkeit des Gegenübers zu erlangen (vgl. Goodwin 2000: 1501). Um die dynamische Entwicklung einer Interaktion also analytisch erfassen zu können, nehmen Goodwins Studien daher auch das Handeln der Adressat*innen in den Blick.

3.1.4 Complex Multimodal Gestalts (Lorenza Mondada)

Mondada (2014a, 2014b, 2018) untersucht menschliche Interaktion aus ethnomethodologisch konversationsanalytischer sowie interaktionslinguistischer Perspektive. Ihre Arbeiten tragen maßgeblich zur Entwicklung der *multimodalen* Konversationsanalyse bei.²² Ähnlich wie Enfield (2009; vgl. Kap. 3.1.2) und Goodwin (2000, 2018; vgl. Kap. 3.1.3) untersucht auch Mondada, wie Sprecher*innen sprachliche und körperlich-sichtbare Ressourcen unter Berücksichtigung ökologischer Aspekte, wie räumlicher Gegebenheiten (z. B. Arbeitsbesprechungen an Tischen, in denen auch Dokumente einbezogen werden können (vgl. u. a. Mondada 2014b)), koordinieren und ihr Handeln reflexiv aufeinander abstimmen. Allerdings weist ihre Konzeptualisierung keine Orientierung an semiotischen Ansätzen auf (vgl. hierzu Kap. 3.1.2, 3.1.3), in denen auch Objekte der Umwelt *semiotische* Ressourcen darstellen; vielmehr versteht Mondada sprachliche und körperlich-sichtbare Ressourcen als *multimodale* Ressourcen, die Interagierende in ihrem und für ihr Handeln einsetzen – jeweils angepasst an räumliche und materielle Gegebenheiten. Die beteiligten Ressourcen, sowohl sprachliche als auch körperlich-sichtbare, werden dabei nicht *a priori* durch die Forscher*innen in ein hierarchisches Verhältnis gesetzt; sondern Mondada geht wie Goodwin (2000) davon aus, dass durch die Orientierung der Interagierenden selbst sichtbar wird, welche Ressourcen sie für ihr Handeln relevant setzen. Anhand ihrer Studien von Alltags- sowie beruflichen Interaktionen, u. a. während des Einkaufens, in gemeinsamen Begehungen von Baustellen, in Arbeitsmeetings, während einer Gartenführung oder beim Aufbau von Kunstwerken in Museen, arbeitet Mondada (2014a, 2014b, 2018) heraus, dass unterschiedliche Ressourcen *gemeinsam* zur Organisation einer Handlung beitragen. Während Goodwin (2000, 2018) grundlegend an der Frage interessiert ist, wie Interagierende ihr Handeln durch unterschiedliche Ressourcen organisieren, legt Mondada (2014a, 2014b) ihren Fokus auf *wiederkehrende* Ressourcenverbünde, die gemeinsam *complex multimodal Gestalts* bilden. Dabei lehnt sie den Begriff der Gestalt u. a. an Auer (2007, 2009) an, der diesen für den Bereich der Syntax der gesprochenen Sprache verwendet. Der zugrunde liegende syntaktische Gestaltbegriff referenziert letztlich auf etablierte Musterhaftigkeiten, die es ermöglichen, dass potenzielle Fortsetzungen der gerade hervorgebrachten Äußerungen sowie auch Beendigungen von *Turns* i. S. von Gestaltschließungen antizipiert werden können (vgl. Auer 2007: 97).²³ Die am Aufbau der syntaktischen Gestalt beteiligten Strukturelemente weisen dabei, u. a. aufgrund ihrer zeitlichen Abfolgen sowie ihres jeweiligen Projektionspotenzials, eine besondere Verbindung untereinander auf. Mondada (2014b) greift sowohl a.) die *Verbundenheit* der einzelnen Strukturelemente als auch b.) die *Projektionskraft*, die komplexe multimodale Gestalten tragen, auf:

a.) In der Rekonstruktion des Einsatzes multimodaler Ressourcen in der sozialen Interaktion zeigt sich, dass sprachliche und körperlich-sichtbare Ressourcen enge Verbindungen miteinander eingehen, sodass sie als zusammengehörig identifizierbar sind und gemeinsam komplexe multimodale Gestalten bilden: Dabei erweisen sich die eingesetzten multimodalen Ressourcen als systematisch (und zeitlich) aufeinander abgestimmt, geordnet und sequenziell eingebettet sowie an ihren jeweiligen ökologischen Kontext angepasst (vgl. Mondada 2014a, 2014b, 2018). Ihre jeweilige Konfiguration, also die Ausprägung der Gestalt, ist dabei abhängig von der übergeordneten Aktivität sowie den materiellen und räumlichen Bedingungen (vgl. Mondada 2018: 86).

²² Mondada (2018: 85f.) weist selbst darauf hin, dass auch frühe konversationsanalytische Arbeiten bereits mit videographierten Daten gearbeitet und ebenfalls multimodale Ressourcen berücksichtigt haben. Allerdings zeigt die Konversationsanalyse lange Zeit einen starken Fokus auf sprachliche Ressourcen.

²³ Auer (2007: 97) entlehnt den Begriff aus der Gestaltpsychologie.

b.) Ebenso wie Goodwin ermöglichen Mondadas Arbeiten aufgrund des konversationsanalytischen Zugangs die Beschreibung der sequenziellen Entfaltung der Handlungen der Sprecher*innen in der Interaktion. Mondada rekonstruiert dabei, dass bzw. wie die multimodalen Ressourcen nicht nur innerhalb der Handlung *eines* Beteiligten *intrapersonell* eng miteinander koordiniert und aufeinander abgestimmt sind. Sie zeichnet auch nach, wie in der interaktionalen Ordnung *reflexive, interpersonelle* multimodale Orientierungen der Beteiligten *aufeinander* sichtbar werden, sodass Sprecher*innenwechsel durch multimodale Gestaltschließungen (vgl. Mondada 2015a) *projizierend* angezeigt werden können oder beispielsweise die Veränderung der Körperpositur einen anstehenden Ortswechsel innerhalb von Ortsbegehungen und somit eine Veränderung der aktuellen Handlung einleiten kann (vgl. u. a. Mondada 2014b: 147).²⁴

Insgesamt zeigt sich in Mondadas Forschungsarbeiten, dass erst mit der konsequenten Erweiterung der Konversationsanalyse, die lange Zeit vornehmlich auf sprachliche Ressourcen fokussiert war, zu einer *multimodalen* Konversationsanalyse sichtbar wird, wie kopräesente Interagierende ihre soziale Interaktion sequenziell organisieren und die dabei fein aufeinander abgestimmten Ressourcen gemeinsam komplexe multimodale Gestalten bilden (Mondada 2014a, 2014b, 2015a). Ihre Konzeptualisierung komplexer multimodaler Gestalten ermöglicht es nun in Erweiterung von Goodwin, auch *wiederkehrende* Ressourcenverbünde zu erfassen, mit denen Interagierende interaktive Handlungen realisieren (vgl. auch Kap. 1). Dieser Befund prägt nachfolgende konversationsanalytischer Arbeiten und trägt zu einem Verständnis bei, dass bzw. wie sich eine Ordnung in der sozialen Interaktion auch in der Verwendung *multimodaler* Ressourcen manifestiert.

Fasst man die Unterschiede zwischen den vorgestellten Konzeptualisierungen zusammen, so zeigt sich, dass sie sich hinsichtlich der beteiligten *interpersonellen Ressourcen*, der *außerpersonellen Ressourcen*, der in den Blick genommenen *Beteiligten* sowie in ihrem *analytischen Vorgehen* unterscheiden: Während Kendon v. a. sprachliche und gestische Ressourcen in den Fokus seiner Untersuchung stellt, beziehen Enfield, Goodwin sowie Mondada eine Vielzahl körperlich-sichtbarer Ressourcen mit ein. Indem sie außerdem systematisch räumliche und materielle Bedingungen (*außerpersonelle Ressourcen*) mit berücksichtigen, ermöglichen sie ein umfassenderes Verständnis der Handlungen von Interagierenden. Ihre Untersuchungen legen zugleich fest, welche außerpersonellen Ressourcen analytisch eingehen: Es handelt sich ausschließlich um Ressourcen bzw. Zeichen, die *für die* und *in der* Interaktion von den Beteiligten selbst relevant gemacht werden (vgl. v. a. Goodwin, Kap. 3.1.3; Mondada, Kap. 3.1.4). Indem außerdem Goodwin und Mondada ihren Blick von einer forschungsseitigen Sprecher*innenzentrierung lösen, werden die Handlungen *aller Beteiligten* in ihrer Funktion für die sich entfaltende Interaktion berücksichtigt. Goodwin und Mondada legen durch detaillierte Transkriptionen die komplexe, emergente Entfaltung von Interaktionen dar und zeigen, wie einsetzende Gesten, sich verändernde Körperposituren etc. interaktive Konsequenzen mit sich bringen können. Mit ihren Arbeiten weisen sie nach, welche interaktiven Auswirkungen multimodal realisierte Handlungen eines*einer Sprecher*in haben können, da sie auch die Reaktion weiterer Beteiligter in ihre Untersuchungen einbeziehen. Dieses *analytische Vorgehen* ermöglicht es, i. S. einer multimodalen Konversationsanalyse, die zeitliche, simultane

²⁴ Mondada (2021a) zeigt darüber hinaus in einer Untersuchung zur sequenziellen Organisation von Paarsequenzen (Bitte – Reaktion), dass beispielsweise in Geschwindigkeitstrainings eine Bitte bzw. Aufforderung durch den Kontext so hochgradig erwartbar wird, dass der Rezipient eine Handlung durchführt, noch bevor die Bitte geäußert wird. Diese Befunde manifestieren zum einen die Relevanz der analytischen Berücksichtigung multimodaler Ressourcen, zum anderen wird deutlich, dass „the detailed examination of earliness can cast new light on fundamental issues such as temporality and sequentiality as deeply related to constitutive features of multimodality and ecology“ (Mondada 2021: 415).

Entfaltung der multimodalen Ressourcen in der Interaktion zu untersuchen und dabei zugleich die sequenzielle, linear fortgesetzte Organisation der Interaktion mit zu berücksichtigen. Darüber hinaus vermag Mondadas Konzeptualisierung komplexer multimodaler Gestalten außerdem aufzudecken, welche Ressourcenverbünde rekurrente Muster bilden und i. d. S. wiederkehrende interaktive Funktionen übernehmen.

Hinsichtlich der Forschungsfrage der vorliegenden Studie, *welche Funktionen multimodale Ressourcen für das mündliche Argumentieren übernehmen* (vgl. Kap. 1), bietet also der vorgestellte analytische Zugang die Möglichkeit, *wiederkehrende Ressourcenverbünde*, die zur Realisierung argumentativer Handlungen (vgl. Kap. 2.2) eingesetzt werden, als bedeutungstragende multimodale Gestalten zu rekonstruieren. Der Einbezug rezipient*innenseitiger Handlungen, der sowohl von Mondada als auch von Goodwin forciert wird, ermöglicht ferner, sie als *aktive Beteiligte* in der Interaktion zu begreifen. D. h., dass auch in Bezug auf sie in mündlichen Argumentationen möglicherweise wiederkehrende Handlungen rekonstruiert werden können. Insgesamt kann durch die Berücksichtigung aller interaktiv Beteiligten sowie der räumlichen Bedingungen etc., die *emergente Entfaltung* argumentativer Handlungen in ihrer Komplexität erfasst werden. Die vorliegende Arbeit wird daher die Konzeptualisierung der multimodalen Gestalten von Mondada zugrunde legen. Darüber hinaus wird in der Datenaufbereitung an gegebener Stelle auf die Notation der Gestenentfaltung, die Kendon (2004) entwickelt hat, zurückkommen.

Ausgehend von der Vorstellung unterschiedlicher Konzeptualisierungen des Verhältnisses sprachlicher und gestischer bzw. körperlich-sichtbarer Ressourcen sowie der anschließenden Schwerpunktsetzung auf Mondadas Konzeptualisierung der komplexen multimodalen Gestalten, das für die vorliegende Studie entscheidend sein wird (vgl. Kap. 3.1), wird das nachfolgende Kapitel (vgl. Kap. 3.2) auf die wichtigsten Befunde zu *Gesten* als visuelle Elemente multimodaler Gestalten fokussieren.

3.2 Gesten als visuelle Elemente multimodaler Gestalten

Da für die interaktive Bearbeitung der argumentativen Entscheidungsfindung insbesondere Gesten mit *pragmatischer* Funktion relevante Funktionen übernehmen, werden die in ihrer multimodalen Gestalt (vgl. Mondada 2014b), ihrer sequenziellen Einbindung, ihrer zeitlichen sowie inter- und intrapersonellen Koordinierung untersucht (vgl. Kap. 1). Im Folgenden wird daher ein einführender Überblick über bestehende Klassifikationssysteme von Gesten sowie über ihre Funktionen gegeben (vgl. Kap. 3.2.1). Da die Klassifikationssysteme in ihrer Genese aus unterschiedlichen Forschungsperspektiven entstanden sind, weisen sie Differenzen in der Beschreibung und Benennung der Gestenformen und -funktionen auf. Ziel der Systematisierung der Klassifikationssysteme ist es somit, einerseits Differenzen sowie Überlappungen aufzudecken sowie andererseits terminologische Inkonsistenzen zu klären. Anschließend wird unter Berücksichtigung des Forschungsinteresses der vorliegenden Arbeit der Blick auf Gesten mit pragmatischer Funktion verengt, da sich diese in den Analysen als besonders relevant für das Argumentieren erwiesen haben (vgl. Kap. 3.2.2).

3.2.1 Gesten: Formen und Funktionen

In der Auseinandersetzung mit den Formen und Funktionen von Gesten ist über die Jahrhunderte²⁵ hinweg eine Vielzahl unterschiedlicher Gestenklassifikationssysteme und Gestentypologien entstanden, deren Heterogenität u. a. durch das jeweilige Forschungsinteresse, die Konzeptualisierung von Gesten und Sprache sowie auch durch technische Entwicklungen beeinflusst ist (vgl. für einen historischen Überblick Müller 1998; Kendon 2004). Das nachfolgende Kapitel wird eine Auswahl von vier Gestenklassifikationssystemen, namentlich von McNeill (1992, 2005), Bavelas et al. (1992, 1995) und Bavelas (1994), Kendon (2004) sowie Streeck (2009b), vorstellen, deren Studien aufgrund videographierter Daten konsequent multimodale Ressourcen einbeziehen und die die Untersuchung der (interaktiven) *Funktionen* von Gesten ermöglichen. Alle vier Klassifikationssysteme berücksichtigen neben Gesten mit semantisch referenziellen Funktionen, die einen Inhaltsbezug aufweisen, auch Gesten mit pragmatischen Funktionen, die den Forschungsschwerpunkt der vorliegenden Arbeit bilden. Um die Genese der einzelnen Klassifikationssysteme nachvollziehbar zu machen, werden sie ihren Forschungsperspektiven entsprechend eingeordnet. Trotz der begrifflichen Heterogenität innerhalb der vier Klassifikationssysteme kristallisieren sich mindestens zwei wiederkehrende Funktionen von Gesten heraus: Semantische referenzielle Funktionen, die an den Äußerungsinhalt anschließen, sowie pragmatische Funktionen, die u. a. zur Strukturierung von Äußerungen beitragen oder Informationen über die Sprechhandlung übermitteln. Die Funktionen werden am Ende des Teilkapitels (Kap. 3.2.1) komprimiert in Form eines tabellarischen Überblicks zusammengefasst.

McNeill (1992, 2005) untersucht die Funktionen von Gesten aus einem kognitionspsychologischen Verständnis heraus. Er begreift dabei Sprache und Gesten als Teil *eines* (psychologischen) Systems (vgl. auch McNeill 1985; vgl. abgrenzend dazu Kendon, Kap. 3.1.1), das Einblicke in das menschliche Denken erlaubt. Diesem kognitionspsychologischen Verständnis folgend können Gesten als Ausdrucksformen innerer kognitiver Prozesse gedeutet werden, die diese sichtbar machen (vgl. McNeill 1985, 1992). McNeills (1992, 2005) Untersuchungen basieren auf experimentellen psychologischen Settings, in denen Nacherzählungen zu einem Videoclip eines Cartoons elizitiert werden (vgl. dazu ausführlicher Kap. 4.2.1). In den Nacherzählungen beobachtet er eine Vielzahl spontan geäußelter sprecher*innenseitiger Gesten. Die identifizierten Gesten weisen sowohl semantisch referenzielle als auch pragmatische Bezüge zu den Äußerungen der Sprecher*innen auf und lassen sich, wie Abbildung 2 zeigt, in insgesamt vier Gestentypen (*Types of Gesture*, vgl. McNeill 1992: 12) unterteilen: in ikonische (*Iconics*), metaphorische (*Metaphorics*) und deiktische Gesten (*Deictics*) sowie *Beat* Gesten und Kohäsionsgesten (*Cohesives*).²⁶

²⁵ Während das vorliegende Kapitel auf Gestenklassifikationssysteme des ausgehenden 20. und beginnenden 21. Jahrhunderts eingeht, zeichnet Kendon (2004) nach, dass das Interesse an Gesten und ihren Funktionen für die Mündlichkeit bis in die klassische Antike zurückreicht.

²⁶ Die Kohäsionsgesten nimmt McNeill (1992: 78ff.) zwar selbst nicht in eine zusammenfassende Darstellung seines Gestenklassifikationssystems auf, da sie allerdings in ihrer Strukturierungsfunktion für den Aufbau von Äußerungen relevant werden, werden sie in der schematischen Übersicht (vgl. Abb. 2) ergänzt.

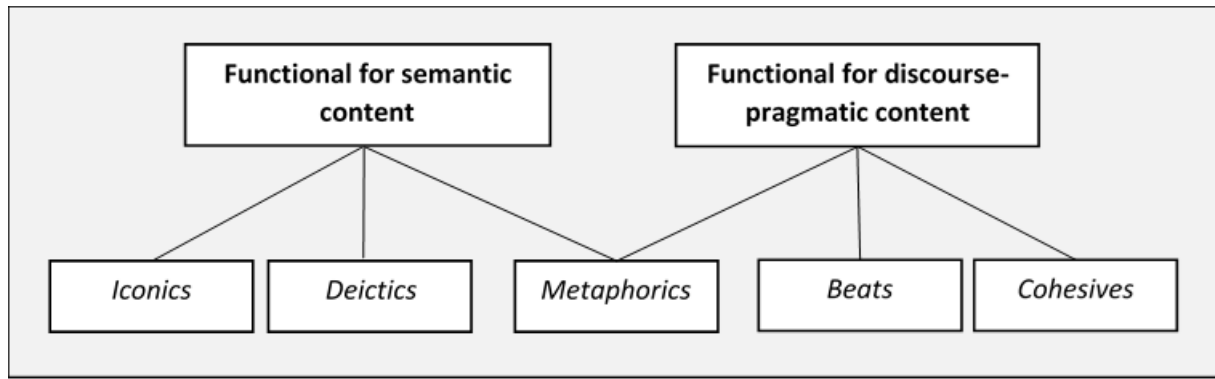


Abbildung 2: Übersicht über McNeills (1992, 2005) Gestenklassifikationssystem (eigene Darstellung)

Ikonische Gesten (Iconics) bilden spezifische Inhalte einer Äußerung ab und können nicht ohne die sprachliche Äußerung verstanden werden (vgl. McNeill und Levy 1982; McNeill 1986: 107, 1992: 145). Dabei können sich die Repräsentationen auf Ereignisse oder Gegenstände der konkreten Welt beziehen – unabhängig davon, ob es sich um imaginierte bzw. fiktive oder reale Ereignisse oder Gegenstände handelt – und den Inhalt der Äußerung illustrieren sowie diesen um weitere Informationen ergänzen (vgl. McNeill 1992: 12ff.).

Metaphorische Gesten (Metaphorics) sind ebenfalls als abbildende Gesten zu verstehen, die allerdings „abstract idea[s] rather than a concrete object or event“ (McNeill 1992: 12) repräsentieren. McNeill (2005: 45) veranschaulicht die Funktion der Geste anhand des folgenden Beispiels aus einer Nacherzählung: Während der Sprecher äußert „‘and the next scene is’, his left hand appeared to be supporting/presenting something“. Die Geste übernimmt an dieser Stelle die metaphorische Bedeutung von „‘I gave you that idea‘“ (McNeill 2005: 45) – ihr kommt somit eine semantische referenzielle Funktion zu. Zugleich wird der abstrakte Inhalt der Äußerung durch die nach oben ausgerichtete Handfläche dem Gegenüber auch präsentierend zur weiteren Aushandlung angeboten. I. d. S. übernimmt die Geste also auch pragmatische Funktionen (McNeill 2005: 45, vgl. auch McNeill 1992: 14f.).

Deiktische Gesten (Deictics) referenzieren auf konkrete oder abstrakte Objekte oder Räume (vgl. McNeill 1992: 18). Mit der Unterscheidung in das Verweisen auf konkrete vs. abstrakte Zeigeeziele hebt McNeill (2005: 40) hervor, dass „[a]bstract deixis creates new references in space“, während „concrete deixis finds references in it“. Das Zeigen auf abstrakte Zeigeeziele weist i. d. S. eine gewisse Ähnlichkeit mit metaphorischen Gesten, denn „[a]bstract pointing gestures imply a metaphorical picture of their own in which abstract ideas have a physical locus“ (McNeill 1992: 18).

Beats sind synchronisiert mit dem Rhythmus der Sprache und können einzelne Wörter oder ganze Äußerungen, mit denen sie auftreten, als relevant markieren, „not for its own semantic content, but for the discourse-pragmatic content“ (McNeill 1992: 15). Diese Gesten übernehmen wichtige Funktionen für die Struktur bzw. Strukturierung der Äußerung – tragen aber im Gegensatz zu ikonischen und metaphorischen Gesten keine abbildende Funktion. Sie unterscheiden sich außerdem in ihrer Art der Ausführung, denn während die vorangegangenen drei Gestentypen in die Phasen der Vorbereitung, des Gestenhöhepunktes sowie der Rückführung unterteilt werden können, besteht ein *Beat* üblicherweise aus „a simple flic of the hand or fingers up and down, or back and forth“ (McNeill 1992: 15; vgl. auch McNeill und Levy 1982). Kennzeichnend für diese Geste ist also ihre bi-phatische Bewegung.

Zu den Gesten mit pragmatischer Funktion gehören auch *Kohäsionsgesten* (*Cohesives*, vgl. McNeill 1992: 16)²⁷: Es handelt sich um Gestenwiederholungen, mit denen Äußerungselemente als zusammengehörig markiert werden – unabhängig von der primären Funktion der Geste. So behalten beispielsweise ikonische Gesten einerseits ihre abbildende Funktion bei, die sie für den Äußerungsinhalt übernehmen, während ihre Wiederholung gleichzeitig auf der Ebene der Äußerungsstrukturierung operiert. McNeill (1992: 16) nennt sie *Discourse Gestures*, die wie die *Beat* Gesten pragmatische Funktionen tragen, da sie relevante Informationen über die Äußerungs- bzw. Diskursstruktur beinhalten (vgl. McNeill 1992: 183).²⁸

Insgesamt unterscheidet die vorgestellte Gestenklassifikation nach McNeill (1992, 2005) also Gesten in ihrer abbildenden (*Iconics*, *Metaphorics*) bzw. nicht-abbildenden (*Deictics*, *Beats*, *Cohesives*) Funktion, wobei die ikonischen, metaphorischen und deiktischen Gesten eine referenzielle (oder semantische) Funktion tragen. Pragmatische Funktionen übernehmen *Beat* Gesten und Kohäsionsgesten sowie darüber hinaus auch metaphorische Gesten.

Bavelas et al. (1992, 1995) sowie Bavelas (1994) nehmen, wie McNeill (1992, 2005), aus einer experimentell-psychologischen Perspektive die Funktionen von Gesten – allerdings in einem dialogischen Setting – in den Blick. Sie unterscheiden dabei *Bewegungen der Hand ohne kommunikative Funktion*, wie beispielsweise das Reiben des Kinns oder das Drehen eines Rings (Bavelas et al. 1995: 397), von *Gesten mit kommunikativer Funktion* (*Conversational Gestures*). Während Erstere Bavelas et al. (1995) zufolge keine kommunikative Funktion tragen, können Letztere z. B. zur Illustrierung eines Äußerungsinhalts (*Topic Related Gestures*) oder zum interaktiven Geschehen beitragen (*Interactive Gestures*), indem sie z. B. den Sprecher*innenwechsel organisieren (vgl. Bavelas et al. 1995). I. d. S. tragen die sogenannten *Conversational Gestures* kommunikative Funktionen. Die Gesten bzw. ihre Funktionen werden im Anschluss an die nachfolgende Übersicht (vgl. Abb. 3) näher vorgestellt.

²⁷ McNeill (1992: 77) führt ebenfalls die sogenannten *Butterworths*, benannt nach Brian Butterworth, an, die u. a. während Wortsuchen auftreten. Diese grenzt McNeill (1992) von den bereits vorgestellten vier Gestentypen ab, da sie nicht äußerungsbegleitend produziert werden.

²⁸ Wenn auch die Funktion der *Cohesives* in ihrer Beschreibung der Strukturierung der Äußerung nachvollziehbar und ersichtlich erscheint, führt McNeill (1992) allerdings ein Beispiel an, in dem er die *Rückkehr* der Hände einer Sprecherin nach einem inhaltlichen Exkurs in ihre Ausgangsposition als *Cohesive* interpretiert. Im Verständnis der Verfasserin der vorliegenden Arbeit handelt es sich hierbei allerdings nicht um eine *Gestenwiederholung* i. S. einer Geste, die in einer bestimmten Phasierung, Bewegungsrichtung, Form etc. wiederholt wird. Allerdings kann die Rückkehr in die Ruheposition eine ähnliche Funktion wie *Cohesives* übernehmen, indem sie eine Rückkehr zum vorangegangenen Gespräch markieren kann.

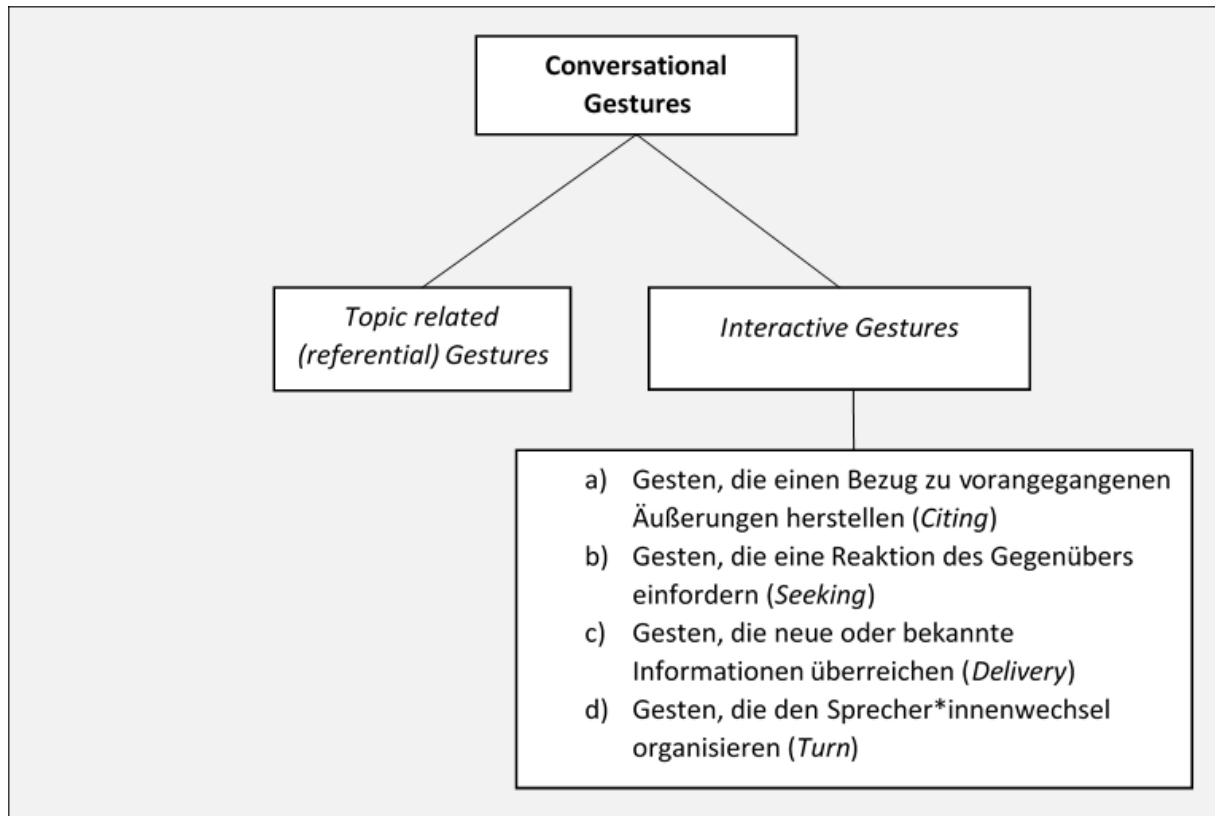


Abbildung 3: Übersicht über Gesten mit kommunikativer Funktion nach Bavelas et al. (1992, 1995) und Bavelas (1994) (eigene Darstellung)

Die sogenannten *Topic Related Gestures* (Bavelas et al. 1992: 472) beziehen sich auf den Inhalt einer Äußerung und können illustrierende, d. h. semantische Funktionen übernehmen, indem sie beispielsweise die Größe eines besprochenen Gesprächsgegenstands abbilden. *Interactive Gestures* sind hingegen nicht an den Äußerungsinhalt gebunden, sondern tragen zur Aufrechterhaltung und interaktiven Koordination der Interaktion bei (vgl. Bavelas 1994; Bavelas et al. 1995: 394). Durch diese Gesten können Interagierende beispielsweise eine gegenseitige Bezugnahme aufeinander anzeigen, die somit für alle Beteiligten auch visuell wahrnehmbar wird.²⁹ Typisch für diese Gesten ist ihre Ausrichtung auf ihr Gegenüber – bleibt dieser formorientierte Aspekt aus, handelt es sich gemäß den Autor*innen nicht um eine interaktive Geste.³⁰ Die interaktiven Gesten weisen insgesamt ein breites Funktionsspektrum auf, wie Bavelas et al. (1992, 1995) darlegen: So dienen die Gesten entweder dazu, einen Bezug zur Vorgängeräußerung anzuzeigen (*Citing*), zur Verständnisabsicherung oder als Hilfesuch (*Seeking*), zur Übergabe (un-)bekannter Informationen (*Delivery*) oder zur Rederechtsorganisation (*Turn*) (vgl. Bavelas et al. 1992: 473, 1995: 397). Die von Bavelas et al. (1992, 1995) herauspräparierten pragmatischen Funktionen der Gesten geben also Auskunft über die Intention der Sprechhandlung und tragen außerdem zur Organisation der Interaktion bei. Dabei ist die Aufgabe interaktiver Gesten, so Bavelas et

²⁹ Bavelas et al. (1992: 472) verweisen darauf, dass bereits in früheren Klassifikationssystemen wie von Efron (1941/1972) oder Ekman und Friesen (1969) Gesten beschrieben worden sind, die den „interactive gestures“ zugeordnet werden können, da sie interaktive Funktionen erfüllen.

³⁰ Die Annahme, dass Gesten nur dann eine interaktive Funktion übernehmen, wenn sie in Richtung des Gegenübers geäußert werden, kann im Kontext weiterer Forschungsarbeiten mittlerweile in Frage gestellt werden. So zeigt beispielsweise Kendon (2004) anhand der *Palm Up Lateral* Gesten, dass eine Geste durchaus interaktive Funktionen übernehmen kann, ohne dass sie auf den*die Rezipient*in ausgerichtet ist.

al. (1992: 470), „to aid the maintenance of conversation as a social system“.³¹ Indem die Forscher*innen die Funktion der Gesten nicht nur sprecher*innenseitig untersuchen, sondern zugleich auf ihre Funktion für die *gemeinsame* Interaktion fokussieren, liefern sie wichtige Erkenntnisse zur Erweiterung bisheriger Befunde zu pragmatischen Funktionen von Gesten (vgl. u. a. McNeill 1992).

Kendon ist aus Perspektive der Gestenforschung daran interessiert, welche (kommunikativen) Funktionen Gesten *für die* und *in der* sozialen Interaktion übernehmen. Davon ausgehend, dass Gesten und Sprache gemeinsam eine kommunikative, soziale Handlung konstituieren (vgl. Kap. 3.1.1), unterscheidet er in seiner Gestentypologie (vgl. Kendon 2004: 225) Gesten mit *semantisch referenzieller* Funktion und Gesten mit *pragmatischer* Funktion. Während Gesten mit semantisch referenzieller Funktion einen direkten Bezug zum Inhalt einer Äußerung aufweisen (vgl. auch oben, Bavelas et al. 1992, 1995), können Gesten mit pragmatischer Funktion die Äußerung von Sprecher*innen qualifizieren, indem sie beispielsweise Äußerungsinhalte modalisieren, abschwächen oder verstärken. Sie können aber auch wichtige Informationen zur *Sprechhandlung* beinhalten und tragen ferner zur *Diskursstruktur* bei. Die nachfolgende Übersicht (vgl. Abb. 4) skizziert diese Unterschiede.

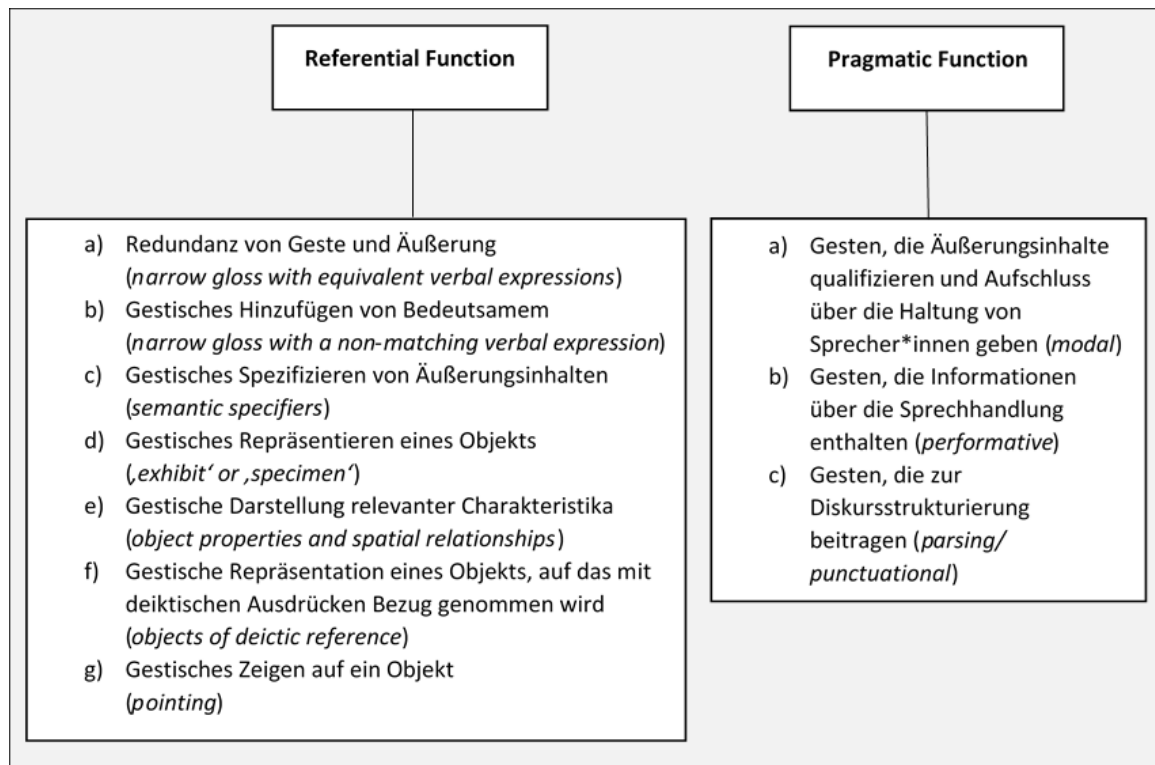


Abbildung 4: Übersicht über die Funktionen von Gesten nach Kendon (2004, 2017) (eigene Darstellung)

Gesten mit *referenzieller Funktion* weisen einen direkten Bezug zu dem Inhalt einer verbalen Äußerung auf. Dabei sind die Ressourcen, d. h. die sprachlichen sowie gestischen Ressourcen, reziprok aufeinander bezogen und tragen gegenseitig zur Bedeutungskonstitution bzw. zu einem Verständnis der gesamten Äußerung bei (vgl. Kendon 2004: 174). Es zeigen sich Unterschiede hinsichtlich des Beziehungsgefüges zwischen den gestischen und sprachlichen Ressourcen (vgl. Abb. 4): So können die Gesten i. S. von Emblemen der simultan produzierten Äußerung *entsprechen*; sie können der Äußerung bedeutende, nicht sprachlich realisierte Inhalte *hinzufügen*; Äußerungsinhalte *konkretisieren* bzw. spezifizieren; sie können Äußerungsinhalte *illustrieren*; sie können relevante Informationen (z. B. über die Größe, Form

³¹ Payrató und Teßendorf (2014: 1532) heben hervor, dass durch die Kategorie der konversationellen Gesten die *interaktive* Funktion von Gesten als wichtiger Bestandteil sozialer Interaktion Berücksichtigung findet.

etc. eines benannten Objekts) *ergänzen* sowie *Objekte* ergänzend *illustrieren* bzw. *Ereignisse enaktieren*, auf die anschließend mit *deiktischen Ausdrücken* Bezug genommen wird (vgl. Kendon 2004: Kap. 10). Zu den Gesten mit referenzieller Funktion gehören ebenfalls verweisende Gesten oder *Pointing Gestures*, durch die Sprecher*innen eine Referenz zu einem konkreten oder abstrakten Objekt herstellen können, das wiederum zur Bedeutungskonstitution der Äußerung beiträgt (vgl. Kendon 2004: 160, 2017: 168). Kendons Studien legen insgesamt ein äußerst differenziertes Bild des *Verhältnisses* zwischen sprachlichen und gestischen Äußerungen dar, in denen Gesten referenzielle Funktionen zukommen. Er deckt dabei die Komplexität auf, in der sprachliche und gestische Ressourcen miteinander verschränkt sind und gemeinsam Bedeutung konstituieren.

Gesten übernehmen allerdings nicht nur Funktionen zur semantischen Unterstützung des Äußerungsinhaltes, wie es Gesten mit referenzieller Funktion tun, sondern sie übermitteln ebenfalls Informationen über die *Implikationen* einer Äußerung bzw. die *Einstellung* der Sprecher*innen, über die *Sprechhandlung* sowie über die *Diskursstruktur*. Kendon (1995a, 2004, 2013, 2017) spricht dabei von Gesten mit *pragmatischer* Funktion, die sich in drei Subfunktionen (*modal*, *performative*, *parsing/punctuational*, vgl. Abb. 4)³² unterscheiden: Gesten mit modalisierender Funktion (*modal function*) können Äußerungsinhalte qualifizieren, indem sie beispielsweise den Inhalt einer Äußerung als Vermutung anzeigen. Sie können aber auch (implizite) Negationen markieren, beispielsweise wenn Gesten, in denen die Handfläche nach unten orientiert ist und die Hand in Richtung der Körperseite bewegt und/oder der Kopf geschüttelt wird, äquivalent oder ergänzend zur sprachlichen Äußerung eingesetzt werden. Werden die Gesten mit besonderem Nachdruck und möglicherweise zusätzlich zu explizit-sprachlichen Negationen geäußert, können sie den Inhalt einer Äußerung intensivieren und geben somit auch Aufschluss über die Haltung eines*einer Sprecher*in (vgl. Kendon 2004, 2013, 2017). Gesten mit performativer Funktion (*performative function*) können Hinweise auf die intendierte Sprechhandlung beinhalten, z. B. indem eine Bitte durch aneinandergelegte Hände (vgl. auch *Praying Hands* Kendon 2004) angezeigt wird. Auch weniger emblematische Gesten können performative Funktionen tragen. Beispielsweise können Interagierende eingebrachte Gesprächsgegenstände auf der geöffneten Handfläche ‚präsentieren‘ und einander zur weiteren Aushandlung anbieten. Bei Gesten mit diskursstrukturierender Funktion (*parsing/punctuational function*) handelt es sich um Gesten wie beispielsweise das alternierende Zurseitelehnen des Kopfes zur Strukturierung einer Liste bzw. der Listenelemente oder aber händische *Beats*, die spezifische Äußerungsteile besonders hervorheben (vgl. Kendon 2017: 171).

³² Die Einordnung der Subfunktionen von Gesten mit pragmatischer Funktion variieren über den langen Zeitraum von Kendons Forschungstätigkeit. In dem Artikel *Gestures as illocutionary and discourse structure markers in Southern Italian conversation* (Kendon 1995: 248) unterscheidet er „*illocutionary marker gestures*“ „[that] serve as a way of indicating the type of interactional move a turn-taker is making“ von „*discourse unit marker gestures*“ „[that] serve to exhibit the role within a discourse of a particular phrase or segment of speech, relative to other segments“. Die *Illocutionary Marker Gestures* wird er in seiner Gestentypologie (2004) den performativen Gestenfunktionen und die *Discourse Unit Marker Gestures* der diskursstrukturierenden Funktion (*parsing*) zuordnen (vgl. Kendon 2004: 159, Fußnote 1). In Kendon (2004) ist die o. g. operationale Funktion, die er in seinem Artikel (2017) anführt, noch der modalisierenden Funktion beigeordnet: „These [Open Hand Prone] gestures may be regarded as *operator* gestures (and may be included in the broader group of *modal* pragmatic gestures, accordingly), just as the spoken language particles of ‚no‘ and ‚never‘ and the like are verbal operators“ (Kendon 2004: 263, Hervorh. i. O.). In der vorliegenden Übersicht wird sich der Auffassung angeschlossen, dass Gesten mit operationaler Funktion weitestgehend modalisierende Funktionen übernehmen. Daher führt die Übersicht die Unterscheidung zwischen performativen, modalisierenden und diskursstrukturierenden Funktionen von Gesten fort, wie sie Kendon (2004) vornimmt.

Insgesamt offenbaren Kendons Studien, die dezidiert das Verhältnis sprachlicher und gestischer Ressourcen in den Blick nehmen, dass die Funktionen von Gesten sich zum einen nur in ihrem direkten Äußerungskontext bestimmen lassen und die Form der Geste nicht *a priori* ihre jeweilige Funktion festlegt; vielmehr kommt ihnen erst in der Orchestrierung mit sprachlichen sowie weiteren körperlichen Ressourcen diese Funktion zu. Es handelt sich also um „a typology of functions, not of gestures“ (Kendon 2004: 225).

Streeck (2009b) entwickelt eine Gestenklassifikation, die sich an Kendons Differenzierung von referenziellen und pragmatischen Funktionen von Gesten anlehnt. Allerdings legt Streeck (2009b, 2015) ein phänomenologisch sowie praxeologisch orientiertes Verständnis von Gesten zugrunde (vgl. Streeck 2009b, 2013, 2015), indem er ihren Ursprung als aus konkreten physischen und kollaborativen Handlungen heraus entstanden begreift (vgl. Streeck 2006). So wird das Präsentieren eines Gesprächsgegenstandes (vgl. oben, Kendon; vgl. Müller 2004; Kendon 2004), der an Rezipient*innen auch zur weiteren Handlung ‚übergeben‘ werden kann, von Streeck (2006, 2009b) aufgrund der Form und der Bewegungsrichtung der Hand im Ursprung auf die manuelle Handlung zurückgeführt, die Menschen vollziehen, wenn sie einander ein physisches Objekt überreichen. Anders als beispielsweise McNeill (1992, 2005) oder Kendon (2004) orientiert Streeck (2009b) seine Gestenklassifikation nicht vornehmlich an der Unterscheidung von referenziellen und pragmatischen Funktionen, sondern an der (Ent-)Kopplung von Geste und Umwelt – und dennoch schimmern die bereits vorgestellten Funktionen auch in dieser Klassifikation durch. Die nachfolgende Übersicht (vgl. Abb. 5) legt die insgesamt sechs Funktionen von Gesten dar.

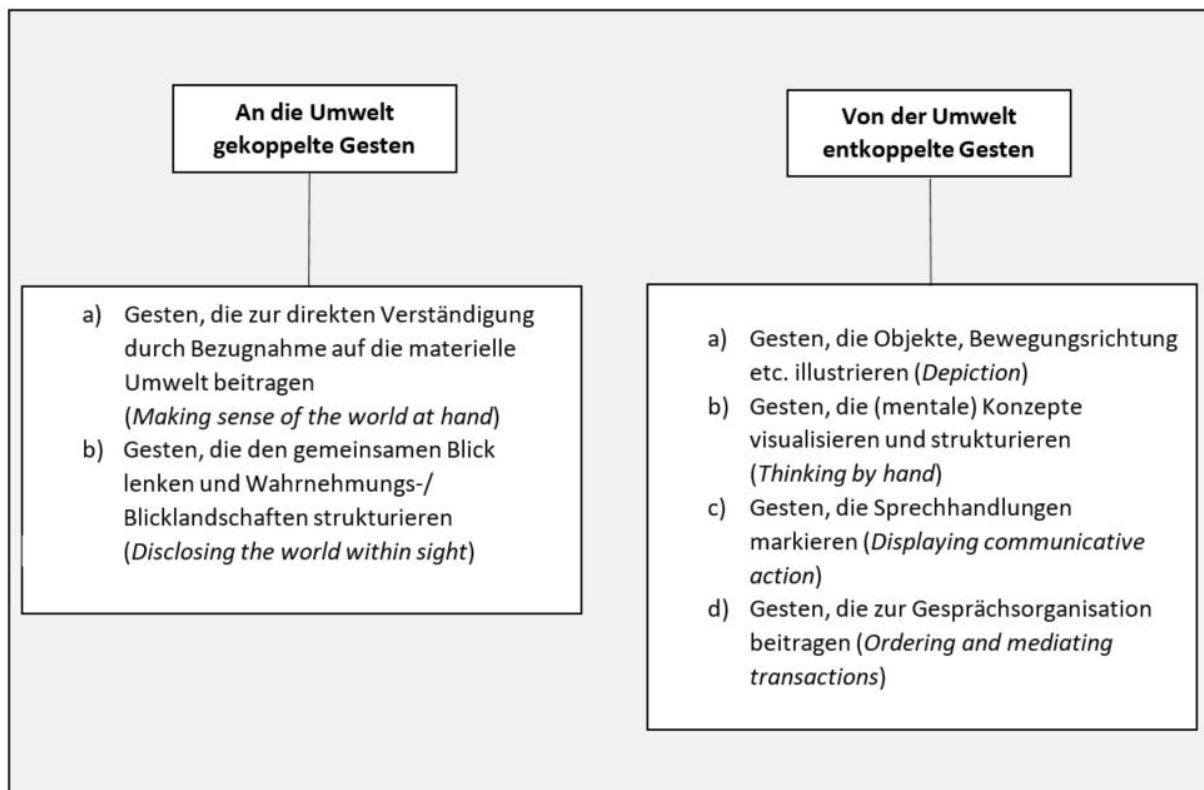


Abbildung 5: Übersicht über die sechs Gestenökologien nach Streeck (2009) (eigene Darstellung)

Streecks (2009b: 8) Gestenklassifikation beruht auf der Beschreibung sechs unterschiedlicher Gestenökologien, die das Verhältnis zwischen „gestural activity“ „with the world, with the concurrent speech, and with the interactants“ zu beschreiben suchen. Der Begriff der Ökologie verweist auf die Eingebundenheit der Geste in ihre Umwelt, in der die Interaktion stattfindet (vgl. auch Streeck 2013). Die folgenden sechs Gestenökologien differenziert Streeck (2009b: 8ff.) voneinander:

1. Die erste Gestenökologie bezieht sich auf Gesten, mit denen Interagierende materielle Gegenstände der Umwelt, die sich in ihrer Reichweite befinden, berühren und in ihre sprachliche Äußerung einbeziehen (*Making sense of the world at hand*) können. Die Gesten können beispielsweise der *Orientierung* dienen, wenn durch eine zeigende Geste ein einzelnes Objekt unterscheidbar von anderen Objekten abgegrenzt wird. Gesten dieser Ökologie können zur (*Er-*)*Klärung* von Äußerungsinhalten beitragen, wenn Sprecher*innen beispielsweise Handlungen enaktieren. Es können auch *taktile* Informationen z. B. über die Beschaffenheit von Objekten übermittelt werden, während die Sprecher*innen, nicht aber die Rezipient*innen, das Objekt selbst berühren. Es kann aber auch das Objekt selbst zum *Demonstrieren* bestimmter Äußerungsinhalte verwendet werden. Die Äußerungen und Gesten der Sprecher*innen sind also unmittelbar an die umgebende materielle Umwelt zurückgekoppelt bzw. mit dieser verbunden (vgl. Streeck 2009b: 59–76). Aufgrund der Funktion der Gesten, mit denen Interagierende Objekte aus der Umwelt zu einer gemeinsamen Verständnisabsicherung und Organisation ihrer Tätigkeiten einbeziehen, bezeichnet Streeck (2009b: 66–69) die Gesten auch als *Clearing Gestures*.
2. Anders als in der ersten Gestenökologie stellen in der vorliegenden Gestenökologie (*Disclosing the world within sight*) Gesten eine Verbindung zu *entfernten* Objekten her. So können beispielsweise Zeigegesten (*pointing gestures*; vgl. Streeck 2009b: 8f., 2015: 426) in ihrer Verlängerung als Vektoren auf das Zeigziel hin dienen und, wenn mehrere Zeigziele vorliegen, können diese gestisch verbunden, strukturiert und arrangiert werden. Auf diese Weise sind auch in dieser Gestenökologie die Gesten sowie die Interagierenden direkt mit der sie umgebenden Umwelt verbunden.
3. Gesten, die einer Repräsentation oder einem Abbild (*Depictive Gestures*) „as a substitute of the world“ (Streeck 2009b: 85) dienen, können Objekte, Handlungen, Formen, Bewegungen, räumliche Gegebenheiten oder Strukturen etc. visualisieren, indem Sprecher*innen beispielsweise Umrisse nachzeichnen, modellieren oder Handlungen mimetisch enaktieren. Dabei entstehen adaptive Referenzen zwischen dem Repräsentierten und dem Äußerungsinhalt, denn „gestures do not usually represent what we see, but what we know about something“ (Streeck 2009b: 150). Da diese Gesten keine direkte Verbindung zur Umwelt herstellen, beschreibt Streeck (2009b: 85) sie als entkoppelt (*uncoupled*) von der Welt.
4. Gesten können auch zur Visualisierung und Strukturierung eines (mental)en Konzepts der Sprecher*innen beitragen (*Thinking by hand: Gesture as conceptual action*). Dabei repräsentieren diese Gesten (*Conceptual Gestures*) keine Objekte, Handlungen etc., sondern zugrundeliegende Konzeptualisierungen von Äußerungsinhalten. Verdeutlicht werden kann dies durch ein von Streeck (2009b: 151) angeführtes Beispiel: Wenn z. B. die Geste eines Sprechers einen Baumstamm visualisiert, während er über ‚Selbstständigkeit‘ spricht, dann ist die Geste kein Abbild des Begriffs, sondern sie visualisiert ein Konzept von Selbstständigkeit als „literally the ability to stand on one’s own“ (Streeck 2009b: 151). Die Gesten resultieren also aus einem mentalen Prozess und dienen sprecher*innenseitig einer Strukturierung innerer Konzepte. Anders als *Depictive Gestures* tragen sie allerdings keine abbildende Funktion für den*die Rezipient*in, was nicht zuletzt darin deutlich wird, dass sich weder Sprecher*in noch Rezipient*in auf die Geste hin orientieren. Auch diese Gesten sind von der Umwelt entkoppelt.

5. Unabhängig von den referenziellen Funktionen können Gesten auch pragmatische Funktionen übernehmen (*Displaying communicative action*). Interagierende können beispielsweise durch Gesten Sprechhandlungen (*Speech-handling Gestures*) markieren wie das Präsentieren einer Äußerung auf der geöffneten Handfläche. Ebenfalls können Sprecher*innen Äußerungsinhalte durch *Beats* hervorheben oder durch abstrakte Zeigegesten auf bereits benannte Äußerungsinhalte zurückverweisen. Außerdem ist es Interagierenden möglich, durch Gesten eine eigene Haltung auszudrücken. Wenn beispielsweise ein Äußerungsinhalt auf einer Handfläche präsentiert bzw. angeboten wird, aber die Finger leicht gekrümmt sind, dann kann die Form der Gestenäußerung, so Streeck (2009b: 188), eine unsichere Haltung des*der Sprecher*in zum Ausdruck bringen, die die Sprechhandlung als „weak offering[]“ (Streeck 2009b: 188) rahmt.
6. Zuletzt können Gesten (*(Meta-)pragmatic Gestures*), die auf die in die Interaktion eingebundenen Interagierenden ausgerichtet sind, zur Organisation der Interaktion beitragen (*Ordering and mediating transactions*), indem die Interagierenden sich beispielsweise gegenseitig berühren, um die Aufmerksamkeit des Gegenübers zu elizitieren oder durch zeigende Gesten einander den *Turn* zuweisen (vgl. Streeck 2009b: 10). Diese Gesten tragen dazu bei, „[to] regulate the actions of others“ (Streeck 2015: 427).³³

Während die ersten vier beschriebenen Funktionen von Gesten referenzielle Funktionen tragen, da sie Bezüge zum Äußerungsinhalt aufweisen, übernehmen die letzten beiden Gesten pragmatische Funktionen, da sie relevante Informationen über Sprachhandlungen, sprecher*innenseitige Haltung oder auch die Diskursstrukturierung enthalten. Ähnlich wie Bavelas et al. (1992, 1995) die kommunikative Funktion von Gesten an der Ausrichtung auf den*die Rezipient*in festmachen, wird nun in Streecks (2009b) Beschreibung der Gestenökologien eine konsequente Berücksichtigung der Umwelt relevant. Dies ermöglicht, auch objektgebundene Gesten in ihren Funktionen für die Interaktion einzubeziehen.

Zusammenfassend zeigen die vorgestellten Gestenklassifikationssysteme und -typologien zwar Unterschiede in der Genese sowie ihrer jeweiligen Strukturierung (abbildende vs. nicht-abbildende, *Topic Related* vs. *Interactive Gestures*, referenzielle vs. pragmatische Gesten bzw. Funktionen von Gesten); jedoch werden trotz der Varianz insgesamt zwei übergeordnete, stabil auftretende Funktionen sichtbar: Gesten mit *referenzieller* sowie Gesten mit *pragmatischer* Funktion. Die als Essenz aus den vorgestellten Gestentypologien gewonnenen Funktionen von Gesten, die allesamt in aktuelle Forschungen hineinreichen (vgl. zsf. u. a. Bohle 2014a), werden im Folgenden (vgl. Tab. 1) in ihren wichtigsten Grundzügen systematisiert und tabellarisch vorgestellt. Dies ermöglicht, die in den vier vorgestellten Klassifikationssystemen teils terminologisch divergierenden und/oder überlappenden Beschreibungen der Gesten und ihre Funktionen übersichtlich zusammenzufassen. Die getrennte Vorstellung der einzelnen Funktionen resultiert aus den vorgestellten Klassifikationssystemen und dient hier der Übersichtlichkeit – in der Interaktion kann eine Gestenäußerung jedoch auch mehrere Funktionen gleichzeitig übernehmen.

³³ Bei den letzten beiden Gestenökologien handelt es sich um eine künstliche Trennung, denn „[m]any gestures blend the display of communicative action with the regulation of the behavior of others, and the distinction between these modes is of only limited analytic value“ (Streeck 2009: 10).

Gesten mit referenzieller Funktion	Gestenbezeichnungen abhängig vom Klassifikationssystem	Gesten mit referenzieller Funktion operieren auf semantischer Ebene, da sie einen Bezug zum Inhalt der Äußerung aufweisen.
	Iconics (McNeill 1992) Topic (related) Gestures (Bavelas et al. 1992, 1995; Bavelas 1994) Referentiell/Representational Gestures (Kendon 2004) Depiction (Streeck 2009b)	• Abbildende Gesten tragen zur Visualisierung spezifischer Aspekte der verbalsprachlichen Äußerung bei (vgl. McNeill 1992; Bavelas et al. 1992, 1995; Bavelas 1994; Kendon 2004; Streeck 2009b). • Sie zeigen eine große Ähnlichkeit zu dem von ihnen Illustrierten in Größe, Form, Bewegung, Umrissen etc. auf (vgl. McNeill 1992: 78; Kendon 2004). • Sie können von Beteiligten der Interaktion aufgrund der Ähnlichkeit zwischen den Gesten und dem Illustrierten sowie aufgrund des Weltwissens der Interagierenden in ihrer Bedeutung enkodiert werden (vgl. Streeck 2009b: 119). • Abbildende Gesten können reale oder imaginierte Objekte modellieren oder nachzeichnen, imaginierte Räume aufbauen, Personen sowie Objekte darin platzieren, (räumliche, örtliche etc.) Relationen darstellen und Handlungen enactieren (vgl. u. a. Kendon 2004: 161ff.; Streeck 2009b).
	Clearing (Streeck 2009b)	• <i>Clearing</i> Gesten beziehen die materielle Umwelt mit ein. • Sie übernehmen u. a. die Funktion, Objekte voneinander zu unterscheiden; durch Enactieren Äußerungsinhalte verständlich zu machen; oder taktile Informationen z. B. über die Beschaffenheit von Objekten zu übermitteln (vgl. Streeck 2009b).
	Deictics (McNeill 1992) Pointing Gesture (Kendon 2004) Pointing (Streeck 2009b)	• Deiktische Gesten stellen zwischen der verweisenden Ressource selbst, z. B. einem Finger oder einer Hand (vgl. Kendon 2004), und dem Referenten einen Bezug her. • Sie können auch mehrere Referenten miteinander in Beziehung setzen sowie, je nach gestischer Ressource, Informationen über den Referenten enthalten. Z. B. können Gesten mit der Handfläche auf räumliche Areale, Indexfingergersten auf spezifische Punkte verweisen (vgl. Streeck 2009b). • Das Zeigeziel kann sich auf die konkret umliegende Räumlichkeit der Interagierenden beziehen, aber auch auf abstrakte Zeigeziele verweisen (vgl. McNeill 1992; Kendon 2004). • Deiktische Gesten treten i. d. R. im Verbund mit sprachlichen deiktischen Ausdrücken, Blickausrichtungen der zeigenden Personen in Richtung des Zeigeziels und Referenten, etc. auf (vgl. Kendon 2004; Streeck 2009b)
	Metaphorics (McNeill 1992) Conceptual Gestures (Streeck 2009)	• Metaphorische (vgl. McNeill 1992) bzw. konzeptuelle Gesten (vgl. Streeck 2009b) bilden abstrakte mentale Konzepte ab. So kann beispielsweise eine Geste, die der Visualisierung eines Baumstammes ähnelt, das Konzept von ‚Selbstständigkeit‘ metaphorisch darstellen (vgl. Streeck 2009b: 151) – und, reziprok, um das Konzept der ‚Stabilität‘ ergänzen. • Indem die Geste zur Verkörperung kognitiver Prozesse beiträgt, wird sie als eine Hintergrundaktivität (vgl. <i>background activity</i> Streeck 2009b: 151) beschrieben, auf die sich weder Sprecher*innen noch Rezipient*innen hin orientieren. • Metaphorische Gesten können auch pragmatische Funktionen übernehmen, wenn sie beispielsweise das abgebildete Konzept zugleich dem*der Rezipient*in präsentieren (vgl. McNeill 1992: 29).

Gesten mit pragmatischer Funktion	Gestenbezeichnungen abhängig vom Klassifikationssystem	Gesten mit pragmatischer Funktion operieren auf der Ebene des „kommunikativen Handelns“ (Streeck 2007: 159) – sie sind nicht an den Inhalt einer Äußerung zurückgebunden.
	<i>Discourse Gestures</i> (McNeill 1992) <i>Cohesives</i> (McNeill 1992) <i>Interactive Gestures</i> (Bavelas et al. 1992) <i>Pragmatic Gestures</i> (Kendon 2004; Streeck 2009b) <i>Meta-pragmatic Gestures</i> (Streeck 2009b) <i>Speech-handling Gestures</i> (Streeck 2009b)	- Gesten mit pragmatischer Funktion können Hinweise auf – die jeweilige Sprechhandlung, – die Einstellung der Sprecher*innen gegenüber der Aussage sowie – die interne Strukturierung der Äußerung beinhalten – und sie können zur Gesprächs- sowie Rederechtsorganisation beitragen (vgl. McNeill 1992; Bavelas et al. 1992, 1995; Bavelas 1994; Kendon 2004; Streeck 2009b).

Tabelle 1: Zusammenfassende Übersicht über die Funktionen von Gesten unter Berücksichtigung der unterschiedlichen Klassifikationssysteme

Aus der Darstellung der Gestenklassifikationssysteme sowie der Beschreibung von Gesten in ihren Formen und Funktionen geht ein durchaus inkonsistentes Begriffsbild hervor, das an dieser Stelle nicht gelöst, aber aus Gründen der Transparenz benannt werden muss: Die *Benennung* der Gesten greift mal Aspekte auf, die direkt mit ihrer Gestalt bzw. ihrer konkreten Funktion verbunden sind (wie z. B. bei *Zeigegesten* sowie *abbildenden* Gesten), mal beschreibt sie die Beziehungen, die eine Geste zu Objekten in der Umwelt bzw. der sprachlichen Äußerung eingeht (z. B. bei *referenziellen* Gesten), und mal wird ihre übergeordnete Funktion für die Interaktion zentral in den Fokus gerückt (z. B. bei *pragmatischen* Gesten). Dabei weisen nicht nur die Benennungen der Gesten selbst Mehrdeutigkeiten auf und führen zu einer Vermischung der Form- und Funktionsebene, sondern hinzutritt, dass diese Attribuierung sowohl den Gesten als auch den Funktionen zugewiesen wird (*referenzielle* Gesten vs. Gesten mit *referenzieller* Funktion). Die Autorin der vorliegenden Arbeit schließt sich Kendon (2004: 225) an, der pointiert herausstellt, dass nicht die *Gesten* pragmatisch oder semantisch seien, sondern dass sich ihre Funktionen immer erst aus dem Kontext, in dem sie von Interagierenden genutzt werden, ergibt. Die vorliegende Arbeit wird somit ihr analytisches Vorgehen daran orientieren, die *Funktionen* von Geste in ihrem Ressourcenverbund einer multimodalen Gestalt für den jeweiligen sequenziellen Kontext zu beschreiben. Denn, um es mit Bavelas (1994: 204) zusammenzufassen, „the goal of analysis should not be to decide in which category we should put a gesture (or all gestures) but rather to discover at least some of the things a gesture is doing at its particular moment in the conversation“.

Im Anschluss an die Vorstellung der unterschiedlichen Gestenklassifikationssysteme wird unter Berücksichtigung des Forschungsinteresses der vorliegenden Arbeit im folgenden Kapitel 3.2.2 auf Gesten mit pragmatischer Funktion fokussiert.

3.2.2 Gesten mit pragmatischer Funktion: Interaktionsbezogene und diskurs- bzw. äuerungsstrukturierende Gesten

Für die kommunikative Gattung des mündlichen Argumentierens kann angenommen werden, dass gerade Gesten mit pragmatischer Funktion eine relevante Rolle zukommt, denn sie liefern wichtige Hinweise z. B. über die Haltung von Sprecher*innen oder Implikationen einer Äußerung, die möglicherweise oberflächensprachlich nicht markiert werden (vgl. Kap. 3.2.1). Da die vorliegende Studie die Forschungsfrage adressiert, welche Funktionen multimodale Ressourcen und v. a. Gesten mit *pragmatischer Funktion* in ihren Ressourcenverbünden für das mündliche Argumentieren übernehmen (vgl. Kap. 1), werden im nachfolgenden Kapitel 3.2.2.1 einschlägige Forschungsbefunde zu Gesten mit pragmatischen Funktionen zusammenfassend vorgestellt, die über die bereits präsentierten Befunde in Kapitel 3.2.1 hinausgehen. Dies ermöglicht, in der Rekonstruktion und systematischen Beschreibung der multimodalen Gestalten in Kapitel 8 auf diese Erkenntnisse zurückzukommen. Da Gesten mit pragmatischer Funktion z. T. eine gewisse Beständigkeit im Verhältnis zwischen der formalen Gestalt und der Bedeutung der Geste aufweisen, werden anschließend in Kapitel 3.2.2.2 auch sogenannte *Gestenfamilien* vorgestellt (vgl. u. a. Kendon 2004; Müller 2004; Bressemer und Müller 2014a, 2014b; Ladewig 2014b). Auf diesen Systematisierungen aufbauend werden dann in Kapitel 3.3. Funktionen pragmatischer Gesten beleuchtet, die für das mündliche Argumentieren relevant werden.

3.2.2.1 Kategorisierung pragmatischer Funktionen

Die im vorliegenden Kapitel vorgenommene Kategorisierung der pragmatischen Funktionen orientiert sich an den dargelegten Funktionsbereichen der tabellarischen Übersicht des vorangegangenen Kapitels 3.2.1 (vgl. Tab. 1). Es werden daran anlehnend die folgenden Funktionen unterschieden: *performative* Funktionen von Gesten, die Informationen über die Sprechhandlung beinhalten; *modalisierende* Funktionen von Gesten, die zur (weiteren) Qualifizierung einer Äußerung beitragen und Hinweise über die

Haltung der Sprecher*innen enthalten; *äußerungsstrukturierende* Funktionen von Gesten, mit denen Interagierende ihren eigenen Beitrag strukturieren können; sowie *interaktionsorganisatorisch* Funktionen von Gesten, die zur Gesprächs- und Sprecher*innenwechselorganisation beitragen. I. S. einer besseren Nachvollziehbarkeit werden einleitend in tabellarischen Übersichten (vgl. Tab. 2-5) konkrete Funktionen der Gesten zusammenfassend vorangestellt. Insgesamt werden hierbei nun v. a. diejenigen Funktionen näher beleuchtet, die von besonderer Relevanz für die empirischen Untersuchungen (vgl. Kap. 8 & 10) sein werden, sodass die Liste der Funktionen an dieser Stelle keinen Anspruch auf Vollständigkeit erhebt.

Performative Funktion

Gesten mit performativer Funktionen geben Hinweise auf die vollzogene Sprechhandlung, wie die nachfolgende Übersicht zusammenfassend skizziert (vgl. Tab. 2).

	Beschreibung	Funktion
Performative Funktion	Anzeige der Sprechhandlung	• Präsentieren/Anbieten von Gesprächsgegenständen • Erbitten einer Rückmeldung • Dissens markierendes Hinterfragen • Zurückweisen • Verwerfen von Äußerungsinhalten

Tabelle 2: Übersicht über performative Funktionen gestischer Ressourcen

Gesten mit performativer Funktion, die auf der Ebene der Sprechhandlung oder auch kommunikativen Handlung agieren, sind bei Streeck (2009b), Kendon (2004) und auch Müller (2004) assoziiert mit „practical everyday actions“ (Müller 2004: 235), die Aspekte ihrer funktionalen Bedeutung beibehalten.^{34,35} So zeigen Müller (2004) und Kendon (2004), dass Sprecher*innen abstrakte diskursive Objekte auf der ausgestreckten und nach oben ausgerichteten Handfläche (*Palm Up* Geste), wie einen materiellen Gegenstand, präsentieren und den Rezipient*innen zur weiteren Diskussion in der Interaktion anbieten. Wird die Geste am Ende eines *Turns* von Sprecher*innen weiter gehalten, kann die Funktion der Geste eine Veränderung erfahren: Die Bedeutung vom Anbieten eines abstrakten diskursiven Gesprächsgegenstandes verschiebt sich dann zum Erbitten einer Rückmeldung (vgl. Streeck 2007): Solange die Geste gehalten wird, besteht für das Gegenüber die Verpflichtung fort, den Gesprächsgegenstand in einem eigenen Beitrag aufzunehmen. Die gestische Ressource trägt hier also zur Elizitierung einer Rückmeldung bei.

Gesten können auch weitere kommunikative Handlungen anzeigen, wie z. B. das Bitten durch sogenannte *Praying Hands* (vgl. Kendon 1995a). Dabei werden die Handflächen vor dem Körper aneinandergelagt. Interagierende können durch gestische Ressourcen ihren Beitrag aber auch als eine neutrale oder auch Dissens markierende Frage rahmen. Im italienischen Sprachraum wird hierzu die sogenannte *Mano a Borsa* eingesetzt, bei der Daumen und Fingerspitzen einer Hand aneinandergelagt werden und die Handfläche nach oben orientiert ist und die Geste vor dem Oberkörper des*der Sprecher*in ausgeführt wird (vgl. Kendon 1995a). Auch Zurückweisungen oder Unterbrechungen anderer Beteiligter können sichtbar angezeigt werden, wie Kendon (2004: 263) für Gesten darlegt, bei denen die Handfläche

³⁴ Vgl. auch Ladewig und Hotze (2023: 94), die bislang fehlende Untersuchungen dieser Alltagshandlungen und deren Zusammenhang zur Genese von Gesten und ihrer Bedeutung kritisch anmahnen. Die Autorinnen legen selbst eine exemplarische Studie zu der Handlung des Schlagens unter Kindern vor, in der sie die Zusammenhänge zwischen der Handlung und der Bedeutungskonstitution der *als-ob*-Geste rekonstruieren.

³⁵ Der Übergang von (manuellen) Handlungen zu Gesten verläuft, so Ladewig und Hotze (2023: 94) in Bezug auf Harrison (2018), über „a stepwise decoupling of the manual movement from its environment and an increase of ‚communicative explicitness‘ (Andrén 2014b: Section 1.2)“.

des*der Sprecher*in vertikal in Richtung des*der Rezipient*in ausgerichtet ist. Durch sogenannte *Throwing Away* Gesten, wie das Anheben und Fallenlassen der Hand, können Interagierende außerdem das Verwerfen von Äußerungsinhalten, entweder der eigenen oder einer vorangegangenen Äußerung, markieren, wie Bressems und Müller (2014a) rekonstruieren.

Den exemplarisch vorgestellten Gesten mit performativer Funktion ist gemein, dass sie anzeigen, welche Sprechhandlung bzw. welche kommunikative Funktion sie für die Äußerung der Sprecher*innen übernehmen: Sie können *anbieten, bitten* bzw. *einfordern, fragen, zurückweisen, unterbrechen, verwerfen* etc.

Modalisierende Funktion

Gesten mit modalisierender Funktion tragen zur Qualifizierung einer Äußerung und Markierung der sprecher*innenseitigen Haltung bei, wie die nachfolgende Übersicht zusammenfassend darlegt (vgl. Tab. 3).

	Beschreibung	Funktion
modalisierende Funktion	Qualifizierung einer Äußerung & Markierung der sprecher*innenseitigen Haltung	• Anzeigen und Verstärken von Affirmation/Negation • Anzeigen von Un-/Sicherheit (epistemisch) • Markieren von ‚delikaten‘ Äußerungsinhalten (affektiv) • Anzeige von Distanzierung

Tabelle 3: Übersicht über modalisierende Funktionen von gestischen Ressourcen

Äußerungsinhalte können mittels gestischer Ressourcen (weiter) qualifiziert werden. So können z. B. Gesten, bei denen die Sprecher*innen die nach unten ausgerichtete(n) Handfläche(n) von der Mitte des Körpers nach außen bewegen, Negationen (sowie auch Affirmationen) einerseits explizit markieren (vgl. Kendon 2004: Kap. 13, 281f.) sowie andererseits Negationen auch intensivierend verstärken.

Körperlich-sichtbare Ressourcen wie Gesten können neben der Qualifizierung von Äußerungen (i. S. von Negationen und Affirmationen) ebenfalls Aufschluss über die epistemische sowie auch affektive Haltung von Sprecher*innen geben. Die epistemische Haltung kann beispielsweise als unsicher identifiziert werden, wenn die Äußerung als „weak offering[]“ (Streeck 2009b: 188) durch eine ungespannte *Palm Up* Geste präsentiert wird. Bei dieser werden die gekrümmten Finger nur lose gehalten (vgl. auch Beaupoil-Hourdel und Debras 2017; Cooperrider et al. 2018). Epistemische Sicherheit kann hingegen beispielsweise durch gestische Evidenzmarkierung ausgedrückt werden, indem Sprecher*innen die nach oben geöffneten Handflächen nach außen bewegen (vgl. *Palm Laterals* Kendon 2004) und mit einem Schulterzucken kombinieren (vgl. Cooperrider et al. 2018), um den Inhalt einer Äußerung als ‚offensichtlich‘ anzuzeigen (vgl. Kendon 2004: 265).

Ferner können Sprecher*innen einen Äußerungsinhalt auch als ‚gefährlich‘ oder ‚delikat‘ markieren, wie Bressems und Müller (2014a: 1584) für die sogenannte *Shaking Off* Geste rekonstruieren, bei der die Sprecher*innen die lose Hand, abgeknickt vom Handgelenk, schnell vor und zurück schütteln. Diese Gesten können im Kontext von Bewertungen, wie z. B. in der Äußerung „billarziose for instance (-) a **terrible** disease“ (Bressems und Müller 2014a: 1584, Hervorh. i. O.), auftreten. Die multimodalen Ressourcen, die gemeinsam mit der sprachlichen Ressource eine potenzielle Gefahr oder einen delikaten Inhalt markieren, geben somit Aufschluss über die affektive Haltung, *affective stance*, der Sprecher*innen (vgl. u. a. Selting 2012; Ruusuvuori 2013) und dadurch auch über eine direkte, affektive und emotionale Involviertheit der Sprecher*innen.

Doch auch Distanzierungen zu Inhalten der eigenen Äußerung sowie auch gegenüber Äußerungen anderer können multimodal markiert werden, z. B. durch *Palm Up* Gesten im Verbund mit weiteren Ressourcen wie dem seitlichen Abknicken des Kopfes, geschürzten Lippen bzw. dem *Mouth Shrug* (vgl. Morris 1994: 165; Debras 2017; vgl. auch Streeck 2009b: 189ff.; vgl. zum Aspekt der Distanzierungen durch *Shrugs* auch Debras und Cienki 2012).

Insgesamt tragen gestische Ressourcen also zur *Qualifizierung* von Äußerungen, zur Markierung der *affektiven* Haltung und Involviertheit der Sprecher*innen gegenüber ihren eigenen oder anderen Äußerungen bei. Darüber hinaus werden sie genutzt, um *epistemische* Un-/Sicherheit eines*einer Sprecher*in anzuzeigen.

Äußerungsstrukturierende Funktion

Wie in der tabellarischen Übersicht 4 illustriert, können gestische Ressourcen zur Strukturierung einer Äußerung beitragen.

	Beschreibung	Funktion
äußerungsstrukturierende Funktion	Strukturierung eines Redebeitrags	• Hervorheben von Äußerungselementen • Markieren zusammengehöriger bzw. fortgesetzter Äußerungselemente • Aufzählen • Kontrastieren bzw. Gegenüberstellen • Lenken der Aufmerksamkeit auf nachfolgende Inhalte

Tabelle 4: Übersicht über strukturierende Funktionen gestischer Ressourcen

Auf Ebene der Äußerungsstrukturierung, d. h. der turninternen Strukturierung können Gesten pragmatische Funktionen übernehmen, indem sie beispielsweise einzelne Äußerungselemente hervorheben, wie u. a. McNeill (1992; vgl. Kap. 3.2.1) für *Beat* Gesten herausstellt. Durch Wiederholungen von Gesten(-elementen) werden außerdem zusammengehörige Elemente und Fortsetzungen von *Turns* und Argumenten identifizierbar (vgl. *Cohesives*, McNeill 1992; vgl. *Palm Up Open Hands* mit rotierenden Bewegungen Müller 2004; vgl. *Cyclic Gestures* Ladewig 2014a). Zählgesten sowie Gestenwiederholungen können auch Abfolgen sowie listenförmige Aufzählungen sichtbar anzeigen (vgl. u. a. Streeck 2009b; vgl. auch Müller 2004; Kendon 2004); und das kontrastive Gegenüberstellen von Äußerungsinhalten kann mittels alternierender Gesten für alle Beteiligten angezeigt werden, beispielsweise indem nacheinander ausgeführte Gesten der rechten und der linken Hand eingesetzt werden (vgl. Müller 1998: 255). Sprecher*innen können aber auch Gesten wie die des erhobenen Zeigefingers in Richtung ihres Gegenübers als „attention getting device“ verwenden (Müller 2004: 242), um die Aufmerksamkeit des Gegenübers auf den dann folgenden Inhalt zu lenken (vgl. auch Inbar 2022).

Zusammenfassend können Gesten also *turnintern*, z. B. durch das Hervorheben von Äußerungselementen sowie durch die Anzeige einer Liste, zur Äußerungsstrukturierung beitragen.

Interaktionsorganisatorische Funktion

Gesten übernehmen auch gesprächs- und rederechtsorganisatorische Funktionen, wie im Anschluss an die zusammenfassende Übersicht (vgl. Tab. 5) näher erläutert wird.

	Beschreibung	Funktion
interaktionsorganisatorische Funktion	Organisation des Gesprächs bzw. des Rederechts	• (frühzeitige) Anzeige einer Turnübernahme • Beanspruchen des Rederechts • Adressieren der Äußerung an spez. Gegenüber • Elizitieren von Rückmeldungen

Tabelle 5: Übersicht über interaktionsorganisatorisch Funktionen von gestischen Ressourcen

Gesten können zur Gesprächs- und Rederechtsorganisation beitragen, wenn sie z. B. als Vorlaufelemente die Turnübernahme eines*einer potenziellen nächsten Sprecher*in bereits ankündigen (vgl. u. a. Streeck und Hartge 1992; Mondada 2007; Streeck 2009a) oder aber über eine Unterbrechung hinweg, z. B. in Form von deiktischen Gesten, gehalten werden und die Sprecher*innen somit anzeigen, dass sie das Rederecht trotz der Unterbrechung durch andere Beteiligte weiterhin beanspruchen (vgl. Mondada 2007). Auch das Unterbrechen von anderen Interagierenden kann durch Gesten markiert werden, wie beispielsweise Kendon (2004: 252) für eine Geste nachweist, bei der die Handfläche auf das Gegenüber hin orientiert ist, um den*die Sprecher*in an der weiteren Entfaltung des *Turns* zu hindern. Dringen Gesten nun beispielsweise in überlappenden Sequenzen in den Gestenraum des Gegenübers vor, können Sprecher*innen dadurch ihren Anspruch auf das Rederecht auch verstärkt durchsetzen (vgl. u. a. Sweetser und Sizemore 2008). Neben der Dauer der eingesetzten Gesten, mit denen das eigene Rederecht sichtbar beansprucht wird, können Interagierende durch ihre körperliche Orientierung sowie die Ausrichtung der Gesten auch anderen Interagierenden das Rederecht zuweisen (vgl. Bavelas et al. 1992, 1995). Wird beispielsweise eine *Palm Up* Geste, mit der ein Gesprächsgegenstand präsentiert worden ist, über das Turnende hinweg gehalten und das Gegenüber z. B. durch eine Blickzuwendung adressiert, kann hierdurch eine Rückmeldung von anderen Beteiligten der Interaktion elizitiert werden (vgl. Streeck 2007). Dabei wird erwartbar, dass das Gegenüber sich auf den weiterhin präsent gehaltenen Gesprächsgegenstand bezieht. Erst, wenn diese Rückmeldung eingegangen ist, kann das Gespräch fortgesetzt werden (vgl. auch Kendon 2004: 273).

Die hier vorgestellten Gesten können also zur Interaktionsorganisation beitragen, indem sie beispielsweise das Bewerben um das Rederecht sichtbar anzeigen und auch auf die Organisation des Sprecher*innenwechsels wirken können.

Im Anschluss an die in diesem Kapitel dargelegten pragmatischen Funktionen gestischer Ressourcen werden nun sogenannte Gestenfamilien vorgestellt, in denen die ‚Familienmitglieder‘ eine gewisse Beständigkeit in ihrer Gestalt sowie ihrer Bedeutung aufweisen.

3.2.2.2 *Gestenfamilien*

Gesten, die jeweils einen semantischen Kern sowie „certain kinesic features“ (Kendon 2004: 281) teilen, werden in sogenannten *Gestenfamilien* zusammengefasst. Kendon (2004), Müller (2004), Bressemer und Müller (2014b) sowie Ladewig (2014a) identifizieren jeweils pragmatische Gestenfamilien, die sowohl auf performativer, modalisierender, äusserungsstrukturierender als auch interaktionsorganisatorischer Ebene operieren. Die in der folgenden Übersicht (vgl. Tab. 6) angeführten Gestenfamilien konnten von den Forscher*innen bislang rekonstruiert werden. Eine Erläuterung der einzelnen Familien erfolgt im Anschluss an Tabelle 6.

	Familie
Kendon (2004)	• <i>G-family (Grappolo Gesture)</i>
	• <i>R-family (Ring Gesture)</i>
	• <i>Open Hand Prone family</i>
	• <i>Open Hand Supine family</i>
Müller (2004)	• <i>Palm Up Open Hand family</i>
Bressem & Müller (2014b)	• <i>Family of Away Gestures</i>
Ladewig (2014a)	• <i>Cyclic Gesture</i>

Tabelle 6: Übersicht zu gängigen Gestenfamilien

Kendon (2004) identifiziert insgesamt vier Gestenfamilien, die sich in der Ausprägung der Handform sowie ihrer Orientierung unterscheiden, innerhalb der Familie aber einen semantischen Kern teilen: die *G-family (Grappolo Gesture)*, die *R-family (Ring Gesture)*, *Open Hand Prone family* sowie die *Open Hand Supine family*. Die Gesten der *G-family* teilen die Form des sogenannten *Fingerbunch*, bei dem alle Finger einer Hand mit dem Daumen zusammengeführt werden. Die Gesten sind eng mit dem thematischen Kern einer Äußerung verknüpft und können diesen hervorheben. Dabei ist ihre interaktive Funktion u. a. davon abhängig, ob die Finger im Gestenhöhepunkt zusammengeführt (*Closure to Grappolo*) oder geöffnet (*Grappolo to Open*) werden. Während die *Closure to Grappolo* Geste in Kontexten verwendet wird, in denen ein neuer Gesprächsgegenstand eingeführt wird, nutzen Sprecher*innen die *Grappolo to Open* Geste v. a. dann, wenn sie einen (abschließenden) Kommentar einbringen (vgl. Kendon 2004: 234). In beiden Ausprägungen wird der Kern einer Äußerung hervorgehoben. Die Gesten der *R-family* teilen in ihrer Form, dass Daumen und Zeigefinger zu einem Ring zusammengelegt werden. Semantische Ähnlichkeit zeigt sich bei ihnen darin, dass sie eine Präzisierung oder Spezifizierung in einer Äußerung markieren. Die Gesten dieser beiden Gestenfamilien operieren auf der Ebene der Äußerungsstrukturierung (vgl. Kendon 2004: 281). Die anderen beiden großen Gestenfamilien bewegen sich hingegen sowohl auf *performativer*, *äußerungsstrukturierender*, *modalisierender* und *interaktionsorganisatorischer* Funktionsebene. Gesten der *Open Hand Prone family* teilen äußerlich eine von den Sprecher*innen abgewandte (entweder nach unten oder in Richtung der Rezipient*innen) ausgerichtete Handfläche, die die Bedeutung von „‘halting, interrupting or indicating the interruption of a line of action““ (Kendon 2004: 281; vgl. zsf. Müller 2017) tragen. In der Gestenfamilie der *Open Hand Supine family* weist die Handfläche hingegen eine nach oben gerichtete Orientierung auf. Je nach Aus- bzw. Bewegungsrichtung tragen die Gesten dieser Familie dazu bei, äußerungsstrukturierend Redebeiträge einzuleiten oder Erklärungen, Kommentare sowie Klarstellungen oder auch Zusammenfassungen präsentierend anzubieten (*Palm Presentation*), oder, wenn die Handfläche auf den*die Rezipient*in ausgerichtet ist (*Palm Addressed*), eine Referenz zwischen dem Äußerungsinhalt und der adressierten Person herzustellen. Wird diese Geste am Ende eines *Turns* gehalten, kann sie in interaktiver Funktion ebenfalls anzeigen, dass eine Reaktion oder Rückmeldung des Gegenübers erwartet wird (vgl. Kendon 2004: 264ff.). *Palm Up* Gesten, die in ihrer Bewegung nach außen geführt werden (*Palm Laterals*), operieren auf modalisierender Ebene, da sie einen Äußerungsinhalt als evident oder offensichtlich markieren.

An Kendons (2004) Untersuchung anknüpfend entsteht Müllers (2004) Studie zur Gestenfamilie der *Palm Up Open Hand Gestures*, die Kendons (2004) *Open Hand Supine family* zuzuordnen ist, und in der sie weitere Funktionen der *Palm Up Open Hand* Gesten rekonstruieren kann. Zu Gesten der *Palm Up Open Hand family* zählen u. a. beidhändig geäußerte *Palm Up* Gesten, mit denen Interagierende ein diskursives Objekt anbieten und eine erwartete Zustimmung intensivieren können. Durch wiederholend nach unten geäußerte Bewegungen der *Palm Up Open Hand* kann eine listenförmige Struktur einer Äußerung angezeigt werden. Wird die *Palm Up Open Hand* Geste mit rotierenden Bewegungen eingebracht, werden Gesprächsinhalte äußerungsstrukturierend kontinuierlich miteinander verbunden. *Palm*

Up Open Hand Gesten können auch alternierend geäußert werden und somit einen Kontrast anzeigen oder aber mit einer weiten, nach außen öffnenden Bewegung hervorgebracht werden, die einerseits ein diskursives Objekt performativ präsentiert und zugleich, in semantisch referenzieller Funktion, eine „[w]ide range of entities“ markieren (Kendon 2004: 252).³⁶

Zwei weitere Gestenfamilien können von Bressem und Müller (2014b) sowie Ladewig (2014a) rekonstruiert werden. Die Gesten dieser Familien sind allerdings nicht durch ihren semantischen Kern und eine kinetische *Form* miteinander verbunden, sondern v. a. durch ihre *Bewegung* bzw. *Bewegungsrichtung*: Die sogenannten *Family of Away Gestures*, die Bressem und Müller (2014b) identifizieren, haben gemein, dass sich die Hand bei allen Familienmitgliedern vom Körper weg bewegt und die Sprecher*innen mit ihnen Negation, Zurückweisung und negative Bewertungen ausdrücken.³⁷ Die Autor*innen zeichnen in ihren Analysen nach, dass diese Bedeutung möglicherweise in der eigentlichen Handlung (vgl. auch Streeck 2009b; Ladewig und Hotze 2023) begründet liegen könnte, ungewollte Objekte oder Handlungen vom eigenen Körper fernzuhalten (vgl. Bressem und Müller 2014b: 1592). Eine weitere Gestenfamilie bildet die *Cyclic Gesture* – die Ladewig (2014a, 2014b) zwar nicht selbst explizit als Gestenfamilie, aber als wiederkehrende (*Recurrent Gesture*) und somit relativ stabile Verbindung von Geste und Bedeutung beschreibt. Auch die Gesten dieser Familie weisen gleiche kinetische Merkmale in ihrer Bewegung auf: Es handelt sich um Gesten, bei der die Hand vom Handgelenk aus in eine mehrmalige, kontinuierliche Rotationsbewegung überführt wird (vgl. Ladewig 2014a). Sie teilen die semantische Bedeutung von Kontinuität (vgl. Ladewig 2014a: 1605; vgl. auch vgl. Müller 2004: 264ff.).

Die in diesem Teilkapitel beschriebenen Mitglieder der Gestenfamilien mit pragmatischer Funktion werden u. a. auch aufgrund des Grades der Konventionalisierung, den sie über unterschiedliche Kontexte hinweg tragen, als wiederkehrende Gesten beschrieben (vgl. u. a. Müller 2017). Je nach kontextueller Verwendung kann es zwar zu lokalen Bedeutungsvariationen kommen, insgesamt zeigen sie aber eine relativ stabile Verbindung von Form bzw. Bewegung(-srichtung) und Bedeutung (vgl. Müller 2017: 277). Müller (2004: 253) führt diese Besonderheit von Gesten mit pragmatischer Funktion unter Bezugnahme auf Kendon (1995b, zit. n. Müller 2004: 253) darauf zurück, „whereas the things to talk about seem infinite, the ways in which we talk about them seem to be limited“.

Zusammenfassend konnte das vorangegangene Kapitel einen Einblick in die vielfältigen pragmatischen Funktionen von Gesten (vgl. auch Kap. 3.2.2.1) geben, in denen u. a. auch relativ feste Form/Bewegungs-Bedeutungs-Verbindungen als sogenannte Gestenfamilien rekonstruiert werden konnten (vgl. auch Kap. 3.2.2.2). Dieses Spektrum an pragmatischen Funktionen soll nun als Hintergrundfolie für das nachfolgende Kapitel dienen, in dem spezifisch auf Funktionen pragmatischer Gesten im Kontext des Argumentierens fokussiert wird.

3.3 Funktionen pragmatischer Gesten für das Argumentieren

In den vorangegangenen Kapiteln wurde gezeigt, dass Gesten gemeinsam mit weiteren Ressourcen komplexe Beziehungsgefüge eingehen, die an dem Aufbau sozialer Interaktion beteiligt sind (vgl. v. a. Kap. 3.1.4), und dass dabei Gesten mit pragmatischer Funktion Einfluss auf den weiteren Verlauf der Interaktion nehmen können, da sie Sprechhandlungen und sprecher*innenseitige Haltungen anzeigen sowie

³⁶ Müller (2004: 254) liefert selbst keine Hinweise darauf, ob es sich um eine untergeordnete Gruppe der von Kendon beschriebenen Gestenfamilie handelt und in welchem Verhältnis die Gestenfamilie der *Palm Up Open Hand* Geste zu Kendons *Open Hand* Familie steht.

³⁷ Auch die von Payrató und Teßendorf (2014) sowie Teßendorf (2014) beschriebenen *Brushing Aside* Gesten zählen zu der Gestenfamilie der *Away* Gesten.

zur Äußerungsstrukturierung und Interaktionsorganisation beitragen (vgl. v. a. Kap. 3.2.2). Inwiefern Gesten mit pragmatischer Funktion gerade für *argumentativ gerahmte Interaktionen* relevant werden können, wird im nachfolgenden Kapitel erläutert. Allerdings zeigt sich in der Zusammenschau der Befunde, dass die Studien nur selten selbst einen expliziten Bezug zur Gattung des Argumentierens herstellen – auch wenn aus den Daten eine argumentative Rahmung der abgebildeten Interaktion hervorgeht. Dieses Desiderat berücksichtigend wirft der abgebildete Forschungsstand relevante Schlaglichter auf die pragmatischen Funktionen von Gesten in argumentativen Kontexten. Die Befunde werden in der anschließenden empirischen Untersuchung (vgl. Kap. 8) kontinuierlich einbezogen.

Die im Folgenden berücksichtigten Studien, die im weitesten Sinne pragmatische Funktionen von Gesten in argumentativen Interaktionen rekonstruieren, nehmen überschneidend ähnliche Phänomenbereiche in den Blick, deren Schwerpunkte wie folgt systematisiert werden können: die *gemeinsame Arbeit an der Argumentation und Kollaboration*; das *Sichtbarmachen der Sprechhandlung*; das *Anzeigen der epistemischen Verantwortung* sowie der *epistemischen Haltung* der Interagierenden und die *Strukturierung des Arguments bzw. der Argumente*. Entlang dieser von den Studien fokussierten Phänomenbereiche werden im Folgenden die pragmatischen Funktionen von Gesten für das mündliche Argumentieren dargelegt.

Gemeinsame Arbeit an der Argumentation und Kollaboration

Im Rahmen argumentativer Sequenzen, in denen Sprecher*innen divergierende Positionen einbringen und „the collaborative continuation of talk is potentially at risk“ (Schönfelder und Heller 2019: 86), können multimodale Ressourcen von Interagierenden genutzt werden, um eine andauernde *gemeinsame Arbeit an der Argumentation* i. S. eines *participation framework* (vgl. Goodwin und Goodwin 2004) zum Ausdruck zu bringen. Dabei meint das Konzept des *participation framework*, dass sich Personen ihr Mitwirken an der laufenden gemeinsamen Interaktion sowie ihr Verständnis der gemeinsamen Aktivität gegenseitig anzeigen. So wurde beispielsweise in Peer-Argumentationen beobachtet, dass die Sprecher*innen reziproke *Palm Up* Gesten nutzen und dadurch das Geben und Nehmen von Argumenten auch visuell sichtbar machen (vgl. Schönfelder und Heller 2019; vgl. hierzu auch unten, *Sprechhandlungen*). Gerade in mündlichen Argumentationen, in denen das Verhandeln inhaltlicher Nicht-Übereinstimmung den Fortgang der Interaktion gefährden kann, erweist sich das Signalisieren einer (fortgesetzten) gemeinsamen Bearbeitung eines aufgetretenen Dissenses oder Problems als relevant. Schönfelder und Heller (2019) rekonstruieren für Peer-Interaktionen, dass die Sprecher*innen wechselseitig Bezug aufeinander nehmen und dabei die *Palm Up* Gesten der vorangegangenen Äußerung wiederholen – auch in nicht-kooperativen Sequenzen. Dadurch markieren die Interagierenden nicht nur die (fortgesetzte) gemeinsame Arbeit an der Argumentation, sondern sie schaffen auch eine kohäsive Verknüpfung zwischen den *Turns* unterschiedlicher Sprecher*innen. In argumentativen Kontexten können interpersonelle gestische Bezugnahmen außerdem zu einer Abschwächung des Grades der Konfrontation beitragen: So können Proponent*innen durch Zeigegesten (vgl. Kap. 3.2.1) auf Opponent*innen Bezug nehmen, ohne eine explizit-sprachliche Referenz zur ursprünglichen Position bzw. zu der Person herstellen zu müssen, die diese geäußert hat (vgl. Jacquin 2015, 2018). Doch Sprecher*innen können gestische Ressourcen nicht nur zur Kontextualisierung einer *gemeinsamen Arbeit* an der Argumentationen einsetzen, sondern durch diese auch anzeigen, dass sie *inhaltlich* miteinander i. S. von Affiliationen (vgl. Stivers 2008) übereinstimmen und in ihrer Bearbeitung des kommunikativen Problems miteinander *kollaborieren*. So hat Schönfelder (2024) anhand von Peer-Argumentationen rekonstruiert, dass heranwachsende Sprecher*innen in Ko-Konstruktionen gemeinsamer Begründungen ganze multimodale Gestalten vorangegangener Sprecher*innen wiederholen. D. h. die Sprecher*innen markieren ihre Kollaboration mit anderen Beteiligten nicht allein durch sprachlich explizierte Formen, sondern auch durch

das Wiederholen gesamter multimodaler Gestalten eines*einer vorangegangenen Sprecher*in (vgl. Schönfelder 2024; vgl. auch Williams 2022). Diese unterschiedlichen Varianten der *gemeinsamen Arbeit* an einer Argumentation, zu denen (reziproke) Bezugnahmen der Sprecher*innen (auch in weniger kooperativen Kontexten) sowie Kollaborationen in ko-konstruierten Begründungen zählen, liegen auf der Ebene interaktionsorganisatorischer Funktionen (vgl. Kap. 3.2.2.1).

Sichtbarmachen der Sprechhandlung

In argumentativen Kontexten tragen Gesten aber auch dazu bei, jeweilige *Sprechhandlungen* zu markieren. So arbeitet Müller (2004) heraus, dass Sprecher*innen Positionen in die Interaktion einbringen und als abstrakte diskursive Objekte auf der nach oben geöffneten Handfläche den Gesprächspartner*innen präsentieren und somit zur *Zustimmung anbieten* oder zur *Weiterverhandlung einladen* können (vgl. Müller 2004: 243). In mündlichen Argumentationen können auch Vorschläge (vgl. Heller 2018a), einzelne Argumente (vgl. Schönfelder und Heller 2019) sowie konkludierende Kommentare (vgl. Heller 2021b) auf der geöffneten Handfläche präsentiert werden. Neben der präsentierenden bzw. anbietenden Funktion, die *Palm Up* Gesten übernehmen können, kann sich diese Funktion über die verstreichende Temporalität hinweg verändern: Wenn Sprecher*innen die Geste über das Ende ihres *Turns* hinweg halten, entsteht für die Rezipient*innen die starke Verpflichtung einer Rückmeldung, z. B. in Form einer Zustimmung.³⁸ Clift (2020) beschreibt diese Funktion für beidhändig, teilweise mit seitlicher Bewegungsrichtung geäußerte und turnfinal gehaltene *Palm Up* Gesten, als eine Art Klimax, mit der die Gesten als „upgrade to a verbal pursuit“ (Clift 2020: 9) fungieren: Ist eine vorangegangene Aufforderung zu einer Rückmeldung fehlgeschlagen, d. h. antwortet das Gegenüber nicht, können *Palm Up* Gesten in einer Erneuerung der Aufforderung den Stillstand der interaktiven Progression, der durch das Gegenüber verursacht wird, indizieren und die Dringlichkeit einer Rückmeldung verstärken. Die Geste übernimmt hier eine performative Funktion, da sie das Einfordern einer Rückmeldung deutlich anzeigt. *Palm Up* Gesten können aber auch im Kontext des Hinterfragens bzw. des Etablierens von Zugzwängen zum Begründen auftreten, die Clift (2020: 13) als „The Palm up in challenges“ beschreibt. Hier markiert die *Palm Up* Geste gemeinsam mit dem Hinterfragen bzw. Explizieren eines Zugzwangs, dass die Interaktion bis zur Beantwortung der Frage (kurzzeitig) angehalten wird. Erst wenn die Frage beantwortet ist, kann die Interaktion fortgesetzt werden (vgl. Clift 2020: 15). Abhängig vom sequenziellen Kontext können *Palm Up* Gesten in argumentativen Sequenzen also sowohl das Anhalten der Interaktion bei ausgebliebener Rückmeldung des Gegenübers sowie das Aussetzen der Progression bis zur Beantwortung einer eingebrachten Frage visualisieren (vgl. Clift 2020: 21). Die Sprecher*innen zeigen durch das turnfinale Halten der Geste somit für alle Interagierenden sichtbar an, dass das diskursive Objekt einen „transitional status“ (Kendon 2004: 274) erlangt, sodass die Interaktion nicht fortgesetzt werden kann, ehe dies nicht verhandelt worden ist. Sprechhandlungen können außerdem auch durch stark konventionalisierte Gesten ausgedrückt werden, wie Kendon (1995a: 262) für die Geste der *Mani Giunte*, die betenden Hände, beschreibt: Diese werden beispielsweise in argumentativen Kontexten eingesetzt, um die Bitte gegenüber anderen Beteiligten zu markieren, einen begründeten Einwand oder ein eingebrachtes Argument zu akzeptieren. Insgesamt zeigt sich in den Studien, dass im Kontext mündlicher Argumentationen körperlich-sichtbare Ressourcen in performativer Funktion Sprechhandlungen sichtbar machen können, wie das Präsentieren von Gesprächsgegenständen, das Markieren einer Bitte sowie (intensivierte) Aufforderungen zu einer Rückmeldung.

³⁸ Vgl. hierzu auch, allerdings nicht in ausgewiesenen argumentativen Interaktionen, Streeck (2007).

Anzeigen der epistemischen Verantwortung

Gestische Ressourcen können in argumentativen Sequenzen auch dazu beitragen, eine (gemeinsame) epistemische Verantwortung gegenüber dem verhandelten Gesprächsgegenstand einzunehmen. Heller (2018a) untersucht, wie Sprecher*innen in Peer-Argumentationen *Palm Up* Gesten gemeinsam mit fazi- alen Gesten nutzen, um eine gemeinsame *epistemische Verantwortung* zu etablieren: Die Beteiligten adressieren in diesem Kontext weitere Gruppenmitglieder durch *Palm Up* Gesten sowie Blickzuwen- dungen am Ende des *Turns*, um die Rezipient*innen zu einem gemeinsamen Nachdenken über und das Verhandeln des präsentierten diskursiven Objektes einzuladen und sich somit auch gegenseitig anzuzei- gen, dass das Lösen eines Problems in einen *gemeinsamen* epistemischen Verantwortungsbereich der Interagierenden fällt (vgl. Heller 2018a: 272). Selbst dann, wenn die Sprecher*innen ihrerseits eine starke epistemische Haltung markiert haben, können die eingesetzten Ressourcen „serve[] to calibrate the stance display and offer the reason as an inspectable object“ (Heller 2018a: 275). Körperlich-sicht- bare Ressourcen fungieren also in argumentativen Sequenzen als Vermittlungsinstanz, die zur Beteili- gung anderer an der Argumentation auch dann einladen, wenn eine Begründung oder Entscheidung auf sprachlicher Ebene bereits abgeschlossen erscheint. Sprecher*innen können ihr Gegenüber aber nicht nur zu einer Beteiligung an der Argumentation *einladen*, sie können durch das turnfinale Halten der Geste z. B. in opponierenden, hinterfragenden Handlungen die *Antwortverpflichtung verstärken* (vgl. oben, *Sprechhandlungen*; vgl. Clift 2020) – die epistemische Verantwortung, sich zu dieser Opposition zu verhalten, wird somit von dem*der Sprecher*in auf das Gegenüber übertragen. Das Übertragen epis- temischer Verantwortung auf das Gegenüber zeigt sich auch in der Studie der von Kendon beschriebe- nen *Mano Giunte* (vgl. Kendon 1995a), der bittenden Hände: Sprecher*innen können mit dieser Geste nicht nur die Sprechhandlung einer Bitte markieren, sondern auch die Entscheidung über die Akzeptanz einer Bitte an das Gegenüber übergeben. Dadurch wird die epistemische Verantwortung sowie die epis- temische Autorität (vgl. Stivers et al. 2011: 13f.) über die Entscheidung, z. B. einen Einwand zu akzep- tieren (vgl. Kendon 1995a: 259f.), für alle Beteiligten sichtbar auf das Gegenüber übertragen. Es wird hier nicht nur die Bitte selbst als Sprechhandlung (vgl. oben, *Sichtbarmachen der Sprechhandlung*) an- gezeigt, sondern auch das etablierte epistemische Gefälle zwischen bittender und entscheidender Person, ohne dass dies sprachlich expliziert werden muss.

Anzeigen der epistemischen Haltung

Gesten mit pragmatischer Funktion können ebenfalls Aufschluss über die *epistemische Haltung*, d. h. eine sicher wissende oder unsichere Haltung (vgl. Heritage 2012b) der Sprecher*innen geben und zur Modalisierung (vgl. Kap. 3.2.2.1) einer Äußerung beitragen. Im Kontext mündlicher Argumentationen können Sprecher*innen „try[] to persuade another that something is correct, or the speaker is trying to alter another’s view of how some event should be interpreted“ (Kendon 1995a: 277). Die epistemische Haltung eines*einer Sprecher*in kann dann dazu beitragen, einen Gesprächsgegenstand als epistemisch abgesichert und/oder unverhandelbar anzuzeigen und somit eine Zustimmung oder Akzeptanz des Ge- genübers forcieren (vgl. u. a. Heritage 2013: 561f.). Eine sichere epistemische Haltung kann u. a. durch das Markieren von Evidentialität bzw. Offensichtlichkeit ausgedrückt werden: Indem Sprecher*innen *Palm Up* Gesten mit lateraler Bewegung äußern, können sie u. a. prospektiv die Antwort auf rhetorische Fragen (vgl. Müller 2004) oder Begründungen (vgl. Heller 2018a) als offensichtlich bzw. auf der Hand liegend markieren. In der Studie von Marrese et al. (2021) zeigt sich ferner, dass die *Palm Up* Geste mit seitlicher Bewegung auch in Kontexten eingesetzt werden kann, in denen sich die Sprecher*innen in einer argumentativen ‚Sackgasse‘ befinden: Indem die eigene Position durch den Einsatz der Geste als evident markiert wird, können Sprecher*innen in argumentativen Sequenzen „enacting having ‚won‘ this point of the argument“, denn „‘nothing further can be said‘“ (Marrese et al. 2021: 5) und somit die

argumentative Sequenz zugleich zu einem Abschluss bringen. Dabei können *Palm Up* Gesten mit lateraler Bewegung in Kontexten diverser pragmatischer Mittel wie Behauptungen, Fragen oder Imperativen sowie ohne sprachliche Ressourcen auftreten, wie die Studie von Marrese et al. (2021) darlegt, und sogar über mehrere *Turns* hinweg als kontinuierliche Markierung von Evidenz aufrechterhalten werden.

Strukturierung des Arguments bzw. der Argumente

Sprecher*innen können nicht zuletzt innerhalb ihrer eigenen Äußerungen Gesten zur Äußerungs- bzw. *Argumentstrukturierung* verwenden. So zeigt beispielsweise Müller (2004), dass die Iteration von *Palm Up* Gesten, die wiederholt nach unten bewegt werden, einzelne Elemente einer Äußerung als zusammengehörig markiert und sie listenförmig verbindet (vgl. Müller 2004: 245f.; vgl. auch *Cohesives* McNeill 1992; Kap. 3.2.1). Diese Form der Kohäsionsmarkierung, die hier deutlich wird, beschreibt auch Jacquin (2017: 252ff.) für wiederholte „grasping gestures“, bei denen der Sprecher*innen die Fingerspitzen einer Hand immer wieder während seiner Äußerung auf den Tisch aufsetzt, oder für die von Kendon (1995a) beschriebene *Grappolo* Geste, die von Sprecher*innen über ihre gesamte Äußerung hinweg *gehalten* werden kann und die somit die einzelnen Äußerungselemente als zusammengehörig markiert. Gesten mit rotierender Bewegung, wie Müller (2004: 246f.) in Bezug auf *Palm Up* Gesten rekonstruiert, können außerdem die kontinuierliche Fortführung eines Arguments markieren. Die Strukturierung *einzelner* Argumente innerhalb eines *Turns* kann ferner durch alternierende *Palm Up* Gesten erfolgen, wenn Sprecher*innen beispielsweise mit Benennen des ersten Arguments erst die eine Hand und mit Benennung eines weiteren Arguments dann die andere Hand zu einer *Palm Up* Geste formen (vgl. Müller 2004: 246ff.). Heller (2021b: 8) beobachtet die alternierende *Palm Up* Geste in Peer-Argumentationen ebenfalls in strukturierender Funktion, allerdings im Kontext des Einbringens und Aufgebens einer ursprünglichen Position. Ferner können Sprecher*innen in Argumentationen *Palm Up* Gesten auch nutzen, um eine Konklusion als Resultat längerer Begründungssequenzen zu markieren (vgl. Heller 2021b). Auch Gesten wie der *Finger Bunch* bzw. die *Closure to Grappolo Geste* (vgl. Kap. 3.2.2.2) werden im Kontext argumentativer Sequenzen äußerungsstrukturierend eingesetzt, um beispielsweise das Thema, das der*die Sprecher*in inhaltlich verhandelt, sowie dessen Relevanz hervor- bzw. von der restlichen Äußerung abzuheben (vgl. Kendon 1995a: 264, 2004: 130).

Die Übersicht über den bisherigen Forschungsstand zu Gesten mit pragmatischen Funktionen zeigt wichtige Befunde zu der Funktionsvielfalt von Gesten, die sie in Abhängigkeit vom interaktiven, argumentativen Kontext und gemeinsam mit anderen Ressourcen übernehmen können. Diese können u. a. in der *gemeinsamen Arbeit an einer Argumentation* reziproke Bezugnahmen der Sprecher*innen markieren (vgl. Schönfelder und Heller 2019) sowie *Kollaborationen* in der gemeinsamen Bearbeitung eines Arguments visuell sichtbar machen (vgl. Schönfelder 2024). Durch interpersonelle deiktische Gesten können die Interagierenden auch den Grad der Konfrontation abschwächen (vgl. Jacquin 2015, 2018). Auf der Ebene der *Sprechhandlungen* können Gesten in ihren Ressourcenverbünden diskursive Objekte wie Argumente, Vorschläge oder konkludierende Kommentare präsentieren (vgl. Müller 2004; Heller 2018a, 2021b; Schönfelder und Heller 2019), eine Verpflichtung zur Rückmeldung markieren (vgl. Clift 2020) sowie den Halt der Progression einer Interaktion anzeigen (vgl. Clift 2020). Das Präsentieren von Begründungen bzw. Argumenten kann dann, in *gemeinsamer epistemischer Verantwortung* (Heller 2018), zur Zustimmung oder weiteren Aushandlung einladen (vgl. Müller 2004). Ferner können die gestischen Ressourcen Aufschluss über die *epistemische Haltung* eines*einer Sprecher*in geben, indem beispielsweise durch *Palm Up* Gesten mit lateraler Bewegung Evidenz markiert wird (vgl. Müller 2004; Marrese et al. 2021; Heller 2021b) und durch diese der Abschluss argumentativer Sequenzen eingeleitet werden kann (vgl. Marrese et al. 2021). Nicht zuletzt tragen gestische Ressourcen zur *Argumentstruk-*

turierung bei, indem sie Aufzählungen bzw. Listen, Turnfortsetzungen sowie das Einbringen und Aufgeben einer Position markieren (vgl. Kendon 1995a; Müller 2004; Jacquin 2017; Heller 2021b) sowie einzelne, inhaltlich besonders relevant Äußerungssegmente aus der umgebenden Äußerung hervorheben (vgl. Kendon 1995a).

Die Befunde des aktuellen Forschungsstandes weisen darauf hin, dass im Rahmen argumentativer Interaktionen gestische Ressourcen eine Fülle pragmatischer Funktionen übernehmen können.³⁹ Gerade in mündlichen Argumentationen, in denen „talk is potentially at risk“ (Schönfelder und Heller 2019: 86), können sie eine wichtige Ressource bilden, die für alle Beteiligten sichtbare Hinweise über die Haltung der Sprecher*innen oder die Sprechhandlungen der Äußerungen enthält, ohne dass diese notwendigerweise auch sprachlich markiert werden müssen. Trotz der breiten Forschungsbefunde zu pragmatischen Funktionen von Gesten im Allgemeinen (vgl. Kap. 3.2.2) zeigen bislang vergleichsweise wenige Studien (vgl. Kap. 3.3) eine Berücksichtigung der diskursiven Gattung bzw. Praktik des Argumentierens (vgl. Kap. 2.2). Eine *systematische* Untersuchung von Gesten mit pragmatischen Funktionen in mündlichen Argumentationen steht m. W. noch gänzlich aus, sodass bislang weitestgehend unbeleuchtet ist, inwiefern Gesten mit pragmatischer Funktion in größeren Verbänden multimodaler Ressourcen an der Realisierung *argumentativer Handlungen* beteiligt sind. Diese Frage adressiert die vorliegende Arbeit.

3.4 Zusammenfassung

Anhand des vorliegenden Kapitels 3 wurde dargelegt, dass für den Aufbau einer gemeinsamen Interaktion nicht nur die Sprache als Ressource zentral ist, sondern dass auch körperlich-sichtbare Ressourcen relevante Funktionen übernehmen. Dazu wurden zu Beginn des Kapitels vier Konzeptualisierungen vorgestellt, die das Verhältnis von sprachlichen und körperlich-sichtbaren Ressourcen (vgl. Kapitel 3.1) in unterschiedlicher Weise beleuchten. Aus diesen konnte begründet abgeleitet werden, dass die vorliegende Arbeit die Konzeptualisierung der komplexen multimodalen Gestalten von Mondada (2014b) zugrundelegt, die u. a. die Beschreibung der emergenten Entfaltung aller beteiligten und durch die Interagierenden relevant gemachten Ressourcen ermöglicht. Mit der anschließenden Aufarbeitung weitläufig etablierter Beschreibungs- und Klassifikationssysteme, mit denen Formen und Funktionen von Gesten beschrieben werden können, wurde dann deutlich, dass diese trotz deutlicher Gegensätze auch Ähnlichkeiten aufweisen: Denn die benannten Klassifikationssysteme beruhen letztlich auf der Unterscheidung von semantisch-referenziellen und pragmatischen Funktionen von Gesten (vgl. Kap. 3.2.1). Für die empirischen Untersuchungen der vorliegenden Arbeit werden die Klassifikationssysteme insofern relevant, als dass aus ihnen eine zusammenfassende Übersicht über Gestenbezeichnungen sowie Funktionsbeschreibungen abgeleitet werden konnte (vgl. Tab. 1), die als Grundlage für die Kapitel 8 und 10 dient. Der Blick wurde anschließend auf pragmatische Funktionen verengt (vgl. Kap. 3.2.2) und unter Berücksichtigung des Forschungsinteresses der vorliegenden Arbeit ein Einblick in Gesten mit pragmatischer Funktion innerhalb argumentativ gerahmter Interaktionen gegeben, die bislang zwar untersucht, aber selten in der Interpretation ihrer Funktionen an ihren argumentativen Kontext zurückgebunden worden sind (vgl. Kap. 3.3).

³⁹ Während das Synchronisieren von Gesten und Gestenbewegungen in Argumentationen relevante Informationen über die Kollaboration in Argumentationen enthält, kann das Einfrieren (vgl. zum *freezing* Hannken-Illjes und Bose 2019; Heller 2021; vgl. auch Marrese et al. 2021) von Körperposituren, entgegengesetzt, deutlich einen Dissens markieren.

In dem Forschungsvorhaben, mündliche Argumentationen systematisch auf die beteiligten multimodalen Ressourcen, mit besonderem Fokus auf gestische Ressourcen, zu untersuchen (vgl. Kap. 1), treten nun zwei grundlegende Beobachtungen besonders hervor, die den jeweiligen forschungsmethodischen Zugang betreffen: Während die Untersuchung mündlicher Argumentationen sowie ihrer interaktiven sequenziellen Entfaltung aus Perspektive der Konversationsanalyse vornehmlich auf die *sprachliche* Ressource fokussiert (vgl. Kap. 2), erfolgt die Rekonstruktion der Funktionen gestischer Ressourcen hauptsächlich unter Vernachlässigung eines differenzierten Einbezugs sprachlicher Ressourcen sowie gattungsspezifischer Merkmale (vgl. Kap. 3; vgl. aber zur Berücksichtigung sprachlicher Ressourcen Goodwin 2000; Mondada 2014b). Die vorliegende Studie adressiert diese Forschungslücke und wird die Funktionen multimodaler Ressourcen für das mündliche Argumentieren über den forschungsmethodischen Zugang der multimodalen Konversationsanalyse (vgl. Kap. 3.1.4; vgl. ausführlicher Kap. 7.1) untersuchen und dabei Gattungsaspekte (vgl. Kap. 2) konsequent berücksichtigen. Zugleich werden die in diesem Kapitel 3 vorgestellten Befunde aus der Gesten- bzw. Multimodalitätsforschung systematisch in die Untersuchung eingebunden.

Da die Studie auch erwerbsbezogene Fragestellungen in den Blick nimmt, wird das folgende Kapitel 4 Befunde zweier bislang vornehmlich getrennter Forschungsstränge zusammenführen, indem Erkenntnisse der sprachlich-diskursiv orientierten Diskurserwerbsforschung zum mündlichen Argumentieren sowie erwerbsbezogene Befunde zur Entwicklung multimodaler Ressourcen vorgestellt werden.

4 Multimodale Ressourcen im Diskurserwerb

Im Hinblick auf die zweite empirische Fragestellung der vorliegenden Studie, *welche Veränderungen sich in einer längsschnittlich angelegten Studie zweier Fokuskinder im Alter von 12, 13 und 15 Jahren hinsichtlich der sprachlich-diskursiven Fähigkeiten sowie des Einsatzes multimodaler Ressourcen zeigen* (vgl. Kap. 1), und somit in Vorbereitung auf Kapitel 10 der Arbeit werden innerhalb des folgenden Kapitels zuerst unterschiedliche Modellierungen von Argumentationskompetenz und ihres Erwerbs sowie entsprechende Kontexte vorgestellt, in denen argumentative Fähigkeiten erworben und ausgebaut werden (vgl. Kap. 4.1). Der Annahme folgend, dass nicht nur *sprachlich-diskursive* Fähigkeiten Teil des Diskurserwerbs sind, werden dann erwerbsbezogene Befunde zur Entwicklung und Funktionen *multimodaler Ressourcen* in diskursiven Praktiken vorgestellt (vgl. Kap. 4.2). Das Kapitel schließt mit einer Zusammenfassung und leitet mit einem Ausblick (vgl. Kap. 4.3) in die Forschungsdesiderate und -fragen (vgl. Kap. 5) über, die die vorliegende Arbeit adressiert.

4.1 Erwerb mündlicher Diskurskompetenz

Mündliches Argumentieren gilt gemeinhin als eine komplexe Diskurspraktik, deren Erwerb nicht bereits mit Beendigung der Primarstufe abgeschlossen ist, sondern der sich bis in die Sekundarstufe I vollzieht (vgl. u. a. Krah et al. 2013; Arendt et al. 2015: 4; Heller und Krah 2015; Quasthoff et al. 2021a). Dabei zeichnet die Forschung ein breites Bild der für den Erwerb und die Entwicklung von Argumentationskompetenz relevanten Facetten und (Teil-)Fähigkeiten. In den Fokus der (u. a. aktual- und ontogenetischen) Studien rückt u. a., in welchem Alter (Klein-)Kinder erste Begründungen sprachlich-diskursiv produzieren können, welche moralischen Fähigkeiten relevant für das Argumentieren werden sowie ob in der Begründung eine Ausrichtung auf das Gegenüber erkennbar ist (vgl. u. a. Miller 1980; Miller und Klein 1981; Völzing 1982; Kline 1998; Kuhn und Udell 2003; Goetz 2010; Zadunaisky Ehrlich und Blum-Kulka 2014; Arcidiacono und Bova 2015; Klein 2017; Arendt 2019).

Im nachfolgenden Kapitel 4.1.1 liegt der Fokus auf Studien, die mit einem konversations- bzw. gesprächsanalytischen Zugriff (Teil-)Fähigkeiten von Argumentationskompetenz anhand von Unterrichtsdaten von Schüler*innen am Ende der Primarstufe (vgl. Vogt 2002) sowie im Sekundarstufenalter (Vogt 2002; Grundler 2011; Krelle 2014) rekonstruieren. Allen Untersuchungen ist durch ihre gesprächsanalytische Ausprägung inhärent, dass die Fähigkeiten der Sprecher*innen immer im Kontext einer gemeinsam hervorgebrachten Interaktion einzuordnen sind. Da die Studien wichtige Einblicke liefern, welche *Teilfähigkeiten* für das Argumentieren relevant werden, aber mögliche *Erwerbsmechanismen* noch nicht freilegen, wird anschließend die empirisch fundierte und erprobte Modellierung von Diskurskompetenz nach Quasthoff (2009) vorgestellt. Diese ermöglicht, erwerbsbezogene Veränderungen argumentativer Fähigkeiten bei heranwachsenden Sprecher*innen zu ermitteln und somit entsprechende *Erwerbsmechanismen* aufzudecken. Das Modell von Quasthoff wird daher in der vorliegenden Studie die Grundlage für die Untersuchung erwerbsbezogener Phänomene in Kapitel 10 bilden. An die Erläuterung des Modells anschließend wird dann anhand der Rezeption des aktuellen Forschungsstandes skizziert, welche Entwicklungslinien sich innerhalb des Argumentationserwerbs Heranwachsender zeigen (vgl. Kap. 4.1.2). Zuletzt wird beleuchtet, inwiefern auch unterschiedlichen Erwerbskontexten, wie der Familie sowie der Schule, eine wichtige Rolle für den Diskurserwerb zukommt (vgl. Kap. 4.1.3).

4.1.1 Modellierungen von Argumentationskompetenz: Argumentative (Teil-)Fähigkeiten und ihr Erwerb

Die nachfolgend vorgestellten Modellierungen von Argumentationskompetenz, die von Vogt (2002), Grundler (2011) und Krelle (2014) entwickelt worden sind, basieren allesamt auf Unterrichtsdaten. Aus diesen rekonstruieren die Forscher*innen induktiv *argumentative Teilfähigkeiten* heranwachsender Schüler*innen. Zugrundeliegend sind eine Querschnittsstudie von Vogt (2002) mit Unterrichtsdaten aus den Jahrgangsstufen vier bis 13 sowie die Studien von Grundler (2011) und Krelle (2014), die jeweils auf eine Alterskohorte in den Jahrgangsstufen acht bzw. neun fokussieren. Während Vogt (2002) die im Unterricht entstandenen Beiträge der Schüler*innen hinsichtlich ihrer argumentativen Fähigkeiten untersucht,⁴⁰ nehmen Grundler und Krelle unterschiedliche Diskussionsformate wie vorbereitete Podiumsdiskussionen, unvorbereitete *ad hoc*-Diskussionen (vgl. Grundler 2011) und vorbereitete Fishbowl-Diskussionen (Krelle 2014) in den Blick. Das Spezifikum der letztgenannten Daten liegt darin, dass der bearbeitete Dissens bzw. das Problem nicht aus der Unterrichtssituation heraus entsteht (wie z. B. in den Daten von Vogt), sondern die Schüler*innen anhand eines vorgegebenen Themas entweder eigene Positionen entwickeln oder ihnen feste Positionen zugewiesen werden, die sie anhand von Unterrichtsmaterial vorbereiten (vgl. zur Methode der Podiums- und Fishbowldiskussionen u. a. Morek 2017; Mundwiler et al. 2017; vgl. auch Kap. 4.1.3, Erwerbskontext Schule).

In der Rekonstruktion der Teilfähigkeiten, die die Schüler*innen zur Bearbeitung eines Problems benötigen, gehen die drei Studien methodisch jeweils unterschiedlich vor und entwickeln spezifische Systematisierungen zur Beschreibung der Argumentationskompetenz: Vogt (2002: Kap. 8.1.1) differenziert in seiner Studie die Analyseebenen der *Struktur-* und der *Prozessebene*. Dies ermöglicht, sowohl zu rekonstruieren, wie die Sprecher*innen ihre Gesprächsbeiträge *intern strukturieren*, als auch zu untersuchen, wie sie ihren Beitrag in die laufende Interaktion *einpassen*. Indem Vogt (2002) ermittelt, welche Sprechhandlung der*die Sprecher*in realisiert, d. h. ob z. B. eine Proposition hervorgebracht und möglicherweise durch eine Begründung gestützt wird (vgl. Vogt 2002: 238f.), kann er zum einen die Struktur in ihrer Komplexität freilegen. Anhand der Rekonstruktion der einzelnen Sprechhandlungen wird etwa sichtbar, welche einzelnen Handlungen der*die Sprecher*in miteinander zu einer Äußerung verknüpft. In der Komplexität, die der Beitrag gewinnen kann, wird schließlich die argumentative Kompetenz eines*einer Sprecher*in manifest (vgl. Vogt 2002: 257). Zum anderen können die Beiträge neben der Strukturebene aber auch auf der Prozessebene untersucht werden. Hier bilden die sequenzielle sowie die interaktive Entfaltung der Äußerung die Grundlage der Analyse. Dies ermöglicht dem Forscher auch, Bezugnahmen der Sprecher*innen aufeinander herauszupräparieren und dabei zu eruieren, ob beispielsweise Positionen neu, d. h. initiiierend eingebracht oder der Beitrag des*der vorangegangenen Sprecher*in reagierend bearbeitet wird (vgl. Vogt 2002: 240f.). Die Anforderung, einen Beitrag initiativ

⁴⁰ Vogt (2002: 1) verwendet zwar selbst den Begriff des Diskutierens anstelle des Argumentierens, definiert ihn aber analog zum Verständnis der Definition des Argumentierens der vorliegenden Arbeit (vgl. Kap. 2) als einen gemeinsamen Versuch, „sich über eine Frage, ein Problem oder einen Sachverhalt Klarheit zu verschaffen und gemeinsam eine Antwort, eine Lösungsmöglichkeit oder eine Deutung zu entwickeln“. Er stellt dabei einen Zusammenhang zwischen der Praktik des Diskutierens und dem Argumentieren her, der sich nach Vogt (2002: 82) so gestaltet, dass das „Diskutieren [...] als Praktik charakterisiert [wird], in der die Beteiligten eine für sie wichtige Frage argumentativ bearbeiten“. Ferner heißt es: „Aus Sicht einer deskriptiven, empirisch fundierten Gesprächsforschung schließlich zeigt sich die Diskussion bzw. das Diskutieren als ein komplexer Prozess der interaktiven Bearbeitung des Argumentierens“ (Vogt 2002: 28). M. E. führt dies zu einer definitorischen Ungenauigkeit, da das Argumentieren selbst eine interaktive Bearbeitung eines Problems darstellt – das Argumentieren selbst kann dabei nicht bearbeitet werden, sondern bietet eine diskursive Praktik, mittels derer etwas *Anderes*, nämlich ein Problem oder ein Dissens, bearbeitet werden kann.

einzubringen, scheint dabei höher bzw. voraussetzungsreicher als einen Beitrag ausschließlich reagierend einzubringen. In dieser ersten deutschsprachigen Studie zum mündlichen Argumentieren im Deutschunterricht wird es durch den gesprächsanalytisch geprägten Zugang möglich, zu erfassen, wie Schüler*innen argumentativ einen Gesprächsgegenstand bearbeiten, und aus den Befunden Anforderungen an mündliche Argumentationsfähigkeiten abzuleiten. Die Studie von Grundler (2011) untersucht die Beiträge von Schüler*innen in *ad hoc*-Diskussionen sowie vorbereiteten Podiumsdiskussionen und entwickelt unter Bezugnahme auf das Sprachkompetenzmodell von Coseriu sowie das Organonmodell von Bühler eine Matrix mit insgesamt *zwölf Teilkompetenzen* (Grunder 2011: 58–68). Diese Modellierung erlaubt es der Forscherin, die einzelnen Kompetenzausprägungen in einer hohen Differenziertheit zu rekonstruieren. Die Teilkompetenzen bilden (analytisch getrennt, aber in der Interaktion miteinander verwoben) die Fähigkeiten der Sprecher*innen auf konversationeller, personaler, kognitiver und sozialer Ebene ab. Dadurch bezieht Grundlers Modell neben sprachlich-diskursiven (konversationellen) Fähigkeiten der Schüler*innen auch ihre dargelegten Wissensbestände (kognitive Ausprägung) sowie die Entwicklung eigener bzw. Berücksichtigung anderer Perspektiven (personale Ausprägung) und ihre Adressat*innenorientierung (soziale Kompetenzausprägung) mit ein. Die Untersuchung der einzelnen Teilfähigkeiten ist dann orientiert an der elokutionellen, einzelsprachlichen sowie expressiven Analyseebene (vgl. Grundler 2011: 65ff.). Grundler (2011) legt somit eine Studie vor, mit der u. a. auch die von den Sprecher*innen eingesetzten sprachlichen Formen (i. S. einzelsprachlicher Fähigkeiten) rekonstruiert und hinsichtlich der argumentativen Anforderungen eingeordnet werden können. Diese Ebene wurde beispielsweise in Vogts Untersuchung, die vornehmlich auf die Komplexität und sequenzielle sowie interaktive Einpassung der Schüler*innenbeiträge sowie das dazu eingebrachte Wissen der Schüler*innen rekurriert, noch nicht in den Blick genommen. Grundlers Studie ermöglicht außerdem, die Ebenen der Kompetenzausprägungen mit der elokutionellen, einzelsprachlichen sowie expressiven Analyseebene in Beziehung zueinander zu setzen. Dadurch können Zusammenhänge zwischen der Ausprägung einzelner Teilfähigkeiten bzw. Kompetenzausprägungen aufgezeigt werden (z. B. dass Sprecher*innen Schwierigkeiten im initiiierenden Einbringen von Beiträgen zeigen (konversationelle Ausprägung), wenn ihr Wortschatz nicht weit genug ausgebaut ist (einzelsprachliche Ebene; Grundler 2011: 390)). Erst indem Grundler (2011) eine bewusste Differenzierung der einzelnen Kompetenzausprägung vornimmt, kann eine Vielzahl an Anforderungen freigelegt werden, mit denen jugendliche Sprecher*innen während des Argumentierens umgehen müssen. Die Studie leistet somit einen wertvollen Beitrag zur Bestimmung einzelner Teilfähigkeiten bzw. Kompetenzausprägungen, die für das mündliche Argumentieren relevant werden, sowie deren Zusammenhänge untereinander. In der letzten Studie adaptiert Krelle (2014) zur Bestimmung argumentativer Fähigkeiten von Schüler*innen ein bereits bestehendes Modell allgemeiner Gesprächskompetenz nach Becker-Mrotzek (2009). Entlang von *fünf Anforderungsbereichen* bzw. sogenannten *Managementaufgaben* entwickelt der Forscher ein Argumentationskompetenzmodell, das er anhand der im Unterricht durchgeführten Fishbowl-Diskussionen absichert und dessen Ziel es ist, anhand der dargelegten Anforderungsbereiche ein Beurteilungsraster für die Argumentationskompetenz von Schüler*innen zu entwickeln. Dabei nimmt Krelle die sprachlich-diskursiven Fähigkeiten der Schüler*innen in den Blick, und untersucht, wie die Sprecher*innen ihren Beitrag sequenziell platzieren (Gesprächs- und Handlungsmanagement), ob sie ihren Beitrag mit einem vorangegangenen in Beziehung setzen (Identitäts- und Beziehungsmanagement), wie sie ihren eigenen Beitrag strukturieren (Sprachmanagement) und welche para- sowie außersprachlichen Mittel sie dafür verwenden (Management von Non- und Paraverbalem). In der letzten Managementaufgabe des Themenmanagements untersucht Krelle (2014: 229) schließlich, ob die Sprecher*innen kohärent „eine Position zum Thema einnehmen und im Gesprächsprozess konsistent vertreten [und ggf. revidieren] können“. Der Autor geht mit seiner Studie die Aufgabe an, die rekonstruierten Teilfähigkeiten mündlichen Argumentierens anhand der fünf Anforderungsbereiche zu operationalisieren. Damit ermöglicht er Lehrkräften,

die argumentativen Fähigkeiten ihrer Schüler*innen in Bezug auf die eingebrachten allgemeinen Wissensbestände, das gattungsspezifische Wissen, das Aufnehmen und Abwägen anderer Positionen sowie in Bezug auf die sequenzielle Platzierung von Beiträgen zu reflektieren und zu bewerten. Er schafft einen Anknüpfungspunkt zwischen der Beforschung argumentativer Fähigkeiten von Schüler*innen und der Überführung der Erkenntnisse in ein schulspezifisches Bewertungsraster der Ausprägung individueller argumentativer (Teil-)Fähigkeiten.

Trotz der unterschiedlichen Systematisierungen argumentativer Kompetenz eint die Studien ihr gesprächsanalytisch orientiertes Vorgehen, das die individuellen Fähigkeiten der Schüler*innen in einem wechselseitigen Verhältnis zu den Beiträgen der anderen Sprecher*innen untersucht. Für das Verständnis, aus welchen (Teil-)Fähigkeiten sich die mündliche Argumentationskompetenz zusammensetzt, scheint es gewinnbringend, die jeweils in den vorgestellten Studien rekonstruierten Teilfähigkeiten von Argumentationskompetenz im Folgenden zusammenfassend vorzustellen.

Die Untersuchungen unterscheiden, ob im interaktiven Vollzug der Argumentation der Beitrag eines*ei-ner Sprecher*in *initiiierend* oder *reagierend* eingebracht wird. Während reagierende Beiträge inhaltlich an vorangegangene Äußerungen sowohl wiederholend als auch expandierend anknüpfen können (vgl. Vogt 2002: 121), erfordert das Initiieren eigener Beiträge, dass die Sprecher*innen z. B. eine Position selbstständig und ohne explizite Bezugnahme auf andere Positionen aufbauen – sie können nicht auf bereits im Gespräch Etabliertes zurückgreifen (vgl. Krelle 2014: 157, 225ff., 2015: 67). Grundler (2011: 389f.) zeigt in ihrer Studie, dass es nicht allen Jugendlichen gelingt, Gesprächsbeiträge initiierend einzubringen: Das Einbringen initiativer Beiträge stellt eine Anforderung dar, die von heranwachsenden Sprecher*innen (erst) mit zunehmender Argumentationskompetenz bewältigt werden kann.

Um sich mit eigenen Beiträgen an der Argumentation beteiligen zu können, ist es erforderlich, dass Interagierende ihre Beiträge *sequenziell* und *interaktiv* in die Interaktion einpassen. Das bedeutet, die Beteiligten müssen zum einen erkennen, wo eine turnübergaberelevante Stelle und somit die Möglichkeit eines Sprecher*innenwechsels (*turn taking*, vgl. Sacks et al. 1974; Krelle 2014)⁴¹ besteht. Dazu muss aber nicht nur ausgelotet werden, wann eine Turnübernahme stattfinden kann, sondern der eigene Beitrag muss zum anderen auch an den vorangegangenen Beitrag (passend) anschließen (vgl. Vogt 2002: 240f.; Grundler 2011: 109). Das erfordert von den Sprecher*innen, u. a. Bezugnahmen auf Vorangegangenes herzustellen, Neues zu initiieren etc. und somit durch den eigenen Beitrag wiederum Voraussetzungen für weitere, daran anknüpfende Gesprächsbeiträge zu schaffen (vgl. Vogt 2002: 241).

Für das Einbringen eines eigenen Beitrages ist über die sequenzielle und interaktive Einbettung hinaus wichtig, dass die Sprecher*innen über *Wissen über die kommunikative Gattung* verfügen. Dieses Wissen umfasst beispielsweise, ob Beiträge entsprechend der Gattung als Einwände, Begründungen oder in Form anderer argumentativer Handlungen eingebracht werden (vgl. Krelle 2014: 44, 228; Grundler 2011: Kap. 5.1.2; Vogt 2002: 304f.). Die Beiträge können in ihrer *internen Strukturierung* dann unterschiedliche Grade der Komplexität aufweisen, die Vogt (2002) als (Struktur-)Formate beschreibt. Er unterscheidet, ob die Beiträge der Schüler*innen allein eine Position beinhalten (Elementarformat), durch eine Begründung gestützt werden (Standardformat) oder durch „z.B. Erläuterungen oder Präzisierungen der Begründung“ erweitert werden (erweitertes Format) (Vogt 2002: 257). Erfolgt zudem eine Vertiefung, indem der*die Schüler*in z. B. Belege liefert, resümiert etc. und diese *explizit* sprachlich realisiert, dann handelt es sich um ein komplexes Format (vgl. Vogt 2002: 257). Grundler (2011: 199)

⁴¹ Obwohl Krelle (2014) diese Anforderung selbst einführt, weist seine Studie diesbezüglich aufgrund des Erhebungssettings in Form von moderierten Fishbowl-Diskussionen folgende Limitation auf: Die Rederechtszuweisung erfolgt stellenweise durch den*die Moderator*in, sodass Aussagen über die individuelle Fähigkeit von Sprecher*innen in Bezug auf das Turnmanagement nur in begrenztem Maße formuliert werden können.

ergänzt in dieser Beschreibung noch das am wenigsten komplexe Format der reinen Meinungsäußerung. Insgesamt zeigt sich in der Beschreibung der Formate eine Progression von satzförmigen zu satzübergreifenden Beiträgen mit zunehmenden Strukturierungsanforderungen an die Sprecher*innen. Zur Realisierung eines Beitrages bedarf es letztlich unterhalb der Strukturierungsebene *sprachlicher, parasprachlicher und außersprachlicher Formen*, mit denen Sprecher*innen z. B. erkennbar machen, welche sprachliche Handlung eingeleitet wird (vgl. zu *weil* in Begründungen Grundler 2011: Kap. 6.2.1), wie sie Modalisierungen zur Abschwächung von Einwänden verwenden (vgl. Grundler 2011: 350f.) oder wie Sprecher*innen ihre Beiträge kohäsiv an die Äußerungen anderer Beteiligter zurückbinden (vgl. Grundler 2011: Kap. 5.2; Vogt 2002: 257) und sich dabei auch körperlich auf den*die Adressat*in hin orientieren (vgl. Grundler 2011: 366). Die Anforderungen, die die Sprecher*innen hierbei bewältigen müssen, bestehen also darin, die sprachlichen, parasprachlichen und außersprachlichen Formen produktiv für die Realisierung und Strukturierung der argumentativen Handlungen einsetzen zu können.⁴²

In argumentativen Interaktionen entwickeln Sprecher*innen sowohl *eigene oder grundsätzlich mögliche Perspektiven* auf ein Problem oder einen strittigen Gesprächsgegenstand. Sie beziehen aber auch *fremde Perspektiven* in ihre eigenen Beiträge ein. Die Anforderung besteht für die Interagierenden nun darin, die eigene (neue) Perspektive von anderen zu entkoppeln bzw. eine Relation (integrierend oder abgrenzend) zwischen anderen und eigenen Perspektiven herzustellen (vgl. Vogt 2002: 89). Dabei können die Interagierenden in unterschiedlicher Weise mit divergenten Perspektiven umgehen: Sie können eine *Perspektivenharmonisierung* vornehmen, indem „zwei konkurrierende Perspektiven“ versucht werden, „miteinander in Einklang zu bringen“ (Grundler 2011: 284). Beteiligte können aber auch die Perspektiven anderer, i. S. v. Fremdperspektiven, kritisieren und somit eine *Perspektivenkritik* äußern (vgl. Grundler 2011: 285) oder eine *Perspektivenhierarchisierung* vornehmen, indem Sprecher*innen „konkurrierende Perspektiven einander über- bzw. unter[ordnen]“ (Grundler 2011: 287). Darüber hinaus können divergente Perspektiven auch *synthetisiert* werden, indem Sprecher*innen aus ihnen heraus „ganz neue[], emergente[] Perspektive[n]“ entwickeln (Grundler 2011: 289). Es zeigt sich, dass diskursiv kompetente(re) Sprecher*innen vermehrt Perspektiven hierarchisieren, Über- und Unterordnungen vornehmen und somit abwägend in Beziehung zueinander setzen (können) (vgl. Grundler 2011: 390f.; vgl. auch Hauser und Luginbühl 2017). Die von den Beteiligten zu bewältigende Anforderung besteht also auch darin, zu prozessieren, an welche Positionen sie anknüpfen, mit wem sie Koalitionen bilden und zu wem sie in Opposition treten können (vgl. Krelle 2014: 265ff.). Sowohl das Einbringen eigener als auch die Bezugnahme auf Perspektiven anderer erfordert, dass die Sprecher*innen nicht nur über Gattungswissen (s. o.), sondern auch über *spezifische Wissensbestände* verfügen (müssen).

Im Umgang mit dem in der Argumentation bearbeiteten Dissens bzw. dem bearbeiteten Problem bringen die Sprecher*innen unterschiedliche Wissensbestände ein. Das Wissen kann sich u. a. aus (vorbereiteten) thematischen Wissensbeständen sowie auch individuellem Erfahrungswissen speisen (vgl. Krelle 2014: 152; Grundler 2011: 193). Der Rückgriff auf unterschiedliche Wissensbestände ermöglicht es den Beteiligten, eigene Argumente zu stützen und die Geltung eines Arguments zu steigern (vgl. Grundler 2011: 193). Dies erfordert von den Sprecher*innen, ihr Wissen in einem dem Kontext angepassten De-

⁴² Krelle (2014: 264) legt dabei normative Kriterien an, um die „[s]prachliche Angemessenheit“, die „[s]prachliche Richtigkeit“ oder auch den angemessenen Einsatz para- und außerverbalen Formen zu ermitteln. Eine dezidierte Darstellung zur Entwicklung der Kriterien der sprachlichen Angemessenheit sowie Richtigkeit expliziert Krelle (2014) allerdings nicht. Aus einer primär gesprächsanalytischen Forschungsperspektive, wie sie beispielsweise auch Grundler (2011) darlegt und der sich auch die vorliegende Studie verschreibt, würde eine Einschätzung über die Angemessenheit sowie sprachliche Richtigkeit einer Äußerung ausschließlich anhand der Rückmeldung der Rezipient*innen vorgenommen werden.

taillierungsgrad und Umfang einzubringen (vgl. Grundler 2011: 252f.). Zu den argumentativen Fähigkeiten zählt nun nicht nur, Wissensbestände darzulegen, sondern auch, dass die Interagierenden Wissensbestände im prozessualen Aufbau der Argumentation vernetzen (vgl. Grundler 2011: 198). So zeigt sich bei kompetente(re)n Sprecher*innen, dass sie vermehrt ihr Wissen flexibel in die *interaktive* Argumententwicklung einbetten können, Wissensäußerungen weniger isoliert einbringen (vgl. Grundler 2011: 226) und somit insgesamt ihrer Position oder ihrem Argument mehr Gewicht verleihen (vgl. Grundler 2011: 229). Weniger kompetente Sprecher*innen hingegen verhandeln die Argumente anderer Beteiligter nicht vertiefend weiter und bleiben somit „an der individuellen Begründungshandlung ihrer Position haften“ (vgl. Grundler 2011: 242). Wenn Sprecher*innen nun auch noch Wissen über die Gesprächspartner*innen einfließen lassen, d. h. Wissen darüber, welche Positionen möglicherweise für das Gegenüber anschlussfähig sind (vgl. Grundler 2011: Kap. 8.1.1.2.1), erhöht sich zudem die Chance, dass die eigene Position weiterverhandelt wird. Argumentative Fähigkeiten umfassen also nicht nur, das eigene thematische Wissen in angemessenem Umfang bzw. Detaillierungsgrad zur Unterstützung des eigenen Arguments einzubringen (vgl. Grundler 2011: 252f.), sondern auch, Wissen über Gesprächspartner*innen in die interaktive Argumententwicklung einzubinden (vgl. Grundler 2011: 345).

Zusammenfassend lässt sich für die drei Modellierungen von Argumentationskompetenz festhalten, dass sie insgesamt die Fähigkeit, argumentieren zu können, als ein Bündel unterschiedlicher Teilfähigkeiten rekonstruieren. Dabei arbeiten sie auf Grundlage von Unterrichtsinteraktionen heraus, wie die Heranwachsenden mit *gattungsspezifischen Anforderungen* umgehen, wie sie ihre Beiträge *sequenziell* platzieren und *intern strukturieren* sowie *Bezüge* zu vorangegangenen Äußerungen sowie den darin enthaltenen Perspektiven bzw. Positionen *herstellen*. Alle drei Studien liefern somit wichtige Beschreibungs- und Analysezugänge, um die individuelle Kompetenzausprägung argumentativer Fähigkeiten heranwachsender Sprecher*innen untersuchen zu können, die auch in dem nachfolgend vorgestellten Modell von Quasthoff (2009) relevant werden. Aufgrund der vorliegenden Settings und der Untersuchung der argumentativen Fähigkeiten ausschließlich im schulischen Kontext verbleibt jedoch zu Teilen unbeleuchtet, wie beispielsweise Sprecher*innen *ohne* eine zuvor fest eingenommene Position *interaktiv* die Rollen von Proponent*in und Opponent*in selbstständig aushandeln, dabei Dissense etablieren, gemeinsame oder voneinander abgrenzende Positionen entwickeln und die auftretenden kommunikativen Anforderungen bewältigen. Darüber hinaus bedarf es in Vorbereitung auf Kapitel 10 einer Modellierung von Argumentationskompetenz, die dezidierte Beobachtungen hinsichtlich potenzieller *Erwerbsphänomene* zulässt. Diesem Anspruch trägt die nachfolgend vorgestellte Modellierung von Diskurs- bzw. Argumentationskompetenz von Quasthoff (2009) Rechnung, mithilfe derer argumentative Erwerbsphänomene untersucht werden können. Es handelt sich um ein streng induktiv ermitteltes, empirisch fundiertes, und breit erprobtes Kompetenzmodell. Es ist bereits erfolgreich für unterschiedliche Gattungen adaptiert (z. B. Erzählen, Erklären, Argumentieren) sowie in diversen Kontexten (u. a. Schule, Familie, Peers) und über eine Altersspanne von der Primarschule bis zur Sekundarstufe erprobt worden (vgl. u. a. Kern 2011; Quasthoff und Krah 2012; Heller und Krah 2015; Morek und Quasthoff 2018; Quasthoff et al. 2019; Quasthoff und Kluger 2021; Morek und Heller 2021). Neben der Möglichkeit, kontext-, gattungs- sowie auch medialitätenvergleichende Untersuchungen durchzuführen (vgl. u. a. Quasthoff et al. 2019; Morek 2021), kann die Modellierung auch für die Bearbeitung erwerbsbezogener Fragestellungen auf Grundlage längsschnittlich erhobener Daten genutzt werden, wie sie im Fokus der explorativen Fallstudie (vgl. Kap. 10) stehen. Im Folgenden wird das Modell näher erläutert. Dabei werden durchaus Überschneidungen zu den oben zusammengefassten Teilfähigkeiten hervortreten, die letztlich durch das gemeinsame Forschungsbestreben, Argumentationskompetenz zu modellieren, sowie den methodischen Zugriff über gesprächs- bzw. konversationsanalytische Untersuchungen erklärbar sind.

Nach Quasthoff (2009) setzt sich Diskurs- bzw. Argumentationskompetenz aus insgesamt drei Teilfähigkeiten zusammen: die *Kontextualisierung*, *Vertexung* und *Markierung*. Diese drei Teilfähigkeiten stehen den Interagierenden als *interne* Ressourcen in der Interaktion zur Verfügung (vgl. auch Katz-Bernstein und Quasthoff 2007; Quasthoff 2009; Morek et al. 2017). Anhand Quasthoffs Modellierung wird es nun möglich, in Gesprächsdaten zu ermitteln, wie die Sprecher*innen – orientiert an der kommunikativen Gattung des Argumentierens – ihre Beiträge intern aufbauen und strukturieren (*Vertexung*), sie sequenziell in die gemeinsame Interaktion einpassen (*Kontextualisierung*) und welche sprachlichen und parasprachlichen Ressourcen (*Markierung*) sie dabei verwenden. Dabei zielt der zugrundeliegende Kompetenzbegriff auf die beobachtete *Performanz* einzelner Sprecher*innen, die sie im Rahmen des interaktiven Geschehens zeigen (vgl. Quasthoff 2009). Die Modellierung *argumentativer* Diskurskompetenz wird im Folgenden entlang der benannten Teilfähigkeiten der Kontextualisierung, Vertexung und Markierung vorgestellt.

Kontextualisierung

Die Fähigkeit, kontextualisieren zu können, erfordert im Rahmen der Diskursfähigkeit des Argumentierens, dass Interagierende im Gespräch aufkommende Dissense oder Problemstellungen erkennen und selbst markieren sowie die gattungsspezifischen globalen Gesprächsaufgaben (1.) Dissens herstellen/Problematisieren, 2.) Begründungspflicht etablieren, 3.) Kernjob Begründen, 4.) Überleiten, 5.) Abschließen; vgl. Kap. 2.2.1) kontextangemessen rezeptiv und produktiv bearbeiten können (vgl. Heller 2012; Morek 2015; Morek et al. 2017; Morek 2017; Quasthoff et al. 2019). Quasthoff et al. (2019) explizieren, dass zu dieser Fähigkeit einerseits gehört, dass die Beteiligten die Diskurseinheit des Argumentierens in den Gesprächskontext, d. h. den *turn-by-turn talk* einpassen können, indem sie entsprechend die Randjobs des Problematisierens und Abschließens sowie des Überleitens bearbeiten. Andererseits umfasst die Fähigkeit neben der (Ein-)Passung in den *turn-by-turn talk* auch, „die Einheit in ihrem Vollzug auf die konkreten [...] Rezipierenden und die je spezifischen Produktions- und Rezeptionsbedingungen“ (Quasthoff et al. 2019: 210) zuzuschneiden. Insgesamt zielt diese Teildimension der Argumentationskompetenz also auf *gesprächsstrukturelle* Anforderungen, die es von den Beteiligten zu bearbeiten gilt (vgl. Quasthoff et al. 2019). Dies schließt auch das Erkennen, Befolgen und eigenständige Etablieren von gattungsspezifischen lokalen und globalen Zugzwängen mit ein „und zwar jeweils mit Zuschnitt auf den jeweils spezifischen sequenziellen und sozialen Kontext“ (Morek 2017: 79). Dies meint, dass Interagierende beispielsweise erkennen, dass ein globaler Zugzwang in Form einer Frage mit dem Interrogativpronomen *warum* im Kontext der kommunikativen Gattung des Argumentierens eine Begründung für eine eingebrachte Position und somit einen global strukturierten, satzübergreifenden Redezug erwartbar macht (vgl. Heller 2012; Quasthoff et al. 2019). Inwiefern soziale Beziehungen der Beteiligten oder auch institutionelle Rahmungen i. S. des von Morek (2017: 79) benannten sozialen Kontexts die sprachlich-diskursiven Anforderungen an die Sprecher*innen verändern können, wird Kapitel 4.1.3 darlegen. Kontextualisierungskompetenz meint also insgesamt die Fähigkeit, „zu wissen, wann welche Art von Äußerung innerhalb eines Gesprächskontextes angeschlossen werden muss“ (Morek und Heller 2012: 88).

Vertexung

Das Vertexen bezieht sich auf die *interne* Strukturierung einer Äußerung „mittels lexikalischer, morpho-syntaktischer, prosodischer und körperlich-visueller Formen“, die „für die Rezipient/innen erkennbar gemacht bzw. markiert [wird]“ (Heller 2021a: 306). In Bezug auf das Argumentieren bedeutet dies, dass Sprecher*innen orientiert an der übergeordneten Gattung „[a]rgumentationstypische Mittel“ einsetzen können (Morek 2017: 79), wie beispielsweise Widersprechen, Infragestellen und auch Begründungshandlungen (vgl. Heller 2012; vgl. Kap. 2.2.2). Die eingesetzten Mittel unterscheiden sich dabei

in ihren strukturellen Anforderungen. Quasthoff et al. (2019: 215) sprechen hier von „Organisationsprinzipien“, d. h. die eingebrachten Beiträge der Interagierenden unterscheiden sich hinsichtlich ihrer lokalen bzw. globalen Strukturierung. So erfordert beispielsweise das Widerlegen von den Sprecher*innen die Fähigkeit, ein global strukturiertes Äußerungspaket hervorzubringen, das einen Bezug zur vorangegangenen Äußerung herstellt und „mit dem eine Gegenbehauptung bewiesen bzw. plausibilisiert wird“ (Heller 2012: 86). Im Gegensatz zum lokalen Widersprechen (vgl. Spranz-Fogasy 2005b) wird hier eine Begründung zur Plausibilisierung der eigenen Position oder des eigenen Vorschlags relevant, die in sich deutlich globaler strukturiert ist und somit in Form eines größeren Äußerungspakets, als sogenannter *Multi-Unit Turn*, eingebracht wird (vgl. u. a. Heller 2012; Quasthoff et al. 2021a). Zur Vertextungskompetenz zählt i. d. S. auch, ob die von den Sprecher*innen eingebrachten Beiträge *kohärent* aufgebaut und strukturiert sind. Während dies für die Gattung des Erzählens bedeutet, das eigene Äußerungspaket chronologisch um einen Planbruch der Geschichte zu arrangieren (vgl. Quasthoff 1980; Hausendorf und Quasthoff 1996/2005), werden für die Gattung des Argumentierens eine logische bzw. hierarchische Ordnung des Beitrags insofern relevant, als dass beispielsweise eigene Positionen eingebracht und durch Beispielszenarios in Begründungen veranschaulicht werden (vgl. Quasthoff et al. 2021a: 22). Dabei greifen die Sprecher*innen auf kontextfreies Wissen über die Gattung des Argumentierens „i. S. eines Sequenzmusters“ zurück (vgl. Quasthoff et al. 2019: 210): Sie verfügen beispielsweise über das gattungsspezifische Wissen, dass in Argumentationen eine Position oder durch eine Begründung gestützt und/oder plausibilisiert werden kann (vgl. Heller 2012) und somit u. a. mögliche Einwände antizipierend entkräftet werden können (vgl. auch Kreuz 2021: 297).

Markierung

Unter der Markierungskompetenz wird die Fähigkeit verstanden, sprachliche, parasprachliche sowie multimodale Ressourcen orientiert an der kommunikativen Gattung nutzen und mit ihnen semantisch kohärente Zusammenhänge innerhalb einer Äußerung herstellen zu können. Der Einsatz der verwendeten Ressourcen übernimmt dementsprechend eine relevante Funktion für die Vertextung sowie die Kontextualisierung (vgl. Quasthoff et al. 2019: 217). So können Sprecher*innen beispielsweise mittels dem Junktor *aber* Adversatives markieren, Modalisierungen wie *vielleicht* oder *ich glaube* einsetzen, um Auskunft über epistemische Haltungen zu geben, oder mittels *weil*-Konstruktionen kausale Zusammenhänge markieren etc. (vgl. u. a. Heller 2012; Krah et al. 2013; Morek 2020). Aufgrund der ursprünglichen Fokussierung des Modells auf sprachliche sowie parasprachliche Ressourcen, entwickelt sich erst seit einigen Jahren eine Ausweitung des Fokus auf die Funktion und den Einsatz *multimodaler* Ressourcen während des mündlichen Argumentierens. So weisen Domenech et al. (2018: 24) im Zusammenhang mit Deutsch-als-Zweitsprachler*innen darauf hin, dass Gesten von den heranwachsenden Beteiligten dazu genutzt werden, um „Aussagen, Gewichtungen und Haltungen [...] [zu] veranschaulichen“. Im Gegensatz zu sprachlichen Ressourcen ermöglichen körperlich-sichtbare Ressourcen, Anteile der Äußerungen „über eine etwas längere Zeitspanne präsent [zu] halten“ (Domenech et al. 2018: 24). So können Sprecher*innen durch eingesetzte abbildende Gesten die Aspekte eines Arguments visualisieren und über den eigenen *Turn* hinweg für alle Interagierenden aufrechterhalten. Multimodale Ressourcen tragen also wesentlich zur Markierung bei. Ferner zeigen sich multimodale Ressourcen nicht nur als hochgradig funktional zur Unterstützung der Lexik, sondern auch hinsichtlich pragmatischer sowie epistemischer Intentionen von Sprecher*innen. In Peer-Argumentationen nutzen Sprecher*innen beispielsweise Gesten mit pragmatischer Funktion, um eine gemeinsame epistemische Verantwortung für ein Argument oder eine Begründung zu markieren (vgl. Heller 2018a) oder das reziproke Geben-und-Nehmen von Argumenten anzuzeigen (vgl. Schönfelder und Heller 2019; vgl. Kap. 3.3). In Peer-Argumentationen kann außerdem das Wiederholen pragmatischer sowie abbildender Gesten sowie wei-

terer Ressourcen unterschiedlicher Sprecher*innen dazu beitragen, dass in kollaborativen ko-konstruktiven Prozessen die Kooperation dieser Interagierenden visuell sichtbar werden (vgl. Schönfelder 2024). Insgesamt liefern diese Befunde also wichtige Erkenntnisse dafür, dass multimodale Ressourcen „einen zentralen Beitrag zur Verständnissicherung und Koordination zwischen den Beteiligten [leisten]“ (Domenech et al. 2018: 24). Dabei zeigt sich unter Rückbezug auf Kapitel 3, dass diese Ressourcen nicht einzeln, sondern in Verbänden komplexe multimodale Gestalten (vgl. Mondada 2014b) bilden, die zur Strukturierung und Kontextualisierung einer Äußerung in der Interaktion beitragen (vgl. dazu u. a. auch Heller 2021a; vgl. Kap. 8 & 10).

Ausgehend von dieser Modellierung argumentativer Diskurskompetenz werden im Folgenden Kapitel 4.1.2 nun bisherige Forschungsergebnisse zu deren ontogenetischer Entwicklung vorgestellt.

4.1.2 Linien des Diskurserwerbs

Die Modellierung interner Erwerbsressourcen von mündlicher Argumentationskompetenz entlang der drei Teildimensionen Kontextualisierung, Vertextung und Markierung ermöglicht es nun, auf Grundlage qualitativer und rekonstruktiver Längsschnittuntersuchungen sich abzeichnende Erwerbslinien heranwachsender Sprecher*innen in einer überblicksartigen Zusammenschau darzulegen. Einbezogen werden Studien, die ihrerseits auf die beschriebene Modellierung von Argumentationskompetenz von Quasthoff (2009; vgl. Kap. 4.1.1) referenzieren. Die Befunde zur Entwicklung der einzelnen Teilfähigkeiten (Kontextualisierung, Vertextung und Markierung) werden analytisch getrennt voneinander vorgestellt. Insgesamt sind die sich abzeichnenden Linien des Diskurserwerbs als *Entwicklungstendenzen* zu verstehen, die aus dem aktuellen Forschungsstand hervorgehen.

In ihrer Longitudinalstudie, die auf Kinder im Alter von 2 bis 4 Jahren⁴³ sowie in einem Vergleichskorpus auf 3- bis 6-jährigen Kindergartenkinder fokussiert, zeigt Arendt (2019), dass diese auf der Ebene der *Kontextualisierung* noch keine bzw. wenig Orientierung an den gesprächsstrukturellen, argumentativen Aufgaben zeigen. Die untersuchten argumentativen Sequenzen sind v. a. auf Spielinteraktionen und Aushandlungen, die im Hier und Jetzt stattfinden, ausgerichtet (Arendt 2019: 361). Im Altersvergleich zeigt sich, dass 3-Jährige Dissense durch wechselseitiges und wiederholtes Insistieren der eigenen Position u. a. in Form von (behaupteten oder erwünschten) Besitzansprüchen anzeigen (vgl. *Ping-Pong-Spiel* bei Arendt 2019; *reziproke Wortbatterien* Komor 2010: 158ff.; *Aufschlagwechsel* bei Knoblauch 1995: 122). Der Dissens wird dabei allerdings seltener durch das Einbringen und Bearbeiten weiterer Jobs ausgehandelt, sondern im Anschluss an das (wiederholte) Opponieren fallengelassen, physisch ausgesagt oder durch das Hinzuziehen von Erwachsenen beendet (vgl. Arendt 2019: 330, 345, 361). Die untersuchten Interaktionen zeigen, dass eine explizite Problematisierung sowie auch Begründungen fehlen, sodass die Autorin die „kontextuelle[] Passung als gering ein[stuft]“ (Arendt 2019: 330). Im längsschnittlichen Vergleich innerhalb der Studie zeigt sich dann, dass die Fähigkeit der Kontextualisierung und somit der rezeptive und produktive Umgang mit den diskursiven Anforderungen mündlichen Argumentierens im Alter zwischen vier und sechs Jahren deutlich zunehmen: Dissense werden markiert, erkannt und zunehmend in den Peer-Interaktionen begründend bearbeitet. Es scheint i. S. der Studie berechtigt anzunehmen, dass Kinder mit zunehmender Kontextualisierungskompetenz Einsicht in die Funktionalität von Begründungen als die eigene Position stärkendes Mittel erhalten und diese gesprächsstrukturelle Aufgabe zunehmend autonom übernehmen. Zwar machen sich die Heranwachsenden in dem von Arendt (2019) untersuchten Erhebungszeitraum auch gegenseitig begründungspflichtig, indem sie eigenständig eine explizite Begründungspflicht etablieren, dies tritt in den Daten aber insgesamt

⁴³ Die Interaktionen der Kinder wurden über einen Zeitraum von 4 Jahren aufgezeichnet (Arendt 2019).

selten auf (vgl. Arendt 2019: 334). Das selbstständige Übernehmen des Abschließens einer argumentativen Sequenz ist in der Studie auch bis zum Alter von 6 Jahren nicht zu beobachten. Im Grundschulalter zeigen Heranwachsende weitere Entwicklungen im Hinblick auf den rezeptiven und produktiven Umgang mit den lokalen und globalen Anforderungen des mündlichen Argumentierens. So weisen u. a. Domenech et al. (2018) anhand videographierter Peer-Argumentationen von Drittklässler*innen nach, dass die Beteiligten autonom einen Dissens hervorbringen und explizit Begründungspflicht einander gegenüber etablieren können. Diese Befunde stützen auch einzelne Studien zu Peer-Argumentationen, aus denen hervorgeht, dass Kinder bereits ab der zweiten Jahrgangsstufe explizit eine Begründungspflicht etablieren und dadurch Begründungen von ihrem Gegenüber elizitieren (vgl. Hauser und Luginbühl 2017; Kreuz 2021). Gerade die Explizierung der Begründungspflicht konnte Arendt (2019) in ihren Daten bis zum sechsten Lebensjahr nicht nachweisen. Berücksichtigt man zudem Studien von Heranwachsenden mit erwachsenen Gesprächspartner*innen, sind ebenfalls Veränderungen in der Kontextualisierungskompetenz zu verzeichnen: So zeigen Jugendliche im Alter der Sekundarstufe I, dass sie mit 12 Jahren die lokalen und globalen Anforderungen noch vornehmlich reaktiv bearbeiten, wohingegen sie mit 15 Jahren vermehrt selbst globale Anforderungen an den*die Erwachsene*n stellen (vgl. Heller und Krah 2015). Insgesamt stellt Quasthoff (2009) gattungsübergreifend eine Entwicklungslinie von einer *reaktiven* zu einer *autonomen* Ausprägung der Kontextualisierungsfähigkeit fest (vgl. auch Quasthoff et al. 2019: 219).

Auf Ebene der *Vertexung* zeigt sich in der Entwicklung der Heranwachsenden, dass es ihnen erst mit zunehmendem Alter sowie entsprechender diskursiver Erfahrung möglich ist, global strukturierte Äußerungspakete in Form von Begründungen, Entkräftungen von Einwänden, Konzessionen oder auch in Form des Widerlegens aufzubauen (vgl. u. a. Heller und Krah 2015; Hauser und Luginbühl 2017; Heller 2017). Während die Äußerungen 3-jähriger Kinder in ihrer internen Struktur noch deutlich lokal organisiert sind und dabei aus satzförmigen opponierenden Redezügen bestehen (vgl. Arendt 2019), gelingt es Kindern im Alter von 6 Jahren eine *Orientierung* an global strukturierten Äußerungen anzuzeigen, wenngleich die eingebrachten Argumente additiv aneinandergefügt und (noch) nicht zu einem einzigen Äußerungspaket zusammengeschnürt werden (Arendt 2019: 346; vgl. zur ‚Reichweite‘ der Vertexungsverfahren Quasthoff et al. 2019: 215). Während Kinder auch im Grundschulalter, wie die Studien von Kreuz und Luginbühl (2020) sowie Kreuz (2021) für Zweitklässler*innen zeigen, in argumentativen Entscheidungsfindungen ebenfalls vermehrt satzförmige Propositionen mit vereinzelt Begründungen einbringen, gelingt es jugendlichen Sprecher*innen der sechsten Klasse in argumentativen Interaktionen unter Peers (vgl. Hauser und Luginbühl 2017; Kreuz und Luginbühl 2020), andere Perspektiven in die eigenen argumentativen Handlungen zu integrieren. Heller und Krah (2015: 17) beobachten, dass Heranwachsende im Alter von 15 Jahren in Eltern-Kind-Argumentationen (vgl. Heller und Krah 2015) dann auch Perspektiven gegeneinander abwägen oder diese hierarchisieren (vgl. hierzu auch Grundler 2011; vgl. Kap. 4.1.1). Die Autor*innen beschreiben das hierbei sichtbar werdende Vertexungsverfahren als zunehmend „integrierend und hierarchisierend“ (vgl. Heller und Krah 2015: 17). Über die untersuchten Gattungen des Erzählens, Erklärens und Argumentierens hinweg zeigt sich ein Erwerbsverlauf von einer zu Beginn *lokal operierenden* Vertexungskompetenz hin zu der Fähigkeit, *global strukturierte, satzübergreifende* Äußerungspakete zu produzieren (vgl. Quasthoff 2009; Quasthoff et al. 2019: 219).

Auf der Ebene der *Markierung* ist in der kommunikativen Gattung des Argumentierens ein Anstieg explizit-sprachlicher Markierungen zu verzeichnen, die die Struktur der eingesetzten Mittel auf der Ebene der Vertexung markieren sowie zur Kontextualisierung beitragen (vgl. Heller und Krah 2015; Arendt 2019). Arendts (2019: 328ff.) qualitative Untersuchung liefert deutliche Hinweise darauf, dass 3-Jährige weder die argumentativen Aufgaben des Problematisierens noch des Begründens explizit markieren, wenngleich sie diese durch Insistieren auf der eigenen Position sowie Opponieren bearbeiten

(vgl. oben, Kontextualisierung). Diese Fähigkeit nimmt bis zum sechsten Lebensjahr jedoch zu, sodass die Sprecher*innen beginnen, sowohl Problematisierungen als auch Begründungen vermehrt explizit-sprachlich beispielsweise durch die Verwendung von adversativen Junktoren wie *aber* oder kausaler Konjunktionen wie *weil* zu markieren (Arendt 2019: 334). Auch in Argumentationen unter Peers der sechsten Klasse zeigen sich diese Formen als Abgrenzungsmarkierungen zur vorangegangenen Äußerung (vgl. Hauser und Luginbühl 2017: 101). Im Alter der Sekundarstufe I werden dann in Eltern-Kind-Argumentationen von Jugendlichen deutlich komplexere sprachliche Ressourcen angewandt wie „argumentative Funktionsverbgefüge (»zur Debatte stehen«), Mikroformeln mit argumentativer Funktion (»der Nachteil ist, dass«) und vermehrt Mittel der Hervorhebung und Gewichtung, (»jeweils schon«, »wirklich«, »unter Umständen«, »mindestens«)“ (Heller und Krah 2015: 17). Auch wächst innerhalb des Erwerbsprozesses das Repertoire an Modalisierungen und Vagheitsmarkierungen an, die Aufschluss über die epistemische Haltung und Verantwortung geben (vgl. Heller 2018a; Morek 2020), und es werden u. a. vermehrt lexikalische Markierungen und syntaktische Anschlüsse zum Aufbau ko-konstruierter Begründungen eingesetzt (vgl. Hauser und Luginbühl 2017; Kreuz und Luginbühl 2020; Kreuz 2021). Quasthoff (2009) beschreibt die sich hier abbildende Entwicklungslinie als eine, die sich gattungsübergreifend von einer *impliziten* zu einer *expliziten* Markierungskompetenz vollzieht (vgl. auch Quasthoff et al. 2019: 219).

Die hier vorgestellten Linien des Diskurserwerbs entlang der internen Ressourcen erweisen sich für die Modellierung von Argumentationskompetenz als hochgradig relevant. Die in den Studien rekonstruktiv ermittelten Teilkompetenzen von Argumentationskompetenz erlauben die Beschreibung, Untersuchung und im Hinblick auf ihr didaktisches Potenzial auch mögliche Ableitungen für die Unterstützung des argumentativen Diskurserwerbs (vgl. u. a. Heller und Morek 2015; Heller 2017; Quasthoff et al. 2019; Morek und Heller 2021; vgl. Kap. 12). Im Folgenden (vgl. Kap. 4.1.3) soll nun dargelegt werden, inwiefern neben den *internen* Erwerbsressourcen der Fähigkeiten zur Kontextualisierung, Textextung und Markierung, auch die (sozialen) Kontexte der Familie und der Schule als *externe* Erwerbsressourcen auf die Entwicklung diskursiver Kompetenzen wirken können.

4.1.3 Externe Erwerbsressourcen: Erwerbskontexte für das mündliche Argumentieren

Die Fähigkeit, mündlich argumentieren zu können, basiert nicht nur auf der Entwicklung der internen Ressourcen, sondern ist ebenfalls bedingt durch unterschiedliche Gesprächskontexte wie der Familie, der Peers oder der Schule. Je nach Kontext variieren dabei u. a. auch die an die Heranwachsenden gestellten diskursiven Anforderungen, wie das vorliegende Kapitel 4.1.3 darlegen wird. Insgesamt erweist sich der Erwerb diskursiver Fähigkeiten als multifaktoriell beeinflusst, sodass die individuelle Ausprägung nicht auf *eine* einzelne Erwerbsbedingung zurückgeführt werden kann, sondern interne und externe Ressourcen *gemeinsam* relevant für die Ausbildung sprachlich-diskursiver Fähigkeiten werden (vgl. u. a. Wild et al. 2012). Exemplarisch werden im Folgenden Befunde aus familialen und schulischen Kontexten vorgestellt, die insofern als externe Erwerbsressource fungieren (können), als dass Heranwachsende in ihnen sprachlich-diskursiv gefordert werden und in den Kontexten zugleich ihre Fähigkeiten unterschiedlich erproben können. Der familiale Kontext ist von besonderer Relevanz für die vorliegende Studie, da sie auf der Grundlage von Eltern-Kind-Interaktionen basiert. Der schulische Kontext wird zusätzlich beleuchtet, um die abgeleiteten erwerbsbezogenen Hypothesen der Fallstudie in Kapitel 10 abschließend auch hinsichtlich möglicher sprachdidaktischer Implikaturen einordnen zu können (vgl. Kap. 12).

Familie

Im spezifischen Kontext familialer Interaktion verweisen Befunde unterschiedlicher Studien darauf, dass Familien einen sprachlich-diskursiven Anregungsgehalt für Heranwachsende bieten können (vgl. Wild et al. 2012; Heller 2012; Morek 2012; Morek 2021). Anhand familialer Tisch- bzw. Mahlzeiten-gespräche (vgl. u. a. Pontecorvo und Fasulo 1997; Heller 2012) zeigt sich, dass je nach Familie unterschiedliche Gattungsrepertoires (vgl. Grundler 2011; Heller 2012) hervorgebracht werden, die den Heranwachsenden das Erproben entsprechender (komplexer) diskursiver Praktiken ermöglichen. In den Tisch- bzw. Mahlzeitengesprächen sowie in elizitierten Eltern-Kind-Argumentationen (vgl. Wild et al. 2012; Heller und Krah 2015) zeigen sich nun Unterschiede hinsichtlich der diskursiven Anforderungen, die die erwachsenen Sprecher*innen an die Kinder bzw. Jugendlichen stellen (vgl. Heller 2012: 156): Während einige Familien den Heranwachsenden im Rahmen habitualisierter Formen der gemeinsamen Interaktion ermöglichen, komplexe globale Diskurspraktiken zu erproben und beispielsweise Dissense zu verhandeln, zeigen sich in anderen Familien Tendenzen zu eher lokal organisierten Diskursaktivitäten wie Frotzeln oder Ermahnen (vgl. u. a. Heller 2012: 106f.), die wenig Erwerbspotenzial für global organisierte Diskursaktivitäten wie das Argumentieren bieten.

Diesbezüglich zeigen auch weitere Studien interessante Befunde hinsichtlich der elternseitig angebotenen ‚Erprobungsräume‘ – nun mit spezifischem Fokus auf das mündliche Argumentieren: Basierend auf einer Longitudinalstudie, in denen Eltern-Kind-Argumentationen in der sechsten, siebten und neunten Klasse erhoben worden sind, rekonstruieren Heller und Krah (2015) drei (elternseitig) habitualisierte Interaktionsmuster, die jeweils ein unterschiedliches Erwerbspotenzial für die Ausbildung argumentativer Fähigkeiten mit sich bringen: das *Fordern und Unterstützen*, das *Raumlassen und Akzeptieren* sowie das *Übergehen und Selberlösen* (vgl. auch Quasthoff und Kluger 2021).⁴⁴ Beim Fordern und Unterstützen handelt es sich um ein Interaktionsmuster, das bereits von Hausendorf und Quasthoff (1996/2005) für das Erzählen sowie von Heller (2012: 158ff.) für das Argumentieren in familialen Tischgesprächen beschrieben worden ist. Hier unterstützen die erwachsenen Sprecher*innen Kinder in der Bewältigung der diskursiven Anforderungen durch Nachfragen, Explizierung der Gattung, in die hineingesteuert wird, etc. Hausendorf und Quasthoff (1996/2005) nennen diese Form der interaktiven Unterstützung diskursiver Praktiken *Discourse Acquisition Support System* (DASS). In dem von Heller und Krah (2015) für das Argumentieren mikroanalytisch rekonstruierten Interaktionsmuster Fordern und Unterstützen stimmen die Eltern ihr sprachlich-diskursives Handeln auf das Kompetenzniveau des Kindes ab, was Hausendorf und Quasthoff (1996/2005: 278) *Finetuning* nennen (vgl. auch Quasthoff und Kern 2007; Quasthoff und Krah 2012), fordern es leicht über den kindseitigen Fähigkeiten (i. S. der proximalen nächsten Entwicklungsstufe n. Vygotsky 1978, zit. n. Heller und Krah 2015) und unterstützen es im Vollzug durch Zuhörer*innenseitige Unterstützungsaktivitäten wie das Setzen von Zugzwängen, interessierte Nachfragen etc. Im Längsschnitt wird das *„Mitwachsen der argumentativen Anforderungen“* (Heller und Krah 2015: 11, Hervorh. i. O.) und somit zugleich das erwerbssupportive Potenzial dieses

⁴⁴ Zur Rekonstruktion der Interaktionsmuster haben die Autorinnen die Eltern-Kind-Argumentationen mit dem Analyseinstrument GLOBE (vgl. Hausendorf und Quasthoff 1996/2005) längsschnittlich untersucht und konnten somit nachvollziehbar darlegen, wie das jeweilige Elternteil mit seinem Kind die Jobs bearbeitet, welche Anteile dabei das Kind übernimmt und welche pragmatischen Mittel zur Bearbeitung der Jobs sowohl eltern- als auch kindseitig eingesetzt werden (vgl. Heller und Krah 2015: 7f.). Durch die längsschnittliche Anlage der Studie konnten die Autorinnen dann ebenfalls ermitteln, inwiefern über den gesamten Erhebungszeitraum von der sechsten bis zu neunten Klasse Veränderungen der diskursiven Fähigkeiten i. S. der Anteile bei der Bearbeitung der Jobs seitens des Kindes zu beobachten sind.

Interaktionsmusters sichtbar, denn es steigen sowohl die elternseitigen Anforderungen an das Kind sowie aber auch die kindseitigen Anteile an der Bearbeitung der diskursiven Aufgabe.⁴⁵ Das zweite Interaktionsmuster nennen die Autorinnen Raumlassen und Akzeptieren (vgl. Heller und Krah 2015: 11). Es unterscheidet sich von dem zuerst beschriebenen Fordern und Unterstützen insofern, als dass dem Kind zwar Gesprächsraum geboten wird, um sich interaktiv an der Bearbeitung der Aufgabe zu beteiligen, jedoch stellt das Elternteil keine sprachlich-diskursiven, d. h. argumentativen Anforderungen an das Kind, die zu einer weiteren Elaboration oder Begründung führen würden. Die eingebrachten Beiträge der Interagierenden bleiben vermehrt unverbunden nebeneinander stehen (vgl. auch Quasthoff und Kluger 2021: 117). In der längsschnittlichen Untersuchung der Realisierung der Gesprächsaufgaben, d. h. der Jobs, zeigt sich dieses Interaktionsmuster zumindest nicht als erwerbshemmend; weitere eindeutige Erkenntnisse lassen sich jedoch nicht ableiten (vgl. Heller und Krah 2015: 16). Das dritte Interaktionsmuster, Übergehen und Selberlösen, ist geprägt durch eine*n erwachsene*n Interaktionspartner*in, der*die primär selbst die Rolle des*der Aufgabenlöser*in und somit „alle zentralen Gesprächsaufgaben [übernimmt]“ (Heller und Krah 2015: 12). Im Längsschnitt zeigt sich ein „geringeres erwerbsupportives Potenzial“, denn die Beteiligung der Heranwachsenden an der Bearbeitung der Aufgaben wächst nur minimal (Heller und Krah 2015: 16). Neben dem Erwerbspotenzial, das sich in der Beteiligungsstruktur manifestiert, die die erwachsenen Sprecher*innen den Jugendlichen in der Longitudinalstudie (6.-9. Klasse) ermöglichen, leiten die Autorinnen ebenfalls Ergebnisse im Hinblick auf das Erwerbspotenzial der internen Ressourcen (Kontextualisierung, Vertextung, Markierung) ab. Während für das Interaktionsmuster Fordern und Unterstützen über den Längsschnitt auch höhere Anforderungen an die Fähigkeit zu kontextualisieren festzustellen sind, indem beispielsweise fiktive Szenarios eingebracht werden, die von dem Kind „interaktiv mitvollzogen werden“ müssen (Heller und Krah 2015: 17), bleiben diese Anforderungen sowohl beim Raumlassen und Akzeptieren wie auch beim Übergehen und Selberlösen aus. Mit den steigenden Anforderungen des Interaktionsmusters Fordern und Unterstützen auf der Ebene der Kontextualisierung erhöhen sich zudem die Anforderungen auf der Ebene der Vertextung. So erfordert beispielsweise das Einbringen imaginierter, fiktiver Szenarios auch eine komplexere Strukturierung der Beiträge, indem beispielsweise durch Konditionalgefüge auf spezifische Folgen eines Szenarios hingewiesen werden (Heller und Krah 2015: 17). Dieser Anstieg der Anforderungen wirkt sich dann auch auf die Ebene der Markierungskompetenz aus, da z. B. eingeführte Konditionalgefüge entsprechender explizit-sprachlicher Markierung bedürfen. Ähnliches lässt sich nicht für die anderen beiden Interaktionsmuster feststellen. Vielmehr scheint die Entwicklung des Kindes auf den Ebenen der Vertextung und Markierung hier von der Eigeninitiative des Kindes abzuhängen, die (wenigen) Gesprächsräume der Interaktion zu nutzen. Heller und Krah (2015: 19) deuten darauf hin, dass kindseitig möglicherweise dann auf Erfahrungen aus anderen Erwerbskontexten, wie argumentativen Peer-Interaktionen sowie Erfahrungen aus dem Unterricht, zurückgegriffen wird. Insgesamt changieren die hier präsentierten Interaktionsmuster in den Eltern-Kind-Argumentationen zwischen erwerbsupportiven und erwerbshemmenden externen Ressourcen, die somit von besonderer Relevanz für die Ontogenese argumentativer Fähigkeiten sind.

⁴⁵ Quasthoff et al. (2021a: 25) beschreiben diesen Mechanismus, der innerhalb des Interaktionsmusters des Forderns und Unterstützens sichtbar wird, mit der Metapher einer Wippe, bei der sich in der „Arbeitsteilung zwischen Kind und Erwachsenem bei der Erledigung der konversationellen Arbeit [...] mit zunehmender Kompetenz der Kinder die Anteile immer mehr angleichen“.

Schule

Auch die Schule gilt als Erwerbskontext diskursiver Fähigkeiten, in dem diese nicht nur außerunterrichtlich mit Peers erprobt werden, sondern der Unterricht selbst potenzielle Erwerbschancen bereithält (vgl. Heller 2017; Heller et al. 2017; Heller 2021a; Leßmann 2021). Die kontextuelle Rahmung unterscheidet sich nun von dem bereits beschriebenen Erwerbskontext der Familie allerdings einerseits im Hinblick auf die institutionellen Rollen der Beteiligten, die auch auf die Interaktionssituation selbst wirken. Denn in Unterrichtsgesprächen treten i. d. R. einzelne Schüler*innen abwechselnd mit der Lehrperson in Interaktion, die den Schüler*innen zuvor das Rederecht erteilt hat – ein Gespräch, in das mehr Beteiligte gleichermaßen involviert sind, findet dabei selten statt (vgl. hierzu auch Morek 2012). Andererseits unterscheiden sich Argumentationen im Unterrichtsgespräch auch insofern von Erwerbskontexten wie dem der Familie, als dass die Diskurspraktik für das fachliche Lernen eingesetzt wird, in dem das „Wissen [...] nicht einseitig von der Lehrerin zur Verfügung gestellt, sondern gemeinsam mit den Schülerinnen und Schülern erschlossen [wird]“ (Heller und Morek 2015: 10). Dabei kann das Argumentieren zum Aushandeln von im Unterricht aufgetretenem Strittigen bzw. divergenten Wissenssystemen oder dem Explorieren von Problemstellungen (vgl. Ehlich 2014) beitragen. In dieser gemeinsamen Erarbeitung des Wissens bieten Unterrichtsgespräche einen Erwerbskontext für den Auf- und Ausbau argumentativer Fähigkeiten – die Diskurspraktik ist somit (genuiner) Teil der im Unterricht stattfindenden Lernprozesse (vgl. Heller 2021a).⁴⁶ Allerdings zeigen Untersuchungen von Unterrichtsinteraktionen, dass Lehrkräfte nur selten explizieren, welche diskursive Praktik sie von den Schüler*innen erwarten bzw. in welche Gattung der jeweilige Zugzwang steuert, sodass schüler*innenseitig Schwierigkeiten entstehen können, die jeweilige diskursive Anforderung zu erkennen (vgl. u. a. Heller und Morek 2015; Leßmann 2021; Quasthoff et al. 2022). Anders als in anderen Kontexten werden den Schüler*innen im Unterricht außerdem häufig nur Möglichkeiten einer sehr lokalen Beteiligung am Aufbau einer Diskurs Einheit zugestanden:

Selbst wenn diskursive Anforderungen gestellt werden, werden sie nicht selten zergliedert, indem die Lehrperson zentrale Jobs übernimmt oder aber diese auf mehrere Schülerinnen und Schüler aufteilt [...]. Ein Grund für dieses Vorgehen könnte sein, dass Lehrkräfte versuchen, möglichst viele Schülerinnen und Schüler am Diskurs zu beteiligen (Heller 2012; Morek 2012). (Leßmann 2021: 286)

Die Schüler*innen werden somit zu „small piece suppliers“ (Quasthoff et al. 2022: 72). Das Erwerbspotenzial wird also auch durch die rar angebotenen Lern- bzw. Erwerbsgelegenheiten, in denen Schüler*innen *größere* Äußerungspakete hervorbringen und das Kontextualisieren, Vertexten und Markieren erproben können, eingeschränkt. Darüber hinaus zeigen sich in untersuchten Unterrichtsdaten häufig keine *interaktiven* Unterstützungsverfahren wie Rückmeldungen seitens der Lehrer*innen, die den Schüler*innen im Prozess des Aufbaus der Diskurseinheit helfen. Vielmehr werden die Zuhörer*innenaktivitäten zu minimalen Rückmeldungen zurückgefahren (vgl. Heller und Morek 2015: 13). Lehrkräfte schöpfen also die Möglichkeiten, Schüler*innen im Erwerb diskursiver Kompetenzen zu unterstützen, in der Unterrichtsinteraktion nicht aus, obwohl sich auch hier potenzielle Erwerbsgelegenheiten zeigen: So können Lehrkräfte Schüler*innen beispielsweise in der Produktion eines Äußerungspakets unterstützen, indem sie durch Nachfragen Reparaturen initiieren und Stellen markieren, an denen Aspekte noch unklar oder zu wenig elaboriert sind (vgl. Leßmann 2021). Es ist also möglich, Heranwachsenden i. S.

⁴⁶ Argumentative (sowie explanative) Anforderungen sind v. a. vermehrt ab der Klassenstufe 5 vorzufinden. Im Primarbereich überwiegen diskursive Praktiken wie das Erzählen und Berichten (vgl. Prediger et al. 2016).

des *Scaffoldings*⁴⁷ in der Interaktion ein Gerüst zur Verfügung zu stellen, das ihnen hilft, komplexere Äußerungspakete auch in der Unterrichtsinteraktion aufzubauen und zu erproben – allerdings scheinen die institutionell geprägten Kommunikationsbedingungen dies zu erschweren (vgl. Heller und Morek 2015: 13f.; Grundler 2022: 585ff.).

Das Argumentieren, das bis hierhin als Lernmedium und somit als Diskurspraktik besprochen worden ist, die Schüler*innen eine Partizipation an Lehr-/Lernprozesse ermöglicht, kann allerdings auch als Lerngegenstand bzw. Lernziel explizit metadiskursiv im Unterricht behandelt werden (vgl. Heller und Morek 2015: 4). Spiegel (1999; 2006b; 2006a), Grundler (2015), Morek (2017) sowie auch Mundwiler et al. (2017) legen Befunde vor, die zeigen, wie Argumentationen (beispielsweise als Pro- und Kontra-Argumentationen in Form von Fishbowl-Diskussionen von Lehrmitteln begleitet) vorbereitet und dann im Unterricht der Sekundarstufe I durchgeführt werden. Kennzeichnend für die Durchführung dieser institutionell geprägten Form des (vorbereiteten) Argumentierens ist, dass

- den Interagierenden feste Rollen von Proponent*innen und Opponent*innen zugewiesen werden;
- die Schüler*innen zu einem vorgegebenen Gegenstand, der potenziell strittig erscheint, eine feste, nicht veränderbare Position einnehmen und diese in der Übungssituation (häufig in Form von exponierten Diskussionsrunden, die vor der gesamten Klasse stattfinden) vertreten sollen;
- die Argumente vorab z. T. in Gruppen erarbeitet und aufgeschrieben werden;
- an die erarbeiteten Argumente seitens der Lehrwerke normative Anforderungen bezüglich des Aufbaus (z. B. dass Behauptungen durch Begründungen abgesichert und durch Beispiele gestützt werden sollen) gestellt werden, die von den Schüler*innen berücksichtigt werden sollen (vgl. Mundwiler et al. 2017);
- und dass die konkrete Aushandlung dann von einem*einer Moderator*in durchgeführt wird, der*die u. a. das Rederecht zuweist, zeitliche Vorgaben der Beiträge koordiniert und dadurch auch Einfluss auf die sequenzielle Organisation nehmen kann (vgl. Morek 2017).⁴⁸

Mundwiler et al. (2017: 103, Hervorh. i. O.) verweisen darauf, dass dieser Form des Argumentierens dann häufig ein „*monologisch-schriftlicher Habitus*“ innewohnt. Mündliche Argumentationen unterscheiden sich aber von schriftlichen Argumentationen insofern, als dass i. S. der Lehrwerke und auch der Lehrkräfte als unvollständig eingeordnete Argumente, in denen „ein Beispiel bereits als Argument“ eingesetzt wird, in mündlichen Interaktionen besonders funktional sein können, „da somit rasch und ökonomisch eine überzeugende Meinung in die Diskussion eingebracht werden kann“ (Mundwiler et al. 2017: 103). Die durch Lehrwerke vermittelten normativen Anforderungen an das mündliche Argumentieren (z. B. in Bezug auf die feste Position von Sprecher*innen, den Verlauf der Argumentationen sowie die Struktur der einzelnen Argumente) legen eine Nähe zu Inhalten der theoretischen Argumentationsforschung nahe (vgl. Kap. 2).

Es wird ersichtlich, dass sich die Vermittlung des mündlichen Argumentierens im Unterricht als Lerngegenstand und Lernziel deutlich von dem familialen Erwerbskontext unterscheidet. Auswirkungen scheinen isolierten Übungssettings v. a. auf die Fähigkeitsdimension der Vertextung zu haben, wie Morek (2017) feststellt. Denn indem Lehrwerke betonen, dass ein ‚gutes‘ Argument aus einer Behauptung oder These, einer Begründung sowie einem Beispiel bzw. Beleg besteht, „wird [...] die Dimension der

⁴⁷ Der Ansatz des *Scaffoldings* ist von Gibbons (2002, zit. n. Quasthoff et al. 2022: 60) ursprünglich für den Zweitspracherwerb entwickelt und u. a. für den Erwerb diskursiver Fähigkeiten adaptiert worden.

⁴⁸ Heller und Morek (2015: 3) nennen dies auch das „isolierte[] Einüben“ von diskursiven Praktiken, bei dem im Gegensatz zur „integrierte[n] Förderung“ die Praktik selbst zum expliziten Lerngegenstand einer Unterrichtsstunde oder -reihe wird.

Vertextung [...] überbetont zu Lasten einer mangelnden Berücksichtigung der Interaktivität und Situiertheit des Argumentierens und somit der *Kontextualisierungsdimension*“ (Morek 2017: 81, Hervorh. i. O.; vgl. auch Grundler 2022: 585ff.). Diese Form des Argumentierens, die im Unterricht explizit als Lerngegenstand bzw. Lernziel vermittelt wird, stellt, Morek (2017) folgend, eine Art Sonderform des Argumentierens dar, die sich deutlich vom faktischen Argumentieren (vgl. Grundler 2015: 47) unterscheidet: Der kommunikative Zweck, wie das Aushandeln eines Dissenses oder Problems, den die diskursive Praktik des Argumentierens eigentlich erfüllt, ist hier überformt vom Übungskontext, sodass das Argumentieren selbst keine „direkte Handlungsfolge“ (Grunder 2015: 49) mit sich bringt.⁴⁹

Resümierend zu den vorgestellten Erwerbskontexten kann festgehalten werden, dass diese mit Morek (2015: 44) als „als komplementäre Diskurssozialisations- und -erwerbskontexte fungieren, die jeweils die Aneignung bereichsspezifischer argumentativer Gattungen ermöglichen“. Dabei befinden sich interne und externe Ressourcen scheinbar in einer wechselseitigen Beziehung miteinander.

Zusammenfassend konnte anhand des vorliegenden Kapitels 4.1 einleitend zentrale Einsichten in die Anforderungen gewonnen werden, die es seitens der Heranwachsenden in mündlichen Argumentationen zu bewältigen gilt. Kapitel 4.1.1 hat dazu Studien vorgestellt, die die schüler*innenseitigen argumentativen Fähigkeiten im Kontext des Unterrichts herauspräparieren und dabei darlegen, wie komplex die diskursive Praktik des Argumentierens ist (vgl. u. a. Entwickeln einer eigenen Position, Strukturierung eigener Beiträge, Bezugnahme auf andere Perspektiven). Im Anschluss wurde dann die Modellierung von Argumentationskompetenz nach Quasthoff (2009) vorgestellt, die es erlaubt, anhand der drei Teilfähigkeiten der Kontextualisierung, Vertextung und Markierung erwerbsorientierten, altersvergleichen- den Fragestellungen nachzugehen. In Kapitel 4.1.2 wurden dann studienbasiert grobe Entwicklungstendenzen nachgezeichnet, die zeigten, dass es heranwachsenden Sprecher*innen mit zunehmendem Erwerb gelingt, kommunikative Aufgaben nicht mehr vornehmlich reaktiv, sondern eigenaktiv autonom zu bearbeiten (Kontextualisierung), dabei die eigenen Äußerungspakete vermehrt global zu strukturieren (Vertextung) und explizit-sprachlich zu markieren (Markierung). Zuletzt wurden unterschiedliche Erwerbskontexte hinsichtlich ihres jeweiligen Erwerbspotenzials für die individuelle Kompetenzausprägung der Heranwachsenden vorgestellt (vgl. Kap. 4.1.3).

Wie sich auch in Kapitel 2 zeigte, beziehen sich die vorgestellten Forschungserkenntnisse (u. a. zur Kompetenzmodellierungen sowie zu den Linien des Diskurserwerbs) zum mündlichen Argumentieren maßgeblich auf die *sprachlich-diskursive* Ebene. Das folgende Kapitel 4.2. wird daher erwerbsbezogene Befunde *multimodaler* Ressourcen vorstellen. Da bisherige Studien allerdings nur vereinzelt Teilaspekte des multimodalen Diskurserwerbs *argumentativer* Fähigkeiten beleuchten, stellt Kapitel 4.2 einleitend Forschungsbefunde zum multimodalen Diskurserwerb des Erzählens und Erklärens vor, für die eine umfangreichere Studienlage vorliegt.

⁴⁹ Vgl. zur Diskussion der Vor- und Nachteile vorbereiteter Argumentationsübungen im Unterricht Grundler (2022: 582ff.).

4.2 Erwerbsbezogene Befunde zur Entwicklung und Funktionen multimodaler Ressourcen in diskursiven Praktiken

Wie in Kapitel 3 bereits ausführlich dargelegt worden ist, entsteht soziale Interaktion durch die Verwendung multimodaler Ressourcen. Diese Erkenntnis berücksichtigen zunehmend auch spracherwerbstheoretische Studien (vgl. u. a. Capirci und Volterra 2008; Rohlfing 2008; Andrén 2010; Morgenstern 2014; Beaupoil-Hourdel et al. 2016; Hübscher und Prieto 2019; Morgenstern und Goldin-Meadow 2022) sowie diskurserwerbstheoretische Studien (vgl. zsf. Guidetti et al. 2014; Colletta 2022), die der Frage nachgehen, welche (kommunikativen) Funktionen multimodale Ressourcen in den jeweiligen Erwerbsprozessen übernehmen. Die vorliegende Arbeit begreift in dieser Linie den Erwerb diskursiver Praktiken als *multimodalen Diskurserwerb* (vgl. Heller 2019), in dem sich sprachliche und körperlich-sichtbare Ressourcen gemeinsam – und über den Spracherwerb hinaus – entwickeln. In der explorativen Fallstudie (vgl. Kap. 10) wird daher untersucht, welche Veränderungen sich u. a. hinsichtlich der Funktion sowie Koordinierung multimodaler Ressourcen mit pragmatischer Funktion in einer Längsschnittstudie (6., 7. und 9. Klasse) bei jugendlichen Sprecher*innen innerhalb des mündlichen Argumentierens mit ihren Eltern zeigen. Da Studien zum multimodalen Diskurserwerb bislang allerdings auf narrative und explanative Diskurspraktiken fokussieren, werden in einem ersten Zugriff einschlägige erwerbsbezogene Forschungsergebnisse zum (Nach-)Erzählen und Erklären dargelegt (vgl. Kap. 4.2.1). Studien, die dezidiert die Entwicklung multimodaler Ressourcen im Diskurserwerb des mündlichen Argumentierens in den Blick nehmen, fallen im Vergleich deutlich schmäler aus und adressieren nur Teilfähigkeiten des Argumentierens – diese werden anschließend zusammenfassend vorgestellt (vgl. Kap. 4.2.2). Hinsichtlich der benannten zweiten Fragestellung, in der die Veränderungen multimodaler Ressourcen sowie ihre Funktionen im Erwerb argumentativer Fähigkeiten mithilfe der Modellierung von Diskurskompetenz nach Quasthoff (2009) untersucht wird, erweist es sich als sinnvoll, die Forschungsergebnisse der Kapitel 4.2.1 und 4.2.2 unmittelbar auch in ihrer Bedeutung für die drei Fähigkeitsdimensionen Kontextualisierung, Vertextung und Markierung einzuordnen (vgl. Quasthoff 2009; vgl. auch Kap. 4.1.1). Das vorliegende Kapitel 4 schließt dann mit einer resümierenden Zusammenfassung sowie der Ableitung erwerbsbezogener Forschungsdesiderate ab (vgl. Kap. 4.3).

4.2.1 (Nach-)Erzählen und Erklären

(Nach-)Erzählen

Erwerbsbezogene Studien zum (Nach-)Erzählen arbeiten häufig mit dem von McNeill (1992: 365ff.) entwickelten Setting, in dem den Proband*innen ein Cartoon (z. B. aus der Zeichentrickserie Tom und Jerry) gezeigt wird, den sie anschließend nacherzählen sollen (vgl. u. a. Colletta et al. 2010; Reig Alamillo et al. 2012; Cristilli 2014; Graziano 2014a). In weiteren Studien werden Kinder aber auch zu Fantasierzählungen (vgl. Ohlhus 2016) oder Erzählungen zu Bilderbüchern (vgl. Heller 2018b, 2019) aufgefordert oder es werden Alltagserzählungen (vgl. Levy und McNeill 2013, 2015) untersucht. Systematisiert man nun in einem ersten Zugriff die Ergebnisse dieser Studien zur multimodalen Entwicklung narrativer Fähigkeiten, so zeigen sich folgende Befunde, die anschließend ausführlicher erläutert werden:

- Jüngere Kinder (19 Monate bis vier Jahre) verwenden v. a. *abbildende Gesten* (vgl. McNeill 1992; Colletta et al. 2010; Cristilli 2014; Heller 2018b, 2019).
- *Gesten mit pragmatischer Funktion* nehmen mit dem Erwerb narrativer Fähigkeiten zu (vgl. McNeill 1992; Colletta et al. 2010; Reig Alamillo et al. 2012; Colletta et al. 2015).
- *Multimodale äusserungsstrukturierende Markierungen* werden vermehrt eingesetzt zur

- Markierung von *Beginn* und *Abschluss* einer Erzählung (vgl. Graziano 2014a, 2014b),
- *Einführung neuer Elemente* (vgl. Reig Alamillo et al. 2012; Graziano 2014a) und
- *Kohäsionsmarkierung* (vgl. Colletta et al. 2010; Reig Alamillo et al. 2012; Graziano 2014a; Cristilli 2014;) bzw. zur globalen *Kohärenzmarkierung* (vgl. Levy und McNeill 2013).
- Multimodale Ressourcen werden vermehrt mit modalisierender Funktion eingesetzt, um die emotionale, mentale aber auch *epistemische Haltung* der Sprecher*innen anzuzeigen (vgl. Colletta et al. 2010; Reig Alamillo et al. 2012; Graziano 2014a, 2014b; vgl. aber Colletta et al. 2015).
- Gestische Ressourcen unterstützen die (noch unsichere) sprachliche *Referenzherstellung* (vgl. Levy und McNeill 2013; Cristilli 2014).

In den zugrundeliegenden Studien zeigt sich insgesamt, dass jüngere Kinder zwischen 19 Monaten und vier Jahren Gesten in (Nach-)Erzählungen verwenden, die vornehmlich abbildende, repräsentationale Funktionen übernehmen (vgl. McNeill 1992; Colletta et al. 2010; Cristilli 2014; Heller 2018b, 2019). Die Sprecher*innen nutzen diese Gesten, um Figuren bzw. Charaktere der Erzählungen zu visualisieren sowie auch, um entsprechende Handlungen und Ereignisse zu modellieren und enactieren (vgl. McNeill 1992; Levy und McNeill 2013, 2015; Heller 2018b, 2019). Dabei lösen sich die jungen Erzähler*innen auf der sprachlichen Ebene nicht von den konkret zu berichtenden Ereignissen: Graziano (2014a) beobachtet beispielsweise, dass sich die Erzählungen von 4-Jährigen, die einen zuvor präsentierten Tom-and-Jerry-Cartoon nacherzählen, noch stark an der Chronologie des Films und der darin enthaltenen Handlungsreihenfolge orientieren. Dies deckt sich mit Befunden zur sprachlich-diskursiven Entwicklung der Erzählfähigkeiten von Kindern, in denen die Sprecher*innen erst mit zunehmendem Alter die Erzählung dann um das erzählwürdige Ereignis herum organisieren können (vgl. Hausendorf und Quasthoff 1996/2005; Quasthoff et al. 2019). Die beobachtbare Orientierung der jüngeren Sprecher*innen an den konkreten Ereignissen scheint eine mögliche Erklärung für die frequente Verwendung abbildender Gesten zu bieten.

Erst mit zunehmenden sprachlich-diskursiven Fähigkeiten, bei denen sich die Sprecher*innen vermehrt von der Chronologie des Gesehenen lösen und Vor- sowie Gleichzeitiges berichten (vgl. Graziano 2014a, 2014b) sowie auch metanarrative Kommentare einbringen können (vgl. Colletta et al. 2010), sind vermehrt Gesten mit pragmatischer Funktion (vgl. Kap. 3.2.2) zu beobachten. Diese tragen zur Vertexung, Markierung sowie Kontextualisierung bei und geben darüber hinaus Auskunft über die epistemische und affektive Haltung eines*einer Sprecher*in, wie nachfolgend dargelegt wird.

Vergleicht man vorab die Veränderungen der Vertexung von Erzählungen auf linguistischer Ebene, so zeigt sich, dass 6-Jährige vergleichsweise kurze Erzählungen mit wenig Konnektoren produzieren, 10-Jährige bereits längere Erzählungen unter Verwendung zunehmender Konnektoren aufbauen können und die Erzählungen erwachsener Sprecher*innen schließlich einen äußerst hohen Grad an syntaktischer Komplexität aufweisen – wenngleich ihre Erzählungen insgesamt kürzer ausfallen und weniger Konnektoren und Anaphern enthalten (vgl. Colletta et al. 2010; Reig Alamillo et al. 2012). Weitet man nun den Blick auf den Einsatz multimodaler Ressourcen zum Aufbau einer Erzählung, so zeigt sich, dass Gesten in pragmatischer Funktion zur Strukturierung des Äußerungspakets beitragen, indem sie *Anfang und Ende* einer Erzählung markieren, *zusammengehörige Abschnitte* einer Erzählung sichtbar machen sowie *neue Elemente* der Erzählung einführen: So können Kinder im Alter von 4 bis 5 Jahren *Palm Up* Gesten mit seitlicher Bewegung einsetzen, um das Ende einer Erzählung zu markieren; ab einem Alter von 6 Jahren führen die Sprecher*innen dann gemeinsam mit Diskursmarkern wie *also* und präsentierenden *Palm Up* Gesten einerseits in die Erzählung ein und nutzen dann die gleiche Geste sowie

eine prosodische Kontur, bei der die Intonation stark fällt, um abschließend das Ende der Erzählung zu markieren (vgl. Graziano 2014b). Auch das Herauslösen sowie Zurückführen von Gesten aus der bzw. in die Ruheposition (vgl. Ohlhus 2016) sowie die Wiederaufnahme von eingangs eingeführten und zum Abschluss wiederholten Gesten (vgl. Heller 2018b) können heranwachsende Sprecher*innen nutzen, um Anfang und Ende einer Erzählung zu markieren. Dabei kann die Wiederaufnahme von Gesten zur Strukturierung einer Erzählung bereits von sehr jungen Kindern (vgl. zur Erzählung eines 19 Monate jungen Kindes Heller 2018b) geleistet werden. Gesten mit pragmatischer Funktion können i. S. der Ver-textung auch zur Gliederung der Erzählung in global-semantische Abschnitte bzw. Strophen (vgl. Ohlhus 2016) beitragen, indem die stetige Wiederholung einer Geste die Zusammengehörigkeit der Abschnittes markiert. Zur Strukturierung von Erzählungen tragen ferner multimodale Ressourcen wie *Palm Up* Gesten (vgl. Graziano 2014a, 2014b) bei, mit denen u. a. Charaktere der (Nach-)Erzählung oder Beschreibungen des Settings eingeführt oder eingeschobene Erklärungen markiert werden (vgl. Graziano 2014b). Allerdings finden sich Gesten in dieser Funktion vermehrt erst bei Kindern im Alter von 8 bzw. 10 Jahren (vgl. Graziano 2014a; 2014b). Auch Übergänge zu neuen Ereignissen werden multimodal von Gesten mit pragmatischer Funktion gemeinsam mit sprachlich explizierten Konnektiven wie *und*, *dann/danach* markiert, wie Reig Alamillo et al. (2012) in Bezug auf schnelle, zur Seite ausgeführte Kopf- oder Handgesten in ihrer altersvergleichenden Studie zeigen. Auch Gesten in dieser diskursstrukturierenden Funktion werden von 6-Jährigen weniger häufig verwendet als von 10-Jährigen (vgl. Reig Alamillo et al. 2012: 18f.).

Auf Ebene der Ver-textung tragen multimodale Ressourcen in mündlichen Erzählungen auch zur *Kohäsionsmarkierung* bei. Insbesondere die Studien von Colletta et al. (2010), Reig Alamillo et al. (2012) und Levy und McNeill (2013) decken auf, dass mit zunehmendem Erwerb narrativer Fähigkeiten die Erzählungen häufiger sprachliche Konnektoren⁵⁰ sowie kohäsionsstiftende Gesten aufweisen.⁵¹ Kohäsionsstiftende Gesten werden in diesen Untersuchungen verstanden als Gestenwiederaufnahmen, sogenannte „catchments“ (Levy und McNeill 2013: 549), mit denen zwischen zwei oder mehr Sätzen Kohäsion hergestellt wird. Aber auch Gesten des abstrakten Zeigens können zur Kohäsion beitragen, denn mit ihnen können Sprecher*innen auf vorher eingeführte Charaktere verweisen. Auch *Beats* werden als Kohäsionsgesten insofern untersucht, als dass sie sprachliche Konnektoren hervorheben können (vgl. Colletta et al. 2010; Reig Alamillo et al. 2012). In Verbindung mit der zunehmenden Fähigkeit, Kohäsionen auch sprachlich zu markieren, nehmen die Autor*innen der altersvergleichenden Studie von Colletta et al. (2010) an, dass der Anstieg diskursstrukturierender, kohäsionsstiftender Gesten von 6- über 10-Jährigen zu Erwachsenen im Zusammenhang mit der generellen Fähigkeit steht, Kohäsion markieren zu können.

Auf der Ebene der Markierung setzen 7- bis 10-jährige Sprecher*innen *Palm Up* Gesten ein, um *Schwierigkeiten in der Äußerungsproduktion*, die sich sprachlich in der Markierung von Längungen sowie Pausen manifestieren, zu überbrücken (vgl. Graziano 2014b: 322). Studienübergreifend wird auf der Ebene der Markierung außerdem der Frage nachgegangen, inwiefern gestische Ressourcen *verstärkend* oder *ergänzend* zur sprachlichen Ressource hinzutreten – inwiefern sich also in der Entwicklung des multimodalen Erzählerwerks auch eine Entwicklung hinsichtlich der Funktion und *Abstimmung* von gesti-

⁵⁰ Die linguistischen Untersuchungen zu sprachlichen Konnektoren im Zusammenhang mit dem Erwerb multimodaler Ressourcen fokussieren auf Konnektoren (vgl. Levy und McNeill 2013), Subordinationen und Anaphern (vgl. Colletta et al. 2010) sowie pronominale und nominale Markierungen (vgl. Cristilli 2014).

⁵¹ Levy und McNeill (2013) sprechen in ihrer Untersuchung von der kohäsionsstiftenden Funktion von Gesten. Dabei schließt die Studie auch Kohärenzmarkierungen mit ein, wenn durch die Wiederaufnahme einer Geste auch zeitlich auseinanderliegende Äußerungsteile als zusammengehörig markiert werden (vgl. hierzu auch Ohlhus 2016).

schen und verbalsprachlichen Ressourcen zeigt. Cristilli (2014) untersucht dazu den Einsatz anaphorischer, d. h. sich wiederholender, repräsentationaler Gesten zur Markierung von Kohäsion. Dabei wird deutlich, dass 6-jährige Kinder die Gesten mehrheitlich in *desambiguierender* sowie *ergänzender* Funktion verwenden, wenn der Referent auf sprachlicher Ebene nicht eindeutig benannt (vgl. Cristilli 2014: 331) bzw. die sprachliche Referenzherstellung (noch) nicht sicher umgesetzt werden kann (vgl. Cristilli 2014: 345). 10-jährige Sprecher*innen verwenden anaphorische Gesten hingegen *integrativ*: Die sprachlich passende Anapher wird durch die anaphorisch eingesetzte Geste expliziter gemacht bzw. verstärkt (vgl. Cristilli 2014: 345).⁵² Cristilli (2014: 345) interpretiert die Befunde so, dass v. a. bei jüngeren Kindern die untersuchten Gesten Schwierigkeiten in der verbalen Referenzherstellung, und somit auf Ebene der Vertextung, ausgleichen können. Interessant erscheint, dass eine eindeutige verbale Referenzherstellung nicht dazu führt, dass die Gesten nicht mehr hervorgebracht werden, sondern dass sich ihre Funktion zu einer modalisierend verstärkenden hin ändert. Die Verwendung multimodaler Ressourcen kann auf der Ebene der Markierung auch Aufschluss über die epistemische sowie die emotionale Haltung geben, in der die Sprecher*innen ihre Erzählung hervorbringen – sie können durch multimodale Ressourcen ihrer eigenen Haltung gegenüber dem Erzählten Ausdruck verleihen können. So verzeichnen beispielsweise Colletta et al. (2010) sowohl hinsichtlich sprachlicher als auch gestischer Markierungen einen deutlichen Anstieg metanarrativer Kommentare sowie emotionaler Bewertungen in Erzählungen von 6- und 10-Jährigen bis hin zu erwachsenen Sprecher*innen. Dabei können multimodale Ressourcen in Form von *Framing Gestures*, wie beispielsweise faziale Gesten, das sprecher*innenseitige Amüsement bezüglich der nacherzählten Sequenz anzeigen (vgl. Colletta et al. 2010). Neben emotionalen Bewertungen und Kommentaren können Sprecher*innen allerdings auch ihre epistemische Haltung verdeutlichen und somit zur epistemischen Modalisierung der Äußerung beitragen: So legt Graziano (2014a, 2014b) in ihren Untersuchungen dar, dass erst 8-Jährige beginnen, Inhalte ihrer Äußerung durch *Palm Up* Gesten mit seitlicher Bewegung und prosodische Markierungen als offensichtlich bzw. evident zu markieren. Die Erzählungen jüngerer Kinder weisen diese Formen der Modalisierung hingegen noch nicht auf (vgl. zum Erwerb der Verwendung von Modalisierungen auch Hübscher et al. 2019).

Neben den Funktionen gestischer Ressourcen auf den Ebenen der Vertextung und Markierung können multimodale Ressourcen Sprecher*innen aber auch darin unterstützen, ihre Erzählung sequenziell im gerade stattfindenden, übergeordneten *turn-by-turn talk* zu platzieren und somit zur Kontextualisierung beitragen (vgl. zur Einbettung von größeren Diskurseinheiten in den *turn-by-turn talk* Quasthoff et al. 2017; vgl. Kap. 4.1.1). So zeigt Morek (2018), wie jugendliche Sprecher*innen, die *Schwierigkeiten* in der *sequenziellen Einpassung*, d. h. der Kontextualisierung einer Erzählung in einer Peer-Interaktion haben, zur Absicherung ihres *Turns* auf aufmerksamkeitsichernde bzw. -erheischende Ressourcen wie den erhobenen Zeigefinger oder das Klatschen in die Hände zurückgreifen. Die eingesetzten multimodalen Ressourcen fungieren hier als Überbrückung von Kontextualisierungsschwierigkeiten und nehmen in der Konsequenz zugleich auch Einfluss auf den Sprecher*innenwechsel. Multimodale Ressourcen können darüber hinaus zur Kontextualisierung der *Sprechhandlung* eingesetzt werden: So nutzen heranwachsende Sprecher*innen ab einem Alter von 6 bis 7 Jahren *Palm Up* Gesten, um Anteile ihrer Erzählung als ‚präsentiert‘ zu kontextualisieren (vgl. Graziano 2014b).

Im Anschluss an die hier vorgelegten erwerbsbezogenen Befunde zum multimodalen Erzählerwerb werden nachfolgend Studienergebnisse zum multimodalen Erklärerwerb vorgestellt.

⁵² Deiktische Gesten werden in diesem Setting wenig verwendet, sodass die Autorin die Auswertung ihrer Befunde auf die anaphorisch verwendeten repräsentationalen Gesten bezieht (vgl. Cristilli 2014: 344).

Erklären

Für die diskursive Praktik des Erklärens liegen zum jetzigen Zeitpunkt, im Vergleich zur Diskurspraktik des (Nach-)Erzählens, verhältnismäßig wenige Studien vor, die die Entwicklung sprachlich-diskursiver und multimodaler Ressourcen beleuchten (vgl. Kern 2003, 2011; Colletta und Pellenq 2009; Reig Alamillo et al. 2012). Es handelt sich um Studien, die prozessuale (vgl. Kern 2003, 2011) sowie kausale Erklärungen (vgl. Colletta und Pellenq 2009; Reig Alamillo et al. 2012) altersvergleichend in den Blick nehmen.⁵³ Es zeigen sich folgende Befunde innerhalb der Entwicklung explanativer Fähigkeiten, die anschließend näher erläutert werden:

- Mit zunehmender *Länge und Komplexität* der Erklärungen steigen auch die eingesetzten multimodalen bzw. gestischen Ressourcen (vgl. Colletta und Pellenq 2009; Reig Alamillo et al. 2012).
- Jüngere Kinder greifen bei Erklärungen deutlich häufiger auf Gesten mit *veranschaulichender* Funktion zurück (vgl. Kern 2003, 2011; Colletta und Pellenq 2009; Reig Alamillo et al. 2012).
- Mit zunehmendem Alter können Kinder auf (eingeführte) *abstrakte* Ideen und Konzepte gestisch *referenzieren*, beispielsweise durch abstrakte Zeigegesten (vgl. Colletta und Pellenq 2009)
- Multimodale Ressourcen tragen mit zunehmendem Erwerb zur *Strukturierung* von Erklärungen bei:
 - Die Zahl der verwendeten *pragmatischen und äußerungsstrukturierenden* Gesten steigt an (vgl. Colletta und Pellenq 2009; Reig Alamillo et al. 2012).
 - Vermittels pragmatischer Gesten können *neue Referenten* eingeführt werden (vgl. Colletta und Pellenq 2009).
 - Die Verwendung *multimodaler Kohäsionsmarkierungen* nimmt zu (vgl. Colletta und Pellenq 2009; Reig Alamillo et al. 2012).

Da sich die Befunde insgesamt auf die äußerungsstrukturierenden Funktionen multimodaler Ressourcen für das Erklären beziehen, erfolgt die Darstellung der Forschungsergebnisse entlang der Ebene der Ver-textung.

Erklärungen, wie auch Erzählungen, werden von einem*einer primären Sprecher*in in Form eines satz-übergreifenden Äußerungspakets hervorgebracht. Im Gespräch werden sie in der Regel dann relevant, wenn sich eine Wissensasymmetrie zwischen Sprecher*innen abzeichnet, deren Behebung wichtig für den Fortgang der Interaktion ist (vgl. Morek 2012; Kinalzik 2024). Die vorliegenden Studien (vgl. Colletta et al. 2010; Reig Alamillo et al. 2012) verweisen übergreifend darauf, dass mit zunehmendem Erwerb explanativer Fähigkeiten die Länge und Komplexität der Erklärungen, die Anzahl der verwendeten sprachlichen Kohäsionsmarkierungen (in Form logischer oder chronologischer Verbindungen (vgl. Colletta und Pellenq 2009: 71f.) sowie die eingesetzten multimodalen Ressourcen ansteigen. Darüber hinaus zeigt sich auch eine Veränderung der Gestenfunktionen: So scheinen jüngere Kinder vornehmlich auf abbildende Gesten zur Veranschaulichung der Erklärung zurückzugreifen, wie u. a. Kern (2003,

⁵³ Dass und inwiefern innerhalb eines Erklärtyps (vgl. zsf. Klein 2009) das jeweilige Explanandum Einfluss auf die eingesetzten Ressourcen nehmen kann, zeigen Mazur-Palandre et al. (2014) in ihrer Studie zu Wie-Erklärungen, in denen Sechsjährige eine Prozesserklärung (*Wie ist XY passiert?*) zu einem Cartoon gegenüber einem Erwachsenen sowie eine Spielinstruktionen für ein zweites Kind (*Wie soll Kind 2 im Spiel weitermachen?*) hervorbringen sollen. Letztere produzieren keine Spielerklärungen, *bevor* ein Spiel gemeinsam aufgenommen wird, sondern es handelt sich um Instruktionen, die *während* des Spiel eines zweiten Kindes fortgesetzt werden, sodass die Erklärenden immer auch die tatsächliche Umsetzung ihrer Instruktionen beobachten können und ggf. *ad hoc* darauf reagieren müssen. Dabei zeigt sich, dass die Erklärenden in den Prozessklärungen deutlich mehr abbildende und in den Instruktionen deutlich mehr pragmatische Gesten, wie performative, interaktive, diskursstrukturierende rahmende (*framing*) Gesten, äußern (vgl. auch Mazur-Palandre und Lund 2016).

2011) und Stude (2003) anhand von Spielerklärungen nachweisen können. Ältere Kinder können dann ihre Erklärungen auch an abstrakten Positionen des Spiels, z. B. Spielregeln, orientieren, in denen Gesten mit abbildender Funktion dann nicht mehr regelhaft eingesetzt werden.⁵⁴

In den Warum-Erklärungen, die Colletta und Pellenq (2009) erhoben haben, wird sichtbar, dass 3- bis 6-Jährige vor allem abbildende Gesten für *konkrete* Referenten verwenden, wohingegen Kinder im Alter 6 bis 11 Jahren zunehmend auf metaphorische und abstrakt repräsentationale Gesten zurückgreifen, um *abstrakte* Konzepte oder Ideen zu visualisieren. So kann beispielsweise das Schließen der Hand zu einer Faust das Konzept, etwas zu bewahren, darstellen (vgl. Colletta und Pellenq 2009: 80). Insgesamt zeigt sich dann in der letztgenannten Alterskohorte, dass sie im Unterschied zu den 3- bis 6-Jährigen vermehrt auch Gesten mit pragmatischer Funktion verwenden, die dann zur Strukturierung des Äußerungspakets beitragen (vgl. Colletta und Pellenq 2009; Reig Alamillo et al. 2012).

Gesten mit pragmatischer Funktion,⁵⁵ die von älteren Sprecher*innen (im Alter von 6 bis 11 Jahren) in der Studie von Colletta et al. (2010) verwendet werden, können insofern zur Strukturierung und somit zur Vertextung der Erklärung beitragen, als dass durch diese neue Inhaltselemente eingeführt werden. So weisen Colletta und Pellenq (2009) sowie auch Reig Alamillo et al. (2012) in ihren Daten nach, dass Sprecher*innen schnelle Kopfgesten zur Seite sowie *Palm Up* Gesten einsetzen, während sie gleichzeitig sprachlich einen neuen, abstrakten Referenten in die Erklärung einführen. Im weiteren Verlauf des Aufbaus der Erklärung können Sprecher*innen sich dann auf die bereits etablierte Geste zurückbeziehen, indem sie durch eine abstrakte deiktische Geste auf diese verweisen – und somit multimodal Kohäsion markieren (vgl. Colletta und Pellenq 2009; Reig Alamillo et al. 2012). Ferner können Sprecher*innen durch die Verwendung von Zählgesten eine (entstehende) Liste verdeutlichen (vgl. Colletta und Pellenq 2009) und somit zur Strukturierung der Äußerung beitragen. Colletta und Pellenq (2009) zeigen, dass dabei vermehrt 11-Jährige kohäsionsmarkierende Gesten einsetzen, während sich die Gesten zur Einführung eines abstrakten Referenten sowie kohäsionsmarkierende Gesten bei 3- bis 6-Jährigen als „quasi-nonexistent“ (Colletta und Pellenq 2009: 87) erweisen. Ähnliche Befunde weist auch die Studie von Reig Alamillo et al. (2012) auf, in der die Autor*innen insgesamt mehr äußerungsstrukturierende (wie z. B. *Beat* Gesten, die Äußerungselemente besonders hervorheben) bzw. kohäsionsmarkierende Gesten (wie z. B. schnelle Kopf- oder Handbewegungen zu einer Seite, die gemeinsam mit sprachlichen Konnektoren geäußert werden) bei den 11-Jährigen als den 6-Jährigen Sprecher*innen nachweisen können.⁵⁶

Zusammenfassend verweisen die Befunde der vorgestellten Studien darauf, dass sich der Erwerb narrativer sowie explanativer Fähigkeiten nicht nur sprachlich-diskursiv vollzieht, sondern multimodale Ressourcen insgesamt ausdrücklich in den Kompetenzerwerb involviert sind. So zeigt sich gattungsübergreifend, dass jüngere Sprecher*innen, wenn sie sich in Narrationen an der Chronologie eines Ereignisses oder in Spielerklärungen an der Chronologie des Spielablaufs orientieren, vornehmlich gestische

⁵⁴ Kern (2003) spricht hier von einem demonstrierenden bzw. explizierenden Erklärstil. Diese Präferenzen, eine Spielerklärung eher durch die Benennung konkreter Spielhandlungen (demonstrierend) oder stärker durch das Benennen von Regeln (explizierend) durchzuführen, scheinen sich allerdings innerhalb des Diskurserwerbs eines Kindes zu verändern, wie altersvergleichende Untersuchungen zeigen (vgl. Stude 2003; Kern 2011): Zum Zeitpunkt der Einschulung sind die Präferenzen für einen demonstrierenden bzw. explizierenden Erklärstil noch ausgeglichen; ein Jahr später jedoch nutzen die Kinder fast ausschließlich den explizierenden Stil in der Realisierung der Erklärung.

⁵⁵ Colletta und Pellenq (2009) sprechen hier unter Rückbezug auf McNeill von *metaphorischen* Gesten. Da diese aber in diskursiver bzw. diskursstrukturierender Funktion eingesetzt werden, werden sie in der vorliegenden Arbeit Gesten mit pragmatischer Funktion (vgl. dazu auch Kap. 3) zugeordnet.

⁵⁶ Reig Alamillo et al. (2012) weisen einen ähnlichen Anstieg für Gesten, die Auskunft über die Haltung eines/einer Sprecher*in innerhalb des Erklärens geben, nach. Aus der Studie geht allerdings nicht hervor, welche Funktion diese Gesten (*Framing Gestures*) im Rahmen der Erklärungen übernehmen, sodass sie nicht in die o. g. Ausführungen aufgenommen worden sind.

Formen der Visualisierung, d. h. Gesten mit abbildender Funktion einsetzen. Mit zunehmender Entwicklung sprachlich-diskursiver Fähigkeiten werden dann umfassende Veränderungen in der Vertextung der jeweiligen kommunikativen Gattung sichtbar: Sprecher*innen beginnen multimodale Ressourcen zu nutzen, um ihre komplexer werdenden Äußerungspakete kohäsiv und kohärent zu strukturieren, in das Äußerungspaket einzuleiten sowie dessen Abschluss zu markieren, neue Referenten einzuführen etc. Dabei scheint sich das Verhältnis der sprachlichen und gestischen Ressourcen mit zunehmendem Alter auf Ebene der Markierung insofern zu verändern, als dass gestische Ressourcen in der Referenzherstellung die sprachliche Äußerung nicht weiter desambiguieren, sondern die dafür eingesetzten Anaphern verstärkend hervorheben können. Sprecher*innen können darüber hinaus auch ihre epistemische oder emotionale Haltung gegenüber den Inhalten des aufgebauten Äußerungspakets durch multimodale Ressourcen anzeigen. Auf Ebene der Kontextualisierung können multimodale Ressourcen mit zunehmendem Diskurserwerb außerdem genutzt werden, um Sprechhandlungen sichtbar zu machen sowie ein Äußerungspaket sequenziell zu platzieren.

Insgesamt liefern die Befunde der vorgestellten Studien deutliche Hinweise darauf, dass sich der Erwerb sprachlich-diskursiver Fähigkeiten gemeinsam mit der Verwendung körperlich-sichtbarer Ressourcen entwickelt. Die Bedeutung der dargelegten erwerbsbezogenen Befunde zur Entwicklung und Funktion der diskursiven Praktik des Erzählens und Erklärens konnte auch hinsichtlich der drei Ebenen von Diskurskompetenz (Kontextualisierung, Vertextung und Markierung, vgl. Kap. 4.1.1) eingeordnet werden. Es zeigen sich dabei (je nach diskursiver Gattung) alters- bzw. erwerbsbedingte Veränderungen in den drei Teilfähigkeiten von Diskurskompetenz.

Da nachfolgende Kapitel 4.2.2 wird nun bisherige Forschungserkenntnisse zu Teilaspekten des multimodalen Diskurserwerbs *argumentativer* Fähigkeiten beleuchten.

4.2.2 Teilaspekte des Argumentierens

Aus einzelnen Fallstudien, die das Argumentieren unter Peers untersuchen, geht hervor, dass jüngere, argumentationsunerfahrene(re) Kinder auch körperlich-sichtbare Ressourcen nutzen, um Opposition anzuzeigen (vgl. dazu bereits Maynard 1985)⁵⁷, z. B. indem Körperposituren eingefroren (vgl. Hannken-Illjes und Bose 2019) oder sprachliche Handlungen körperlich durch das Zuhalten des Mundes unterbunden werden (vgl. Kraft und Lanzen 1995).⁵⁸ Trotz der hier nur exemplarisch benannten Befunde und obwohl längsschnittliche Untersuchungen zum sprachlich-diskursiven Erwerb des mündlichen Argumentierens vorliegen (vgl. Kap. 4.1), ist der *multimodale* Argumentationserwerb bislang noch nicht *systematisch* untersucht worden. Lediglich zu argumentativen *Teilfähigkeiten* existieren längsschnittliche Studien. Aufgrund der äußerst begrenzten Studienlage können im Folgenden nur zwei Studien referiert werden: eine spracherwerbsbezogene Studie, die den Erwerb des Negierens untersucht (vgl. Beaupoil-Hourdel et al. 2016), das in mündlichen Argumentationen gerade für das Herstellen eines Dissens oder

⁵⁷ Auch Maynard (1985) untersucht, inwiefern junge Sprecher*innen im Alter der ersten Klasse Dissens etablieren und aushandeln. Körperliche Handlungen, wie das Weglegen eines Buntstiftes während des gemeinsamen Malens, können beispielsweise als sogenannte „arguable moves“ (Maynard 1985: 3) bzw. Auslösehandlungen Widersprüche initiieren.

⁵⁸ Die Autorinnen legen zwar dar, dass bis zum Alter des Schuleintritts diese Form oder „Technik“ des (körperlichen) Opponierens (Kraft und Lanzen 1995: 95) abnimmt und die Sprecher*innen zunehmend beginnen, auch begründete Widersprüche etc. zu produzieren, allerdings handelt es sich um eine zusammenfassende Ergebnispräsentation, aus der für die vorliegende Arbeit keine weiteren Aussagen zum multimodalen Diskurserwerb abgeleitet werden können.

in opponierenden Handlungen relevant wird (vgl. Kap. 2.2.2); sowie eine Studie zum Erwerb der pragmatischen *Slapping* Geste, mit der Interagierende die Handlungen des Gegenübers unterbrechen und somit Opposition oder Protest markieren können (vgl. Ladewig und Hotze 2023).

Beaupoil-Hourdel et al. (2016) untersuchen innerhalb einer Longitudinalstudie aus Perspektive des Spracherwerbs die multimodale Entwicklung von Negationen bzw. die Markierung von Nicht-Wissen eines englischsprachigen Kindes im Alter von 10 Monaten bis zu 4 Jahren. Vor Einsetzen des Gebrauchs verbaler Ressourcen werden Negationen in Form von Ablehnungen eines erwachsenenseitigen Vorschlags oder epistemischer Unsicherheit bezüglich erwachsenenseitiger Fragen durch körperlich-sichtbare Ressourcen ausgedrückt, indem z. B. der Blickkontakt unterbrochen oder die Handlungen des Erwachsenen wie das Füttern abgelehnt (vgl. Beaupoil-Hourdel et al. 2016: 110f.; vgl. jüngst auch Klatt 2024) und Objekte gestisch weggedrückt werden (vgl. Beaupoil-Hourdel et al. 2016: 111). Mit dem Einsetzen verbaler Ressourcen sowie der Verwendungsmöglichkeit der Negationspartikel *no* zeigt sich nun, dass die sprachlichen und körperlich-sichtbaren Ressourcen kombiniert werden können. Zur Verneinung schüttelt das Kind beispielsweise den Kopf (vgl. Beaupoil-Hourdel et al. 2016: 113) oder dreht im Kontext epistemischer Negationen, i. S. von *ich weiß nicht*, beide Handflächen mit einem Schulterzucken nach oben (vgl. zu *Shrugs* auch Debras 2017). Insgesamt zeigt sich innerhalb des Negationserwerbs eine Ablösung von rein körperlichen Handlungen hin zur Ausbildung eines abstrakten Zeichensystems, mithilfe dessen Heranwachsende Negationen ausdrücken können. Dabei werden die Gesten im weiteren Erwerbsverlauf dann nicht mehr ausschließlich redundant zur sprachlichen Äußerung eingesetzt, sondern „they are used to enrich the verbal utterance of the child“ (Beaupoil-Hourdel et al. 2016: 117). Da die Studie allerdings spracherwerbstheoretische Fragen adressiert und ermittelt, wie sich Formen des Negierens von praktischen Handlungen zur Ausbildung eines abstrakten Zeichensystems entwickeln, liegen diese Befunde in einer Phase der Entwicklung, die dem Erwerb sprachlich-diskursiver Fähigkeiten, die die vorliegende Arbeit fokussiert, vorausläuft.

Ladewig und Hotze (2023) hingegen adressieren eine ältere Alterskohorte von 4- bis 6-Jährigen und untersuchen hier die Entwicklung von Gesten mit pragmatischer Funktion. Der Fokus liegt dabei auf der Geste des *Slapping* bzw. Schlagens, die die Autorinnen in Alltagsinteraktionen deutschsprachiger Kinder im Vorschulalter beobachten (vgl. Ladewig und Hotze 2023: 98). Dabei zeigt sich, dass das Schlagen in argumentativen Sequenzen eingesetzt werden kann, um u. a. die Handlungen anderer zu beenden (vgl. u. a. Ladewig und Hotze 2023: 99ff.). Anhand der von den Autorinnen angeführten Beispiele kann angenommen werden, dass die Geste zum einen als Mittel eingesetzt wird, um einen Dissens herzustellen bzw. ein Verhalten eines anderen Kindes zu problematisieren, sowie zum anderen in Begründungssequenzen als „expressive function of protest“ (vgl. Ladewig und Hotze 2023: 104). Es zeigt sich dann im Erwerb, dass ältere Interagierende das Schlagen nicht mehr als konkrete *Handlung* (Ladewig und Hotze 2023: 107), sondern nun mit einem deutlich höheren Abstraktionsgrad (vgl. Ladewig und Hotze 2023: 111) als *Geste* ausführen, die eine „as-if quality“ (Ladewig und Hotze 2023: 101) besitzt. Aufgrund der Gestenform, -bewegung und -funktion nehmen die Autorinnen an, dass die *Slapping* Geste einen Vorläufer der noch weiter schematisierten *Throwing Away* Geste bildet (vgl. Ladewig und Hotze 2023: 97), mit der Sprecher*innen Themen aus der weiteren Interaktion ausschließen können. Ihre Bewegung wird dabei noch weiter als bei der *Slapping* Geste reduziert; und sie wird dann v. a. von erwachsenen Sprecher*innen verwendet (vgl. auch Bressemer und Müller 2014a). Mit dieser Untersuchung liefern die Autorinnen wichtige Einsichten zu der Frage, wie sich bei heranwachsenden Sprecher*innen der Erwerb von Gesten mit pragmatischer Funktion (aus konkreten Handlungen heraus) vollziehen kann.

Insgesamt ermöglichen beide Studien erste Einblicke in den Erwerb von Gesten mit pragmatischer Funktion. Dabei konzentrieren sich Beaupoil-Hourdel et al. (2016) auf spracherwerbstheoretische Fragen und setzen hinsichtlich der zweiten Fragestellung der vorliegenden Arbeit vor der Entwicklung sprachlich-diskursiver Fähigkeiten an. Die Untersuchung von Ladewig und Hotze (2023) fokussiert zwar auf ältere Kinder und hier auf den Erwerb von Gesten mit pragmatischer Funktion, allerdings ordnen sie ihre Befunde, trotz der deutlich argumentativen Rahmung in den Daten, nicht hinsichtlich ihrer Funktion für den Argumentationserwerb ein. Eine umfassende Untersuchung der Entwicklung sprachlich-diskursiver Fähigkeiten unter Berücksichtigung multimodaler Ressourcen steht bislang noch aus.

4.3 Zusammenfassung und Ausblick: Gesten als Teil multimodaler Gestalten beim Argumentieren und im Diskurserwerb

In Vorbereitung auf die Untersuchung von Gesten mit pragmatischen Funktionen innerhalb eines *multimodalen Argumentationserwerbs* (vgl. Kap. 10) wurden in dem vorliegenden Kapitel 4 zwei relevante Forschungsstränge zusammengeführt: Unter Berücksichtigung qualitativer Studien zur Ermittlung der Argumentationskompetenz von Heranwachsenden wurde eine Modellierung zur Erfassung von Diskurs- bzw. Argumentationskompetenz (vgl. u. a. Quasthoff 2009; Morek 2017) vorgestellt, mit der auch Erwerbsphänomene untersucht werden können. Dabei ermöglicht diese Modellierung, sowohl die sequenzielle Passung (Kontextualisierung) als auch die interne Strukturierung der Äußerungen (Vertexung) sowie die dafür eingesetzten Ressourcen (Markierung) auch altersvergleichend zu untersuchen (vgl. Kap. 4.1.2). Bislang fokussiert das Modell allerdings vornehmlich auf die Ebene der Sprache. Mit der dann anschließenden Rezeption des Forschungsstandes (vgl. Kap. 4.2) zur Entwicklung multimodaler narrativer und explanativer Diskurskompetenz konnten auch Befunde in die Teilfähigkeiten der Kontextualisierung, Vertexung und Markierung einbezogen werden, die sich auf erwerbsbedingte Veränderungen *multimodaler*, aber v. a. *gestischer* Ressourcen beziehen. Bereits hier zeigt sich, dass eine Integration multimodaler Ressourcen in das Modell von Diskurskompetenz fruchtbar ist und zudem notwendig erscheint, wenn Erwerbsphänomene diskursiver Praktiken umfassend nachvollzogen werden sollen. Wendet man den Blick nun hin zu den Befunden der Studien in Kapitel 4.2.1, so wird hinsichtlich der Entwicklung und Funktion v. a. gestischer Ressourcen eine globale Tendenz deutlich: Für das (Nach-)Erzählen und Erklären scheinen jüngere Kinder frequent Gesten mit abbildender Funktion zu nutzen; erst mit höherem Alter verwenden sie vermehrt auch pragmatische Gesten, die v. a. zur Vertexung eingesetzt werden. Ableitend aus diesen Studienergebnissen zu den diskursiven Praktiken des (Nach-) Erzählens und Erklärens kann angenommen werden, dass heranwachsende Sprecher*innen die Verwendung von Gesten mit pragmatischer Funktion innerhalb des Diskurserwerbs erst erlernen.

Erwerbstheoretische Studien zur Entwicklung und Funktion multimodaler Ressourcen beim *Argumentieren* erweisen sich hingegen als rar. So beleuchten wenige Untersuchungen (vgl. Beaupoil-Hourdel et al. 2016; Ladewig und Hotze 2023) die multimodale Entwicklung *einzelner Teilfähigkeiten* und zeichnen nach, wie Interagierende, mit zunehmender Abstraktion, Negation und Ablehnung ausdrücken können (vgl. Kap. 4.2.2) – eine Rückbindung dieser beobachteten Entwicklung an die diskursive Praktik des Argumentierens liegt hier allerdings nicht vor.

Insgesamt offenbart sich an dieser Stelle eine Forschungslücke hinsichtlich der Untersuchung eines *multimodalen Argumentationserwerbs*. Dieses Desiderat adressierend wird die vorliegende Arbeit im zweiten Teil der empirischen Studie (vgl. Kap. 10) anhand einer Fallstudie explorativ untersuchen, wie sich der Einsatz und die Funktionen multimodaler Ressourcen innerhalb des Erwerbs argumentativer Fähig-

keiten verändern. Dazu wird das Modell von Argumentationskompetenz (vgl. Kap. 4.1.1) um multimodale Ressourcen erweitert, um dann auf Grundlage längsschnittlich erhobener Eltern-Kind-Argumentationen, in denen die Fokuskinder 12, 13 bzw. 15 Jahre alt sind, anhand qualitativer längsschnittlicher Untersuchungen die Veränderungen auf Ebene der Kontextualisierung, Vertextung und Markierung zu ermitteln – die multimodal realisiert werden.

5 Forschungsdesiderate und Fragestellungen

Aus den vorangegangenen Kapiteln (vgl. Kap. 2, 3 & 4) werden nun die zentralen Forschungsdesiderate abgeleitet und in forschungsleitenden Fragen zusammengefasst. Als strukturierende Begleitung werden die Fragestellungen in den nachfolgenden empirischen Kapiteln (vgl. Kap. 8 & 10) wiederkehrend benannt und in den Analysen adressiert, bevor die Befunde abschließend zusammenfassend eingeordnet werden (vgl. Kap. 11). Da sich das Forschungsvorhaben der vorliegenden Arbeit aus zwei Schwerpunkten (*multimodale Ressourcen in der kommunikativen Gattung des mündlichen Argumentierens* und *multimodale Ressourcen im Erwerb der Diskurskompetenz*) zusammensetzt, wird die Vorstellung der Desiderate sowie die daraus abgeleiteten Fragestellungen dieser Zweiteilung folgen.

Forschungsdesiderate & Forschungsfrage I: Multimodale Ressourcen beim mündlichen Argumentieren

Anhand der referierten empirischen (gesprächsanalytischen) Studien zum mündlichen Argumentieren (vgl. Kap. 2) konnte gezeigt werden, dass dieses in Alltagsinteraktionen anders verläuft, als es lange von der theoretischen Argumentationsforschung idealiter konzeptualisiert worden ist. Bisherige gesprächsanalytische Studien zum mündlichen Argumentieren konnten bereits sowohl die *globale Struktur* des Argumentierens und die dabei für die Beteiligten anfallenden Gesprächsaufgaben (vgl. Heller 2012; Kap. 2.2) als auch die zur Bearbeitung der Jobs eingesetzten pragmatischen Mittel bzw. argumentativen Handlungen und sprachlichen wie parasprachlichen Formen, die die einzelnen Sprecher*innen verwenden, rekonstruieren. Allerdings fokussieren diese bislang primär auf die linguistische Untersuchung der *sprachlichen* Ressourcen (vgl. Kap. 2). Demgegenüber stellen Forschungsarbeiten, die sich mit Funktionen körperlich-sichtbarer sowie umgebender (*environmental*) und materieller Ressourcen befassen, keinen systematischen Gattungsbezug zum Argumentieren her (vgl. Kap. 3). Erst eine Handvoll Studien führt diese beiden Forschungsstränge zusammen und zeigt, dass und inwiefern körperlich-sichtbaren Ressourcen in mündlichen Argumentationen eine relevante Rolle zukommt. In diesen Studien zeichnet sich u. a. ab, dass die Interagierenden sich auch dann, wenn während des Argumentierens „the collaborative continuation of talk is potentially at risk“ (Schönfelder und Heller 2019: 86), durch körperlich-sichtbare Ressourcen die weitere gemeinsame Arbeit an der kommunikativen Aufgabe anzeigen. So können sie einander z. B. durch gestische Ressourcen zur Beteiligung an der gemeinsamen Aushandlung einladen (vgl. Kap. 3). Eine systematische Untersuchung, welche Funktionen im Spezifischen *Gesten mit pragmatischer Funktion* für das mündliche Argumentieren in der Realisierung argumentativer Handlungen (vgl. Kap. 2) übernehmen, steht m. W. bislang noch aus. Die hier sichtbar werdenden Forschungsdesiderate adressiert das vorliegende Forschungsprojekt. Es ergibt sich die folgende Forschungsfrage, die in Kapitel 8 untersucht wird.

Welche Funktionen übernehmen multimodale Ressourcen für das mündliche Argumentieren?

Der Fokus der Studie liegt auf der Untersuchung von Gesten mit pragmatischen Funktionen in ihren jeweiligen Ressourcenverbünden, die u. a. gemeinsam mit dem Blick multimodale Gestalten (vgl. Mondada 2014b, vgl. Kap. 3) aufbauen und wiederkehrend an der Realisierung argumentativer Handlungen beteiligt sind (vgl. Kap. 8).

Forschungsdesiderate und Forschungsfrage II: Multimodale Ressourcen im Erwerb argumentativer Fähigkeiten

Die vorgestellten gesprächsanalytischen Modellierungen von Diskurs- bzw. Argumentationskompetenz zeigen ebenfalls eine vornehmliche Orientierung an den *sprachlichen* Ressourcen auf (vgl. Kap. 4.1). Auch das dieser Arbeit zugrunde liegende Modell von Diskurskompetenz nach Quasthoff (2009), das

die sequenzielle Passung (Kontextualisierung) sowie die Strukturierung (Vertextung) der Äußerung einzelner Beteiligter mittels sprachlicher sowie parasprachlicher Formen (Markierung) ermöglicht zu untersuchen, ist auf die Untersuchung *sprachlicher* Ressourcen ausgerichtet. Erst vereinzelt nehmen Studien in der Beschreibung argumentativer Fähigkeiten von Heranwachsenden auch Multimodalität in den Blick (vgl. u. a. Domenech et al. 2018), wenngleich eine gleichzeitige *erwerbsorientierte* Perspektive, i. S. der Veränderungen der Kompetenz, selten mit adressiert wird (vgl. aber Heller 2018a).

Studien, die auf die Entwicklung multimodaler Ressourcen fokussieren, untersuchen bislang ausschließlich narrative und explanative Gattungen. Zusammenfassend zeigen sie, dass jüngere Kinder sowohl während des Erzählens als auch des Erklärens v. a. abbildende Gesten verwenden. Erst mit zunehmendem Alter, in dem auch sprachlich-diskursive Veränderungen sichtbar werden, nutzen die Heranwachsenden vermehrt Gesten mit pragmatischer Funktion, die zur internen Strukturierung der Äußerungspakete (i. S. der Vertextung), zur Modalisierung (i. S. der epistemischen oder auch emotionalen Haltung eines*einer Sprecher*in) sowie auch zur sequenziellen Platzierung (i. S. der Kontextualisierung) beitragen (vgl. Kap. 4.2). Die Untersuchungen adressieren dabei Kinder im Alter bis zu 11 Jahren.

Vergleichbare Untersuchungen liegen m. W. für das mündliche Argumentieren derzeit noch nicht vor. Nur vereinzelte Studien geben Hinweise darauf, dass sich die Funktionen eingesetzter Ressourcen mit zunehmendem Alter verändern: So entwickeln sich beispielsweise im Alter von 4 bis 6 Jahren aus vormals praktisch vollzogenen Handlungen wie dem *slapping* (vgl. Ladewig und Hotze 2023) Gesten mit abstrakter Bedeutung in pragmatischer Funktion, mit der die Sprecher*innen Ablehnung anzeigen können (vgl. Kap. 4.2.2). Mit der Untersuchung von Colletta und Pellenq (2009) zeigt sich ferner, dass multimodalen Ressourcen mit pragmatischer Funktion in mündlichen Argumentationen eine relevante Rolle zukommt. Ihre gattungsvergleichende Übersicht zum Vorkommen von Gesten innerhalb von Erklärungen, Erzählungen, Beschreibungen und Debatten, die eine Nähe zum Argumentieren aufweisen, legt dar, dass in sogenannten debattierenden Sequenzen Gesten mit pragmatischer Funktion häufiger vorkommen als beispielsweise abbildende Gesten und doppelt so häufig wie in Erklärungen oder Erzählungen (vgl. Kap. 4.2). Aufgrund des Fokus ihrer Studie auf andere diskursive Gattungen wird dieser Befund von den Autor*innen allerdings nicht näher beleuchtet. Insgesamt ergeben sich nun zwei Forschungsdesiderate: Es wird angenommen, dass gerade Gesten mit pragmatischen Funktionen für das mündliche Argumentieren relevant werden. Allerdings adressieren bisherige Erwerbsstudien diese Gattung bislang nicht. Ferner zeigt sich, dass die referierten Studien zum Erklären und (Nach-)Erzählen die Entwicklung querschnittlich bis zu einem Alter von 11 Jahren untersuchen. Da aber aus Studien zum argumentativen Diskurserwerb bekannt ist, dass auch Jugendliche (vgl. Kap. 4.1.3) ihre sprachlich-diskursiven Fähigkeiten weiter ausbauen, erweist es sich als notwendig, diese Alterskohorte mit zu adressieren. Aus diesen Desideraten ableitend, nimmt die vorliegende Studie in Kapitel 10 die nachfolgende erwerbsbezogene Forschungsfrage in den Blick.

Welche Veränderungen zeigen sich in einer längsschnittlich angelegten Studie zweier Fokusinder im Alter von 12, 13 und 15 Jahren hinsichtlich der sprachlich-diskursiven Fähigkeiten sowie des Einsatzes multimodaler Ressourcen?

Hinsichtlich des erwerbsorientierten Forschungsinteresses basiert die zweite empirische Untersuchung (vgl. Kap. 10) auf längsschnittlich erhobenen Eltern-Kind-Argumentationen, die zu drei Erhebungszeitpunkten (im Alter von ca. 12, 13 und 15 Jahren) elizitiert worden sind (vgl. dazu ausführlicher Kap. 6). Die Untersuchung baut auf den Beschreibungskategorien auf, die entlang der Forschungsfrage I (vgl. oben) entwickelt werden, und untersucht in zwei Fallstudien die Argumentationskompetenz der Heranwachsenden anhand der drei Facetten von Diskurskompetenz (vgl. Kap. 4.1.1).

Insgesamt ist es also das Ziel des Forschungsprojektes, die interaktiven Funktionen von Gesten in ihren Ressourcenverbünden und unter Berücksichtigung der kommunikativen Gattung des Argumentierens zu rekonstruieren. Ausgehend von einem längsschnittlich erhobenen Datenmaterial (vgl. Kap. 6) werden auf den systematischen Befunden aufbauend (vgl. Kap. 8) in den anschließenden Fallstudien (vgl. Kap. 10) potenzielle Veränderungen der sprachlich-diskursiven sowie multimodalen Ressourcen untersucht und daraus ableitend erste Hypothesen sowie weiterführende Fragestellungen zum multimodalen Diskurserwerb entwickelt. Die Längsschnittuntersuchungen können dementsprechend zu einem besseren Verständnis des Erwerbs argumentativer Fähigkeiten beitragen und somit u. a. für die gezieltere Unterstützung des multimodalen Diskurserwerbs z. B. im Rahmen des Deutschunterrichts fruchtbar gemacht werden. Das Forschungsprojekt setzt dabei an der Schnittstelle von multimodaler Konversationsanalyse und der (multimodalen) Diskurserwerbsforschung an (vgl. Kap. 7).

6 Daten

Im Nachfolgenden wird zunächst das Datenkorpus vorgestellt (vgl. Kap. 6.1), auf dessen Grundlage die Untersuchungen durchgeführt werden. Aufgrund der unterschiedlichen Schwerpunkte der Forschungsfragen (vgl. Kap. 5), werden zuerst die Daten des *querschnittlichen* Datenkorpus (Forschungsfrage I, interaktionsbezogene Untersuchung) sowie anschließend des *längsschnittlichen* Datenkorpus (Forschungsfrage II, erwerbsbezogene Untersuchung) vorgestellt. Anschließend wird ein Einblick in die Datenerhebung gegeben (vgl. Kap. 6.2), der u. a. einordnet, inwiefern das Arrangement des Erhebungssettings Einfluss auf die verwendeten multimodalen Ressourcen nehmen kann. Zuletzt wird das Vorgehen der Datenaufbereitung skizziert (vgl. Kap. 6.3).

6.1 Datengrundlage

Die Grundlage der vorliegenden Untersuchung bildet ein über vier Jahre und drei Erhebungszeitpunkte erhobenes, längsschnittliches Datenkorpus mündlicher Eltern-Kind-Argumentationen. Zu den Erhebungszeitpunkten befanden sich die Heranwachsenden in den Jahrgangsstufen 6 bis 9 der Sekundarstufe I und waren ca. 12, 13 und 15 Jahre alt.⁵⁹ Insgesamt beteiligten sich an der Studie 33 Eltern-Kind-Dyaden aus zwei deutschen Großstädten. Die Daten wurden über Entscheidungsaufgaben elizitiert, in denen die Eltern-Kind-Dyaden zu einer begründeten Auswahl über den gewonnenen Preis eines Preisausschreibens (Jahrgangsstufe 6), mögliche Aktivitäten mit Austauschschüler*innen (Jahrgangsstufe 7) oder das Ziel einer anstehenden Abschlussfahrt (Jahrgangsstufe 9) kommen sollten (vgl. 6.2).

Die Datenauswahl für die vorliegende Arbeit erfolgt, aufgrund der zwei unterschiedlich orientierten Fragestellungen (vgl. Kap. 5), in zwei voneinander unabhängigen Schritten.

Die vorliegende Arbeit fokussiert im ersten Teil der qualitativen Untersuchung (vgl. Kap. 8), in dem es um die systematische Beschreibung der multimodalen Realisierung argumentativer Handlungen geht (vgl. Kap. 5, Forschungsfrage I), auf die Eltern-Kind-Interaktionen der neunten Klasse. In diesem Alter agieren Heranwachsende in mündlichen Argumentationen bereits als kompetente(re) Beteiligte und werden auch von erwachsenen Beteiligten vermehrt als solche adressiert (vgl. Quasthoff 2015: 307; vgl. auch Quasthoff et al. 2015; Heller und Krah 2015; Quasthoff und Kluger 2021). Daher wird die Untersuchung der Daten dieses Erhebungszeitpunktes keine Unterscheidung zwischen eltern- und kindseitig multimodal realisierten argumentativen Handlungen vornehmen. Für die Studie wurde im Rahmen der ersten Fragestellung (vgl. Kap. 5) ein Korpus aus insgesamt 20 Eltern-Kind-Interaktionen des dritten Erhebungszeitpunktes gebildet, zu dem die jugendlichen Sprecher*innen ca. 15 Jahre alt sind und gemeinsam mit ihren Eltern die Auswahl eines von vier Zielen einer Klassenfahrt am Ende der neunten Klasse verhandeln. Die erhobenen mündlichen Argumentationen der Dyaden haben eine Gesamtlänge von insgesamt ca. 138 Minuten ($\emptyset$ = 07:15 Minuten), wobei die kürzeste Interaktion 03:21 Minuten und die längste Interaktion 18:28 Minuten dauert (vgl. Tab. 7). Die Auswahl der 20 Eltern-Kind-Dyaden aus den insgesamt 33 Interaktionen erfolgte nach dem Kriterium des Ressourceneinsatzes. Es wurden somit solche Daten ausgewählt, in denen ein*e Sprecher*in Gesten einsetzt. Zugleich wurden Eltern-Kind-

⁵⁹ Die Daten stammen aus der qualitativen Datenerhebung des BMBF-Forschungsprojekts *FUnDuS I und II (Die Rolle familialer Unterstützung beim Erwerb von Diskurs- und Schreibfähigkeit in der Sekundarstufe I)*; Leitung: Prof. 'in Dr. Uta Quasthoff (TU Dortmund) und Prof. 'in Dr. Elke Wild (Universität Bielefeld)), das im Zeitraum 2009-2016 vom Bundesministerium für Bildung und Forschung (BMBF) gefördert worden ist (Förderkennzeichen 01 GJ 1207/A + B.). Für einen Überblick über das Gesamtprojekt vgl. u. a. Wild et al. (2012), Krah et al. (2013), Quasthoff et al. (2015), Quasthoff et al. (2021c).

Dyaden aus der Studie ausgeschlossen, in denen die Interagierenden (im Korpusvergleich) einen besonders hohen Gesteneinsatz aufweisen, der deutlich von den Vergleichsdaten abweicht.⁶⁰

Im zweiten Teil der empirischen Untersuchung (vgl. Kap. 5, Forschungsfrage II) erfolgt zuerst eine Voranalyse des Gesamtkorpus über alle drei Erhebungszeitpunkte hinweg. Entlang der elternseitig etablierten Interaktionsmuster (vgl. Kap. 4.1.3) wurde die Beteiligungsstruktur der Heranwachsenden (vgl. Kap. 2, GLOBE) ermittelt. Die Orientierung an den Interaktionsmustern versprachen dabei, möglichst diskursstrukturell unterschiedlich starke Kinder ermitteln zu können. Pro Interaktionsmuster wurden dann zwei Kinder ausgewählt ($n = 6$), die im Rahmen weiterer Voranalysen interpersonelle Unterschiede hinsichtlich der drei Teilfähigkeiten von Argumentationskompetenz (Kontextualisierung, Vertextung, Markierung; vgl. Quasthoff 2009; vgl. Kap. 4.1.1) zeigen. Die dabei beobachteten Unterschiede können in diskursstrukturellen Profilen der Heranwachsenden, „die empirisch Ausmaß und Art der Varianz in den mündlichen Diskursfähigkeiten der Kinder rekonstruieren“ (vgl. Kern 2003: 36, Fußnote 5; vgl. auch Kern 2003; Quasthoff und Kern 2007; Quasthoff et al. 2019: 59), zusammengefasst werden (vgl. dazu ausführlicher Kap. 10.1.1). Auf Grundlage dieser Voranalysen wurden dann zwei Fokuskinder mit maximal unterschiedlichen diskursstrukturellen Profilen ausgewählt: Dabei zeigen die Fokuskinder jeweils performativ Varianzen im Umgang mit den kommunikativen Anforderungen im Gespräch auf den Ebenen der Kontextualisierung, Vertextung und Markierung. In den mikroanalytischen Fallstudien der Daten dieser zwei Fokuskinder wurde anschließend rekonstruiert, ob und wie sich mit Veränderungen der *sprachlichen-diskursiven* Fähigkeiten auch der Einsatz der *multimodalen* Ressourcen von der sechsten bis zur neunten Klasse verändert. Wenngleich hier eine Trennung zwischen den sprachlich-diskursiven Fähigkeiten und multimodalen Ressourcen suggeriert wird, handelt es sich ausschließlich um eine analytische Entkopplung, da in der vorliegenden Arbeit multimodale Ressourcen grundsätzlich als Teil des argumentativen Diskurserwerbs verstanden werden (vgl. Kap. 10). In der Untersuchung wird dann die Vergleichbarkeit der Daten über die drei Erhebungszeitpunkte zum einen durch ähnliche, aber thematisch jeweils auf das Alter abgestimmte Aufgabenstellungen sowie zum anderen durch die gleich bleibende Konstellation der Eltern-Kind-Dyade gewährleistet (vgl. Quasthoff 2021). Die nachfolgende Tabelle 7 gibt einen Überblick über das zugrundeliegende Datenkorpus.

	Anzahl (n) der Eltern-Kind-Interaktionen	Dauer der Aufnahmen (hh:mm:ss)
Korpus Forschungsfrage I (Alter der Heranwachsenden: ca. 15 Jahre)	$n = 20$	02:17:40
Korpus Forschungsfrage II (Alter der Heranwachsenden: ca. 12, 13 und 15 Jahre)	$n = 6$	00:30:37
GESAMT		02:48:17

Tabelle 7: Übersicht über das Datenkorpus

⁶⁰ Sowohl die Daten, in denen Sprecher*innen fast keine multimodalen Ressourcen nutzen, als auch die Daten, in denen eine überdurchschnittlich hohe Verwendung der multimodalen Ressourcen vorzufinden ist, können in Anschlusspublikationen eingehender untersucht werden. Innerhalb des vorliegenden Promotionsprojektes würden sie aber im Hinblick auf die Frage, welche multimodalen Ressourcen von Sprecher*innen funktional und regelhaft für den Aufbau argumentativer Handlungen verwendet werden, zu einer Verzerrung der Analyseergebnisse führen.

6.2 Datenerhebung

In dem zugrundeliegenden Datenkorpus sind über einen Zeitraum von vier Jahren zu drei Erhebungszeitpunkten Eltern-Kind-Interaktionen erhoben worden, in denen die Beteiligten gebeten worden sind, in Bezug auf die o. g. Fragen (Auswahl eines Preises, einer Aktivität mit Austauschschüler*innen, des Ziels einer Abschlussfahrt) jeweils zu einer begründeten Entscheidung zu kommen. Dabei enthielten die Entscheidungsaufgaben vier Optionen, aus denen die Beteiligten eine begründet auswählen sollten. Die Aufgabenformate wurden über alle drei Erhebungszeitpunkte hinweg konstant gehalten. Ausgangspunkt bildet jeweils eine offene Fragestellung, die hochgradig erwartbar macht, dass die Beteiligten sie argumentativ bearbeiten werden. Dies kann sowohl kollaborativ explorierend, d. h. im Aufbau gemeinsamen Wissens, als auch konfrontativ-persuasiv geschehen, wenn beispielsweise in der Aushandlung divergierende Geltungsansprüche auftreten (vgl. u. a. Mundwiler et al. 2019). Diese Rahmung der Gesprächssituation als (eher) kollaborativ bzw. konfrontativ geschieht durch das Interaktionsverhalten der Beteiligten selbst (vgl. u. a. Goodwin 2007) und kann sich innerhalb der Interaktion kontinuierlich verändern (vgl. u. a. Kreuz 2021: 47) – die kollaborative bzw. konfrontative Rahmung ist dabei nicht bereits durch den vorgegebenen Gesprächsanlass selbst ausgelöst.

Die Datenerhebung hat im Zuhause der Familien stattgefunden und wurde i. d. R. am Ess- oder Wohnzimmertisch durchgeführt. Eine Videokamera wurde den Beteiligten schräg gegenüber als Standkamera positioniert (vgl. Abb. 6); diese videografische Aufzeichnung ermöglicht die Untersuchung der multimodalen Ressourcen im Rahmen dieser Arbeit.⁶¹ Die Beteiligten erhielten durch die Versuchsleiter*innen die Aufgabe, zu einer gemeinsamen und begründeten Entscheidung bezüglich der zukünftigen, hypothetischen Ereignisse zu gelangen, in die von dem*der Versuchsleiter*in durch die phrasenhafte Formulierung *Stellt euch vor* eingeleitet worden ist. Die Interaktion war dabei keiner zeitlichen Beschränkung ausgesetzt. Die Versuchsleiter*innen haben außerdem den Raum verlassen und waren nicht an der Interaktion beteiligt, sodass die Eltern-Kind-Dyaden selbstständig für die gesamte Durchführung sowie auch das Abschließen (vgl. Kap. 2) der argumentativen Entscheidungsfindung verantwortlich waren. Während der Aufzeichnung lag das Material, auf dem die Aufgabenstellung sowie die entsprechenden Auswahloptionen vermerkt waren, zwischen den Interagierenden auf dem Tisch und konnte von ihnen zur Reorientierung oder auch als Argumentationsgrundlage die gesamte Entscheidungsfindung über genutzt werden. Die nachfolgende Abbildung 6 illustriert das prototypische Erhebungssetting.

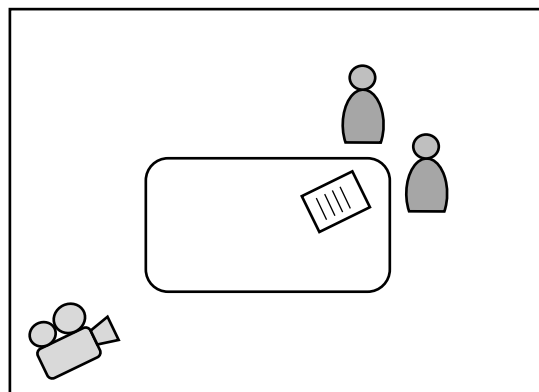


Abbildung 6: Aufnahmesetting der Eltern-Kind-Dyaden

⁶¹ Da sich das Forschungsinteresse im FUnDuS-Projekt nicht auf multimodale Ressourcen richtete, führen die Lichtverhältnisse in manchen Aufnahmen zu einer eingeschränkten Sichtbarkeit der Ressourcen.

Da die Forschungsfragen (vgl. Kap. 5) der vorliegenden Arbeit sich mit den Funktionen multimodaler Ressourcen in der Interaktion befassen, scheint es an dieser Stelle relevant, auf die ökologischen Gegebenheiten (vgl. Mondada 2014b) des Settings einzugehen: Im Vergleich zu dynamischen Interaktionskontexten wie z. B. in Interaktionen während des Laufens (vgl. u. a. Mondada 2013) erweisen sich die vorliegenden Interaktionen durch die sitzende Position der Beteiligten relativ stabil und in gewisser Weise vorstrukturiert (vgl. Müller und Bohle 2007). Gemeinsam bilden die Interagierenden eine sogenannte *F-Formation*, „in which the space between them is one to which they have equal, direct, and exclusive access“ (Kendon 1990b: 209; vgl. auch Ciolek und Kendon 1980). Das bedeutet, dass die Interagierenden durch die Ausrichtung ihrer Körper selbst einen Interaktionsraum herstellen (vgl. Hausendorf et al. 2016: 21), in dem und durch den definiert wird, welche Beteiligten Teil der *fokussierten Interaktion* (vgl. Goffman 1963) sind. Dieser Interaktionsraum, der in Abbildung 7 illustriert wird, wird in dem beschriebenen Aufnahmesetting davon beeinflusst, dass die Eltern-Kind-Dyaden am Tisch nicht nebeneinander sitzen, sondern über Eck, sodass ein *L-Arrangement* (vgl. Kendon 1990b: 213) entsteht, das es den Beteiligten ermöglicht, sich aufeinander zu orientieren (vgl. abgrenzend dazu das *side-by-side* Arrangement, in dem die Interagierenden direkt nebeneinander sitzen, vgl. u. a. Schefflen 1964; Kendon 1990b; Tiittula 2007; Auer und Zima 2021). Den geteilten Handlungsraum, der sich dann zwischen den Beteiligten der F-Formation befindet, nennt Kendon (1990b: 212) *o-space*. In diesem liegt in den vorliegenden Daten das Material mit der verschrifteten Entscheidungsaufgabe, wie der nachfolgenden Abbildung 7 zu entnehmen ist.⁶²

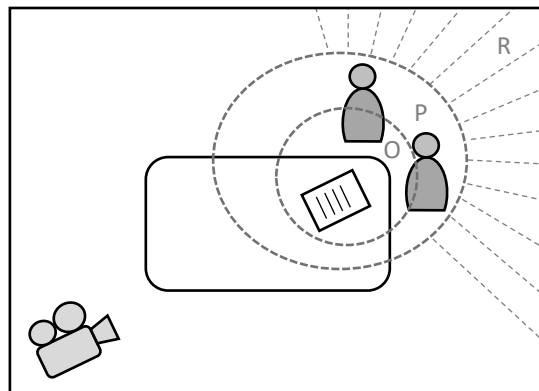


Abbildung 7: Struktur der F-Formation (adaptiert nach Kendon 1990)

In dem beschriebenen Arrangement produzieren die Interagierenden dann innerhalb des sogenannten (individuellen) *Gestenraums* (vgl. u. a. McNeill 1992; Müller 1998; Sweetser und Sizemore 2008) Handgesten. McNeill (1992: 89) unterteilt den Gestenraum einer sitzenden Person schematisch in die vier Bereiche *Center-Center*, *Center*, *Periphery* und *Extreme Periphery*, wie die nachfolgende Abbildung 8 zeigt.

⁶² Kendon (1990b) unterscheidet noch zwei weitere Räume voneinander: den *r-space* und den *o-space*. Personen, die nicht an der Interaktion beteiligt sind, die sich aber in räumlicher Nähe zu den Interagierenden aufhalten, befinden sich in dem sogenannten *r-space*. Der sich zwischen dem inneren *o-* und äußersten *r-space* befindliche Raum wird *p-space* genannt, so Kendon (1990b: 233). In diesem Raum befinden sich die Interagierenden, die körperlich die F-Formation bilden, selbst (vgl. Abb. 7).

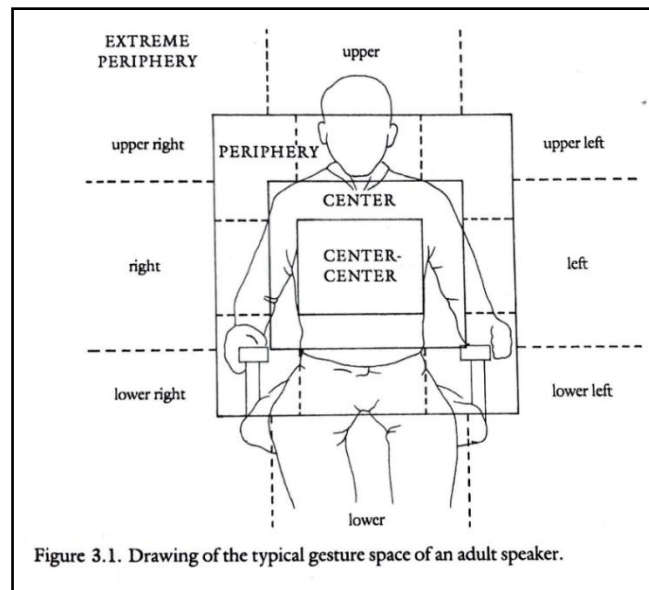


Abbildung 8: Unterteilung des Gestenraums bei sitzenden Personen aus McNeill (1992: 89)

Dabei produzieren sitzende Personen, wie die Interagierenden der vorliegenden Studie, ihre Gesten i. d. R. im Bereich der Körpermitte (*Center-Center* und *Center*). Abhängig von räumlichen Gegebenheiten, etwa wie nah die Interagierenden beieinander sitzen oder ob durch eine Tischplatte der untere Gestenraum beschränkt wird, nutzen die Interagierenden den Gestenraum unterschiedlich – die eingesetzten Ressourcen werden jeweils „adjusted to the local ecology“ (Mondada 2015a: 303). Von einem *persönlichen* Gestenraum der Beteiligten, der eine deutliche Nähe zum eigenen Körper der Interagierenden aufweist, wird ein *interpersoneller* Gestenraum unterschieden, der zwischen den Beteiligten liegt. In diesem können alle Beteiligten Gesten äußern (vgl. Sweetser und Sizemore 2008: 25f.). Die Gesten der Beteiligten können aber auch in den persönlichen Gestenraum der *anderen* hineinreichen, wenn beispielsweise die Aufmerksamkeit eines*iner bestimmten Beteiligten eingefordert wird (vgl. Sweetser und Sizemore 2008: 26). Dabei ist die Unterteilung zwischen persönlichem und interpersonellem Gestenraum insgesamt fließend. Setzt man die hier beschriebenen persönlichen und interpersonellen Gestenräume in Beziehung zu dem oben beschriebenen *o-space*, wird deutlich, dass die Gestenäußerungen der Interagierenden – solange sie sich in einer gemeinsamen, aufeinander orientierten Interaktion befinden – im gemeinsamen *o-space* geäußert werden. Die hier getroffene Unterscheidung der Gestenräume wird in der vorliegenden Arbeit dann relevant, wenn Interagierende tatsächlich in den persönlichen Gestenraum anderer vordringen und dies interaktive Konsequenzen für die argumentative Handlung hat.

6.3 Datenaufbereitung

Die videographierten Daten wurden grundlegend nach den Konventionen des Gesprächsanalytischen Transkriptionssystems 2 (GAT 2) nach Selting et al. (2009) unter Einbezug der Transkription multimodaler Ressourcen nach Mondada (2014b, 2018) verschriftet. Die Transkripte wurden als Basistranskript mit entsprechenden Erweiterungen des Feintranskripts angefertigt, in denen die multimodalen Ressourcen, die zur Realisierung der argumentativen Handlungen eingesetzt worden sind, jeweils unter dem Verbaltranskript notiert werden (s. u.; vgl. Transkriptionskonventionen, Anhang). Die Beschreibung der geäußerten Gesten orientiert sich dabei u. a. an den von Kendon (2004) etablierten *Formbezeichnungen* der Gesten (z. B. *Palm Up* Geste), während für den Blick der Interagierenden angegeben wird, in welche

Richtung er erfolgt (z. B. *blickt zu Mutter*), und für Kopfgesten die *Bewegung* sowie ihre *Bewegungsrichtung* benannt wird (z. B. *lehnt Kopf Richtung Sohn*). Da die Ressourcen auch über das Ende der sprecher*innenseitigen Äußerungen hinweg gehalten werden können (vgl. Kap. 3), wird die Zugehörigkeit der jeweiligen Ressource ebenfalls mit einer personenbezogenen Sigle versehen (vgl. Mondada 2018). Die multimodalen Ressourcen sind mit senkrechten Strichen (||) mit dem verbalsprachlichen Transkript aligniert (vgl. Heller 2018a)⁶³, die auch Informationen über die Gestenphasierung (vgl. Kendon 1980, 2004; vgl. auch Kap. 3.1.1), d. h. ihre Vorbereitung sowie die Rückkehr der Ressource in eine Ruheposition, enthalten können. Da die sprachliche Beschreibung in der Transkription sowohl die Entfaltung der körperlich-sichtbaren Ressourcen als auch ihre Gesamtgestalt z. T. nur unzureichend abbilden kann (vgl. Mondada 2018), werden die Transkripte um Standbilder ergänzt, die durch eine computergestützte Software verfremdet worden sind (vgl. Heller 2018b: 249).⁶⁴ An gegebener Stelle werden die Standbilder um Symbole wie beispielsweise Pfeile ergänzt, um die Bewegungsrichtungen sichtbar zu machen. Der Zeitpunkt der Entnahme des Standbildes wird darüber hinaus im Verbaltranskript gekennzeichnet. Dabei bilden sowohl die Auswahl der Standbilder als auch die Ergänzung um weitere Symbole einen interpretativen Eingriff in die Daten (vgl. Stukenbrock 2009: 150; Deppermann 2013: 3), denn der*die Forscher*in entscheidet, welches Standbild beispielsweise aufgrund seiner illustrierenden Funktion Eingang in das Transkript findet. In der vorliegenden Arbeit dienen die Standbilder in erster Linie den Leser*innen zur Nachvollziehbarkeit der eingesetzten Ressourcen. Die Analysen wurden von der Verfasserin dieser Arbeit hingegen auf Grundlage der Transkripte sowie der Videoaufzeichnungen durchgeführt (vgl. Gülich und Mondada 2008: 117).⁶⁵

I. S. der Lesbarkeit (vgl. Selting et al. 2009) werden nur diejenigen Anteile einer Sequenz durch die Notation der körperlichen Ressourcen ergänzt, die gemäß dem Forschungsinteresse der vorliegenden Untersuchung (vgl. Kap. 5) zur Konstituierung einer argumentativen Handlung sowie ggf. zur Untersuchung von Anschlusshandlungen beitragen und Gesten mit pragmatischen Funktionen enthalten. Hinter dieser Auswahl verbergen sich bereits zwei analytische Schritte: die Identifizierung *argumentativer Handlungen* (vgl. Kap. 2.2.2), sowie, und dies betrifft die Transkription selbst, die Auswahl der zu *notierenden multimodalen Ressourcen*. Denn

[d]ie Transkription des Multimodalen ist notwendigerweise selektiv: Angesichts der zahlreichen simultan ablaufenden multimodalen Handlungen müssen diejenigen ausgewählt werden, die für das untersuchte Phänomen relevant sind (Gülich und Mondada 2008: 117).

Die Phänomene sind i. d. S. durch das Transkribieren selbst immer schon „einem Transformations- und Reduktionsprozess unterworfen“ (Stukenbrock 2009: 147), sodass das Anfertigen eines Transkripts selbst samt den abgebildeten Ressourcen „a thoroughly reflexive and interpretative accomplishment“ (Deppermann 2013: 3) darstellt. Welche multimodalen Ressourcen nun in diesem Reduktionsprozess transkribiert werden und in welcher Granularität (vgl. Deppermann 2013: 3), ist einerseits abhängig von

⁶³ Vgl. zu unterschiedlichen Formen der multimodalen Transkription u. a. Goodwin (1981); Heath (1984); Streeck (1993); Kendon (2004); Gülich und Mondada (2008); Stukenbrock (2009); Heath und Luff (2013); Mondada (2018).

⁶⁴ Mondada (2018: 90) verweist darauf, dass durch das Einfügen von Standbildern die Möglichkeit für die Leser*innen entsteht, die sich *analytisch*, d. h. einzeln entfaltenden Ressourcen in ihrer *synthetischen* Gesamtgestalt zu erfassen.

⁶⁵ Bereits die Positionierung der Kamera, der gewählte Bildausschnitt, die Fokussierung der Kamera etc. stellen interpretative Eingriffe in die Daten dar. Dabei vermag die Kamera nur ein unvollständiges *Abbild* der sozialen Interaktion herzustellen, denn sowohl außerhalb des Objektivs liegende räumliche bzw. materielle Ressourcen entgehen Forscher*innen in dem Videomaterial als auch beispielsweise olfaktorische bzw. sensorische Informationen (vgl. u. a. Mondada 2021).

dem zu untersuchenden Phänomen (vgl. Kap. 5) sowie andererseits davon, welche Ressourcen die Interagierenden selbst für ihre Handlungen relevant machen (vgl. Mondada 2018). Im Fokus der vorliegenden Arbeit werden die multimodalen Ressourcen transkribiert, die gemeinsam eine argumentative Handlung realisieren und in fein abgestimmter Koordinierung multimodale Gestalten bilden. Dabei werden die relevant gemachten multimodalen Ressourcen der Gesprächsbeteiligten unabhängig von ihrer ‚Sprecher*innenrolle‘ untersucht. Dies ermöglicht, auch argumentative Handlungen zu erfassen, die ausschließlich durch körperlich-sichtbare Ressourcen realisiert werden.

Innerhalb der Aufbereitung wurden die Daten anonymisiert, d. h. dass alle Informationen aus den Daten getilgt wurden, die Aufschluss auf die Identität der Beteiligten geben (vgl. auch Deppermann 2008). In diesem Zuge wurden auch die Siglen der Beteiligten der Eltern-Kind-Interaktionen entsprechend ihren familialen Rollen, wie Mutter (*MU*) und Sohn (*SO*), gewählt.

Sowohl die Transkription als auch die Datenverwaltung erfolgte mit der Software MAXQDA, die der computergestützten qualitativen Datenanalyse dient. Diese wurde eingesetzt, um die Befunde der mikroanalytischen, induktiv rekonstruktiv vorgenommenen Untersuchungen zu den o. g. Fragestellungen (vgl. auch Kap. 5) in ersten Kollektionen zu bündeln und mit fortschreitenden Analysen die Kollektionen auszudifferenzieren und dabei auch Neusortierungen vorzunehmen. I. d. S. dient die Software MAXQDA der Forscherin v. a. als administratives Instrument, um die daten- bzw. gesprächsübergreifenden Befunde zu systematisieren.

7 Methode

Um den in Kapitel 5 vorgestellten Fragestellungen in den empirischen Kapiteln 8 und 10 nachgehen zu können, bedarf es vorab der Klärung der forschungsmethodischen Grundlagen, denen die vorliegende Arbeit folgt. Diese Funktion übernimmt das vorliegende Kapitel 7. Da die Arbeit sich zwischen Ansätzen der multimodalen Konversationsanalyse sowie der (multimodalen) Diskurserwerbsforschung bewegt, werden zunächst beide Zugänge vorgestellt (vgl. Kap. 7.1). Auf dieser Einordnung aufbauend wird dann das Vorgehen der Datenanalyse der empirischen Untersuchungen erläutert (vgl. Kap. 7.2).

7.1 Forschungsmethodische Einordnung

Zur Klärung der ersten Fragestellung, die an den Funktionen multimodaler Ressourcen für das mündliche Argumentieren interessiert ist (vgl. Kap. 5, Forschungsfrage I), legt die Arbeit den Ansatz der multimodalen Konversationsanalyse zugrunde. Diese ermöglicht eine Untersuchung linguistischer Phänomene in der Interaktion unter Berücksichtigung körperlich-sichtbarer Ressourcen. Da die multimodale Konversationsanalyse aus der ethnomethodologisch fundierten Konversationsanalyse heraus entstanden ist, wird einleitend dargelegt, welches Erkenntnisinteresse die Konversationsanalyse verfolgt – und wie sich dieses mit der *multimodalen* Konversationsanalyse verändert hat.

Die ethnomethodologische Konversationsanalyse ist daran interessiert, wie durch sprachliches Handeln soziale Wirklichkeit entsteht (vgl. u. a. Bergmann 2001; Deppermann 2008). Dabei geht sie der Frage nach, wie es Sprecher*innen gelingt, gemeinsam soziale Interaktion aufzubauen (vgl. Bergmann 2001). Durch sequenzanalytische Untersuchungen, in denen die Interaktion Redezug um Redezug nachvollzogen wird, zeigt sich, dass trotz des allgemeinen Eindrucks, dass soziale Interaktionen ‚ungeordnet‘ oder ‚chaotisch‘ seien, „die Kommunikationspartner ihre Handlungen methodisch und systematisch hervorbringen“ (Gülich und Mondada 2008: 17): Es werden Ordnungsstrukturen in der sozialen Interaktion sichtbar. Diese entstehen aufgrund sogenannter konditioneller Relevanzen, d. h. spezifische Handlungen, wie z. B. Fragen, machen Folgehandlungen, wie z. B. Antworten, regelhaft erwartbar. Diese regelhaft miteinander auftretenden Handlungen werden auch Paarsequenzen oder *adjacency pairs* (vgl. Schegloff und Sacks 1973: 73ff.; Schegloff 2007: 13ff.) genannt. Es zeigt sich, dass Interagierende ihr Handeln in der Interaktion an solchen Erwartbarkeiten orientieren und das Gespräch *turn-by-turn* aufbauen. In wechselseitiger Orientierung aufeinander zeigen sich die Sprecher*innen an, wie sie die Vorgängerausäußerung interpretieren, und bilden wiederum durch ihren eigenen Beitrag Anschlussmöglichkeiten für die Fortsetzung der Interaktion.

Lange Zeit stehen vornehmlich *sprachlich* realisierte soziale Handlungen im Fokus konversationsanalytischer Forschung (vgl. aber Schegloff 1984; 1998)⁶⁶, sodass sich auch die entstehenden Transkriptionssysteme (vgl. u. a. Jefferson 1979; Selting et al. 1998, 2009) zur Untersuchung sozialer Interaktion vornehmlich an den sprachlichen und prosodischen Ressourcen der Beteiligten orientieren, die sich linear entfalten und sequenziell organisiert sind.

⁶⁶ Bereits Schegloff (1984) verweist auf die projizierende Funktion von Gesten, die u. a. Turnübernahmen und somit Sprecher*innenwechsel anzeigen können.

Bei der Einbindung körperlich-sichtbarer Ressourcen in die Untersuchung von Interaktionen handelt es sich nun keinesfalls um eine „neue Analysemethode oder eine neue Methodologie, sondern [die Multimodalität] bezeichnet eine konstitutive Qualität von Interaktion“ (Mondada und Schmitt 2010: 22). Bereits frühe konversationsanalytische bzw. konversationsanalytisch orientierte Arbeiten beziehen Multimodalität in die Untersuchung von Interaktionen mit ein und prägen dadurch auch das Verständnis der Relevanz von Körperlichkeit in der Interaktion (vgl. u. a. Goodwin 1979, 1981, 1986; Goodwin 1980; Schegloff 1984, 1998; Heath 1984, 1985; Streeck und Hartge 1992). Ausgehend von diesen frühen Arbeiten entwickelten sich forschungsmethodische Zugänge einer multimodalen Konversationsanalyse (vgl. u. a. Stivers und Sidnell 2005; Gülich und Mondada 2008; Stukenbrock 2009; Streeck et al. 2011; Deppermann 2013, 2018; Mondada 2014b, 2016, 2018; Hofstetter und Keevallik 2020). Diese legen u. a. dar, wie die Integration von Sprache und Körperlichkeit (sowie Räumlichkeit) unter einer konversationsanalytischen Perspektive gelingen kann. Denn mit Einbezug körperlich-sichtbarer Ressourcen, die *simultan* (vgl. Deppermann 2013; Mondada 2021a) geäußert werden können und somit gleichzeitig mit den sich linear *nacheinander* entfaltenden sprachlichen Ressourcen organisiert werden (vgl. Birkner et al. 2020: 235), entstehen entscheidende Konsequenzen für die Entwicklung einer multimodal orientierten Konversationsanalyse. Der Einbezug simultaner multimodaler Ressourcen zeigt, dass zeitgleich mit den Handlungen, die sich in der sprachlichen Linearität aufbauen, weitere Handlungen körperlich-sichtbar angezeigt werden können. In die Untersuchung der *sequenziellen* Entfaltung einer Interaktion muss nun die der *simultanen* integriert werden (vgl. Deppermann 2013; Mondada 2014b). Mit der Anerkennung körperlicher Ressourcen als wichtige Bestandteile von Interaktionen sowie der Berücksichtigung räumlicher sowie auch materieller Bedingungen der Interaktion wird auch eine Priorisierung sprachlicher Ressourcen aufgegeben. Daraus resultieren sowohl konzeptionelle als auch methodische Implikationen: So nimmt Mondada (2014b) konzeptionell eine Gleichwertigkeit der vorkommenden Ressourcen im Gespräch an, in der die körperlichen Ressourcen den sprachlichen Ressourcen nicht untergeordnet sind, sondern in der sie *gemeinsam*, i. S. der *Complex Multimodal Gestalts* (vgl. Mondada 2014b; vgl. Kap. 3.1), Handlungen und somit soziale Interaktion aufbauen. Die multimodalen Ressourcen werden dabei in ihren Verbänden von den Beteiligten der Interaktion „holistisch wahrgenommen“ (Gülich und Mondada 2008: 97). Die Auffassung einer Gleichwertigkeit der Ressourcen (vgl. Mondada 2014b; vgl. Kap. 3.1.4) erfordert wiederum einen kritischen Umgang mit bestehenden Begriffen, wie beispielsweise *Sprecher*innen* und *Rezipient*innen* bzw. *Adressat*innen*. Es scheint daher angemessen, so Deppermann (2013: 3), von beispielsweise *Beteiligten* (*participants*) zu sprechen (vgl. hierzu auch Deppermann und Schmitt 2007). Methodisch wiederum erfordert der Einbezug multimodaler Ressourcen, dass die zu untersuchenden Daten videographiert vorliegen, sowie ein Transkriptionssystem, das es ermöglicht, das *lineare* sowie *simultane* Entfalten sprachlicher *und* körperlicher Ressourcen sowie die Bezugnahmen auf räumliche und materielle Gegebenheiten bzw. Ressourcen abzubilden, sobald sie von den Interagierenden in ihrem Handeln relevant gemacht werden (vgl. u. a. Goodwin 2000; Schmitt und Hausendorf 2016; Mondada 2018). Es wird nun möglich, nicht nur das sprachliche Handeln der Interagierenden zu untersuchen, sondern auch zu berücksichtigen, dass soziale Interaktion in körperlicher Kopräsenz der Interagierenden innerhalb eines Raumes, der unterschiedliche Materialitäten beinhaltet, stattfindet. So entstehen u. a. Arbeiten zur Koordination im Raum, z. B. bei Stadtführungen oder in Tanzstunden (vgl. u. a. den Sammelband von Schmitt 2007), zur Organisation und Koordination von *multiactivities*, d. h. gleichzeitig ablaufender Aktivitäten (vgl. u. a. den Sammelband von Haddington et al. 2014) sowie in jüngster Zeit auch zu Berührungen (vgl. u. a. Mondada und Tekin 2020). Durch diese Öffnung der multimodalen Konversationsanalyse wird der Untersuchungsgegenstand der Interaktion nun „als ein auf unterschiedlichen Ausdrucksebenen hervorgebrachtes, von allen Beteiligten zu jedem Zeitpunkt gemeinsam konstituiertes und praxeologisch gerahmtes Ereignis begriffen“ (Deppermann und

Schmitt 2007: 22). Da u. a. die Funktion bzw. Bedeutung körperlich-sichtbarer Ressourcen allerdings häufig implizit und dadurch ambig verbleibt, verweist Deppermann (2013: 4) darauf, dass

the meaning of such phenomena and their import for the interaction at hand does not become clear from close interactional analysis of participants' displays alone. [...] Close analysis of visible and audible multimodal interaction organization therefore often has to be complemented by ethnography.

Um also in der hervorgebrachten sozialen Interaktion u. a. auch die Funktion multimodaler Ressourcen rekonstruieren zu können, wird der forschungsmethodische Ansatz der multimodalen Konversationsanalyse durch Erkenntnisse der Ethnographie sowie Praxeologie angereichert (vgl. Deppermann 2013; Streeck 2013, 2016). I. d. R. greifen Studien im Bereich der multimodalen Konversationsanalyse außerdem auf Befunde aus den Forschungsbereichen der Gestenforschung und z. T. auch aus der Kognitionspsychologie zurück, die im Rahmen der vorliegenden Arbeit in Kapitel 3 vorgestellt worden sind.

Die Konversationsanalyse ist v. a. an der induktiven, d. h. von den Daten selbst ausgehenden Rekonstruktion der sich in der Interaktion entfaltenden sequenziellen Geordnetheit interessiert und fokussiert dabei auf die in der Interaktion manifest werdenden kommunikativen Handlungen, die einen Blick individuelle Fähigkeiten der Gesprächsbeteiligten ausspart (vgl. Hausendorf und Quasthoff 2005). Daher muss im Hinblick auf die zweite leitende Fragestellung (vgl. Kap. 5, Forschungsfrage II) das gesprächs- bzw. konversationsanalytische Paradigma um einen forschungsmethodischen Ansatz erweitert werden, mithilfe dessen längsschnittliche Untersuchungen von *Erwerbsphänomenen* möglich werden. Dafür wird in der vorliegenden Arbeit der Ansatz der *Interaktionalen Diskursanalyse* (IDA) relevant (vgl. u. a. Morek et al. 2017). Es handelt sich um einen Ansatz, der zwischen der Gesprächsforschung, Diskurserwerbsforschung und Soziolinguistik verortet ist und mithilfe dessen induktiv-rekonstruierend interaktive Verfahren untersucht werden können, mit denen Interagierende größere Diskurseinheiten wie das Argumentieren umsetzen (vgl. Heller 2012; Morek 2012; Morek et al. 2017: 16; Quasthoff et al. 2017: 85). Das Interesse, diskurserwerbstheoretische Fragen konversationsanalytisch rückzubinden, ist bedingt durch die Annahme, dass Diskurserwerbsprozesse „nicht unabhängig und losgelöst von Interaktionsprozessen“ (Hausendorf und Quasthoff 2005: 588) verstanden werden, sondern sie werden verstanden als aufeinander bezogene Prozesse. Mit der IDA wird es möglich, sowohl die *interpersonelle Koordinierung* und somit *gemeinsame* Beteiligung der Interagierenden im Aufbau der argumentativen Diskurseinheit zu rekonstruieren (vgl. Kap. 2.2) und dabei gleichzeitig die *Anteile* und *Fähigkeiten der einzelnen Sprecher*innen* in den Blick zu nehmen (vgl. Kap. 4.1.1). Es können sowohl konversationsanalytische als auch erwerbsorientierte Phänomene mit diesem Ansatz untersucht werden. Die IDA ist dabei „strikt teilnehmer- und funktional orientiert“ sowie „auf die Analyse globaler Diskurspraktiken zugeschnitten“ (Morek et al. 2017: 18). Dabei erweist sich die Arbeit mit elizitierten Daten als durchaus sinnvoll, auch wenn sich dieses Vorgehen in einem Spannungsverhältnis zu einem streng konversationsanalytischen Paradigma (vgl. u. a. Goodwin und Heritage 1990: 289) befindet. Denn im Hinblick auf die Untersuchung erwerbsorientierter Phänomene ermöglicht eine Vergleichbarkeit der Daten, „Aspekte der Varianz argumentativer [...] Praktiken sowie deren Erwerb analytisch zu beschreiben“ (Morek et al. 2017: 18) und vergleichend intra- sowie interindividuelle Unterschiede aufzudecken (vgl. Quasthoff 2021).

Die hier vorgestellten forschungsmethodischen Zugänge werden im Nachfolgenden (vgl. Kap. 7.2) für das Vorgehen in der Datenanalyse fruchtbar gemacht.

7.2 Vorgehen in der Datenanalyse

Das Vorgehen zur Untersuchung der Daten folgt, wie oben beschrieben, einem induktiv-rekonstruktiven Forschungszugang, innerhalb dessen, aufgrund des Forschungsinteresses an der Multimodalität in Argumentationen, sehr feinkörnige konversationsanalytische Mikroanalysen durchgeführt werden. Angesichts des zweigeteilten Forschungsinteresses verlangt das Vorgehen in der empirischen Untersuchung zwei unterschiedliche Zugänge, die im Folgenden entlang der grundlegenden Fragestellungen der Studie vorgestellt werden.

Die vorliegende Studie ist daran interessiert, die Funktionen multimodaler Ressourcen für das mündliche Argumentieren zu rekonstruieren (vgl. Kap. 5, Forschungsfrage I). Das Interesse der Studie richtet sich dabei auf die multimodalen Gestalten, die an der Realisierung argumentativer Handlungen beteiligt sind und die anhand des multimodalen konversationsanalytischen Zugriffs (vgl. Kap. 7.1) in induktiv-rekonstruktiven Analysen herausgearbeitet werden können (vgl. Mondada 2014b, 2018). Der erste Teil der Datenanalyse erfolgt unter Rückgriff auf das Beschreibungsinstrument GLOBE (vgl. Hausendorf und Quasthoff 1996/2005; Heller 2012; vgl. Kap. 2), mithilfe dessen in der Globalstruktur der jeweiligen Argumentationen die argumentativen Handlungen rekonstruiert werden können, die im Zentrum der Analyse stehen (vgl. Kap. 2 & Kap. 5). Dabei gehen die Beobachtungen, i. S. eines konversationsanalytischen Paradigmas, „from the data themselves“ (Schegloff und Sacks 1973: 291) aus, um nicht Gefahr zu laufen, „Vorannahmen als Ergebnisse zu reproduzieren“ (Deppermann 2008: 19). Es werden also in einem ersten Schritt die argumentativen Handlungen, wie das Beziehen einer Position, das Begründen oder das Einwenden (vgl. Kap. 2.2.2) aus den Daten der argumentativen Entscheidungsfindungen der Eltern-Kind-Interaktionen zum dritten Erhebungszeitpunkt herauspräpariert. Die Analyse berücksichtigt dabei aufgrund der fortgeschrittenen Adoleszenz und damit verbunden sprachlich-diskursiven Erfahrungen in unterschiedlichen Kontexten (wie Familie, Schule, vgl. Kap. 4.1.3) sowohl die Beiträge der jugendlichen sowie auch der erwachsenen Sprecher*innen. Gemäß dem Forschungsinteresse fokussieren die Analysen dabei nicht ausschließlich auf *sprachlich*, sondern auf *multimodal* realisierte argumentative Handlungen, mit der die Interagierenden die kommunikativen Aufgaben bearbeiten. Die dabei rekonstruierten Phänomene werden in Kollektionen zusammengeführt (vgl. hierzu auch Hausendorf und Quasthoff 2005: 600; Heritage 2011) und durch rekursiv durchlaufene Analyseprozesse auch intersubjektiv abgesichert (vgl. Deppermann 2008). Anhand mikroanalytischer Untersuchungen wird die temporale sowie sequenzielle Entfaltung der an den argumentativen Handlungen beteiligten Ressourcen nachvollzogen. Dieses Vorgehen ermöglicht nun präzise Einsichten in die *Funktionen* gestischer Ressourcen in ihren Ressourcenverbünden für die jeweilige argumentative Handlung. Dies ermöglicht eine Ausdifferenzierung der Kollektionen sowie außerdem eine von konkreten Interaktionsdyaden unabhängige Beschreibung der rekonstruierten multimodalen Gestalt (vgl. Kap. 8).⁶⁷

Im zweiten Teil der empirischen Untersuchung (vgl. Kap. 5, Forschungsfrage II) erfolgt unter Rückgriff auf die Interaktionale Diskursanalyse (IDA) dann eine Fokussierung auf erwerbsbezogene Phänomene. Dazu erfolgt die Auswahl zweier Fokuskinder für die explorativen Fallstudien in drei Schritten. In einem *ersten Schritt* werden die 33 Eltern-Kind-Interaktionen des Hauptkorpus über alle drei Erhebungszeitpunkte hinweg hinsichtlich der Beteiligungsstruktur der Kinder, d. h. inwiefern sie den Interaktionsraum selbstständig nutzen (können) (vgl. Kap. 2.2, GLOBE), sowie der Verwendung von Gesten in pragmatischer Funktion gesichtet. Da davon ausgegangen wird, dass für eine erwerbsbezogene, vergleichende

⁶⁷ Mundwiler et al. (2019: 325) beschreiben, dass konversationsanalytische Studien durch das Bilden von Kollektionen bereits lange informale Kodierungen vornehmen. Dieses Verständnis aufgreifend arbeitet auch die vorliegende Studie mit Kodierungen von induktiv rekonstruierten Kollektionen der multimodal realisierten argumentativen Handlungen (vgl. zum Kodieren in der Konversationsanalyse auch Stivers 2015).

Untersuchung die Auswahl von Kindern mit möglichst differierenden Fähigkeiten sinnvoll erscheint, werden in einem *zweiten Schritt*, mit Fokus auf den dritten Erhebungszeitpunkt, sechs Kinder entlang der Interaktionsmuster *Fordern und Unterstützen*, *Raumlassen und Akzeptieren* sowie *Übergehen und Selberlösen* ausgewählt (vgl. Heller und Krah 2015). Die Orientierung entlang der Interaktionsmuster, die sich als unterschiedlich erwerbssupportiv erweisen (vgl. Heller und Krah 2015), erhöht die Wahrscheinlichkeit, sowohl diskursstrukturell stärkere als auch diskursstrukturell schwächere Kinder ermitteln zu können. Diese Hypothese bestätigend zeigen sich in der Untersuchung der Argumentationskompetenz dieser sechs Kinder Unterschiede in Bezug auf die drei Ebenen der Argumentationskompetenz (vgl. Kap. 4.1.1). Diese umfassen die Fähigkeit, globale Zugzwänge, wie das Einfordern einer Begründung, erkennen, bedienen und produzieren zu können (*Kontextualisierung*); satzübergreifende Äußerungspakete, wie eine Begründung, erkennen, aufbauen und sie intern gemäß der Gattung des Argumentierens strukturieren zu können (*Vertexung*); sowie dazu auf die entsprechenden Formen der Markierung, wie das Anzeigen einer Kausalität durch beispielsweise die Konjunktion *weil*, rezeptiv und produktiv zuzugreifen zu können (*Markierung*). Trotz der performativen Varianzen im Umgang mit den kommunikativen Anforderungen im Gespräch (vgl. zum Kompetenz-/Performanzbegriff u. a. Quasthoff 2009; Kreuz und Luginbühl 2020) zeigen sich auch interindividuelle Ähnlichkeiten. Diese werden zu kindübergreifenden diskursstrukturellen Profilen zusammengefasst (vgl. auch Quasthoff et al. 2019: 63).

Die Ermittlung der diskursstrukturellen Profile dient dabei nun dem *dritten Schritt*: der Auswahl der Fokuskinder. Anhand der diskursstrukturellen Profile können zwei Fokuskinder mit maximal kontrastiven diskursstrukturellen Profilen ausgewählt werden, die jeweils Unterschiede hinsichtlich der Kontextualisierungs-, Vertexungs- und Markierungskompetenz zeigen. Um dem benannten Forschungsinteresse nachzugehen, werden in den Fallstudien nun zwei voneinander getrennte Analyseschritte durchgeführt (vgl. zur analytischen Trennung Kap. 6.1): Es werden zuerst detailliert die sprachlich-diskursiven Entwicklungslinien der Fokuskinder über alle drei Erhebungszeitpunkte hinweg rekonstruiert. Da Erwerbsstudien zum Argumentieren außerdem zeigen, dass Sprecher*innen mit ansteigenden argumentativen Fähigkeiten auch vermehrt epistemische Modalisierungen z. B. zur Abschwächung oder Verstärkung einer Begründung nutzen (vgl. Kreuz und Luginbühl 2020; Morek 2020), werden diese ergänzend in den Fallstudien auf der Ebene der Markierung untersucht. Aufbauend auf der systematischen Beschreibung der multimodalen Gestalten in Kapitel 8, die an der Realisierung argumentativer Handlungen beteiligt sind, wird nun möglich, die *multimodalen* Entwicklungslinien der Fokuskinder über alle drei Erhebungszeitpunkte hinweg zu rekonstruieren. Im Mittelpunkt der Untersuchung steht dabei die argumentative Handlung des Begründens. Durch diese forschungsmethodische Entscheidung wird es einerseits möglich, die Entwicklungslinien der Forschungskinder anhand einer für das Argumentieren überaus relevanten Handlung nachvollziehbar zu machen. Andererseits wird angenommen, dass die Realisierung gerade dieser Handlung sich über den Erwerb verändert, denn tendenziell beginnen Heranwachsende mit zunehmendem Diskurserwerb z. B. Begründungen global strukturierter aufzubauen, sie initiiierend einzubringen und sequenziell in den Gesprächskontext einzupassen (vgl. Kap. 4.1.2). Wie die sprachlich-diskursive Entwicklungslinie wird auch die *multimodale* Entwicklungslinie anhand des Modells der Argumentationskompetenz untersucht. Fokussiert wird hier nun auf multimodale Gestalten, die dazu beitragen, argumentative Beiträge zu *kontextualisieren* (z. B. indem eine Begründung auf der geöffneten Handfläche präsentiert und zur weiteren Aushandlung angeboten wird), Beiträge zu *vertexen* (z. B. indem durch Gestenwiederaufnahmen Bestandteile einer Begründung kohäsiiv miteinander verknüpfen) sowie Beiträge epistemisch abschwächend oder verstärkend zu *modalisieren* (indem z. B. durch das Hochziehen der Schultern und das Aufdrehen der Hände der Inhalt einer Begründung als unsicher markiert wird). Die Untersuchung berücksichtigt in diesem Vorgehen grundsätzlich die Ebene der Markierung. Da sowohl die sprachlich-diskursiven als auch die multimodalen Entwicklungslinien

anhand der gleichen Analyseebenen untersucht werden, wird es dann möglich, die Zusammenhänge zwischen den sprachlich-diskursiven und multimodalen Entwicklungslinien zu ermitteln und aus diesen mögliche Hypothesen zum multimodalen Diskurserwerb abzuleiten.

8 Interaktionsbezogene Analysen der multimodalen Gestalten zur Realisierung argumentativer Handlungen für das Argumentieren

Entsprechend dem Interesse der ersten Forschungsfrage, *welche Funktionen multimodale Ressourcen für das mündliche Argumentieren übernehmen* (vgl. Kapitel 5), werden im Folgenden die rekonstruierten multimodalen Gestalten vorgestellt, die an der Realisierung argumentativer Handlungen wiederkehrend beteiligt sind. Der Fokus der Studie liegt dabei auf der Untersuchung von Gesten mit pragmatischen Funktionen in ihren jeweiligen Ressourcenverbünden.

Der Aufbau des Kapitels orientiert sich an den argumentativen Handlungen, die mittels GLOBE (vgl. Kap. 2 & 7.2) rekonstruiert werden konnten. Es handelt sich um die argumentativen Handlungen des *Beziehens einer Position* (vgl. Kap. 8.1), *Begründens* (vgl. Kap. 8.2), *Einwendens* (vgl. Kap. 8.3), *Beipflichtens bzw. schwachen Zustimmens* (vgl. Kap. 8.4) sowie des *Einforderns von Rückmeldungen* (vgl. Kap. 8.5). In der Untersuchung der Daten wird außerdem sichtbar, dass Interagierende multimodale Ressourcen zur *epistemischen Modalisierung* argumentativer Handlungen einsetzen. Da epistemische Modalisierungen beispielsweise dazu beitragen, den Grad der Verhandelbarkeit einer eingebrachten Position zu indizieren, sind sie äußerst relevant für das mündliche Argumentieren. Die multimodalen Gestalten werden zur Modalisierung unterschiedlicher argumentativer Handlungen, d. h. beim Beziehen einer Position, Begründen oder Einwenden eingesetzt. Daher werden die *epistemischen Modalisierungen* im letzten Kapitel (vgl. Kap. 8.6) handlungsübergreifend vorgestellt. Das Kapitel wird ferner in Vorausschau auf die erwerbsbezogene Untersuchung (vgl. Kap. 10) eingezogen, denn der Einsatz epistemischer Modalisierungen verändert sich innerhalb des Erwerbs argumentativer Fähigkeiten (vgl. Kap. 4.1.2). Indem also nachfolgend untersucht wird, wie kompetente(re) Sprecher*innen multimodale Gestalten zur epistemischen Modalisierung verwenden, können diese – sowie ihre Veränderungen – auch in der anschließenden Fallstudie berücksichtigt werden. Im Anschluss an die Befunde zur epistemischen Modalisierung schließt das Kapitel mit einer zusammenfassenden Übersicht über die rekonstruierten Funktionen der multimodalen Gestalten, bevor dann der Ertrag der Forschungsergebnisse für eine multimodale Gattungsforschung diskutiert wird (vgl. Kap. 9).

Zu Beginn der einzelnen Unterkapitel und unter Rückgriff auf die theoretischen Grundlagen (vgl. Kap. 2.2) wird die jeweilige argumentative Handlung (vgl. Kap. 8.1-8.5), an deren Realisierung die multimodalen Gestalten beteiligt sind, knapp in ihren wichtigsten Grundzügen vorgestellt – und, sofern vorhanden, mit Studien in Beziehung gesetzt, die wichtige Hinweise zum Einsatz multimodaler Ressourcen hinsichtlich der Handlung liefern. In dem Kapitel der handlungsübergreifend auftretenden multimodalen Gestalten (vgl. Kap. 8.6) erfolgt ihre Einordnung nicht über eine spezifische argumentative Handlung, sondern über die rekonstruierten, übergeordneten pragmatischen Funktionen (epistemische Modalisierungen), für die sie funktional eingesetzt werden.

Nach einer kurzen Beschreibung der rekonstruierten multimodalen Gestalten sowie der Funktionen, die ihnen für das mündliche Argumentieren und den Aufbau argumentativer Handlungen zukommen, werden sie dann anhand ausgewählter Sequenzen der mündlichen Argumentationen veranschaulicht. Da die multimodalen Gestalten in ihren Konfigurationen (u. a. in den beteiligten Ressourcen sowie auch der zeitlichen Entfaltung) durchaus einen Grad an Variabilität aufweisen, werden z. T. mehrere Sequenzen vorgestellt, die diese transparent machen und zugleich die Ähnlichkeit der jeweiligen Gestalten darlegen. Aus der Beschreibung der multimodalen Gestalt heraus erfolgt ebenfalls eine Einordnung, ob sie performative, modalisierende, äusserungsstrukturierende und/oder interaktionsorganisatorische Funktionen für die argumentative Handlung übernehmen (vgl. Kap. 3.2.2).

8.1 Position beziehen

Das Beziehen einer Position stellt für Interagierende in argumentativen Entscheidungsfindungen eine etablierte Handlung dar, mit der sie eigene Dis-/Präferenzen gegenüber den zur Auswahl stehenden Optionen benennen, diese zur weiteren Aushandlung bereitstellen oder aber auch Optionen von der weiteren Entscheidungsfindung ausschließen können (vgl. Kap. 2.2.2; vgl. auch Grundler et al. 2020). In den vorliegenden Daten meint dies, dass die Interagierenden eine Präferenz gegenüber einem zur Auswahl stehenden Ziel der Abschlussfahrt einbringen. Dabei ist die argumentative Handlung i. d. R. satzförmig organisiert. Sie kann aber durch Begründungen plausibilisiert und somit zu größeren Äußerungspaketen ausgebaut werden. Abhängig von der epistemischen Modalisierung können Interagierende ihre Position als einen vorläufigen Vorschlag oder aber als eine abgesicherte Entscheidung einbringen (vgl. Heritage 2012b; Heller 2018a; Kreuz und Luginbühl 2020; Morek 2020).

Das vorliegende Kapitel zeigt nun, wie in den argumentativen Entscheidungsfindungen Interagierende innerhalb einer Positionsbekundung das Ziel der Klassenfahrt, auf das sie sich beziehen, durch deiktische Gesten auf das Material (vgl. Kap. 6.2) sichtbar machen und somit das Bezugsobjekt der eigenen Dis-/Präferenz verstärkt hervorheben (vgl. Kap. 8.1.1). Darüber hinaus können Interagierende durch *Throwing Away* bzw. *Brushing Away* Gesten (vgl. Bressemer und Müller 2014a; vgl. Kap. 3.2.2.2) Optionen aus der weiteren Entscheidungsfindung *ausschließen* (vgl. Kap. 8.1.2).

8.1.1 Hervorheben von Optionen beim Beziehen von Positionen

Es zeigt sich in den vorliegenden Daten, dass an der Realisierung der argumentativen Handlung des Position-Beziehens wiederkehrend multimodale Gestalten beteiligt sind, die zu einer spezifischen *Ausprägung* der argumentativen Handlung führen. So nutzen die Interagierenden beispielsweise die multimodale Gestalt des Hervorhebens einer Option innerhalb des Beziehens einer Position, bei der die Sprecher*innen ihre Position, entweder als Stellungnahmen (z. B. *wir nehmen X*) oder als Bewertungen (z. B. *ich finde X schön*), in die Interaktion einbringen und gestisch auf die entsprechende Option als abstrakten Referenten (vgl. Clark 2003)⁶⁸ ihrer Positionsbekundung verweisen. Dabei wird die deiktische Geste auf das Material zeitlich aligniert mit der Benennung der Option (z. B. *Rom*). Die Geste übernimmt hier allerdings nicht primär die Funktion, zur Lokalisierung und Identifizierung eines (ambigen) Bezugsgegenstandes beizutragen (vgl. Stukenbrock 2015); vielmehr kommt es zu einer doppelten, d. h. semantisch redundanten Bezugnahme, indem die Option gleichzeitig benannt und auch eine gestische Referenz hergestellt wird, so dass die Option als Teil der Positionsbekundung verstärkt hervorgehoben wird (vgl. Kendon 2004; Colletta et al. 2010). Die Geste übernimmt also eine pragmatische, äußerungsstrukturierende Funktion (vgl. Kap. 3.2.2.1; vgl. auch Kendon 2017: 169).⁶⁹ Die nachfolgende Tabelle 8 enthält eine Übersicht über die an der multimodalen Gestalt beteiligten Ressourcen sowie ihre pragmatische Funktion. Weitere Ressourcen, wie Fokusakzentuierungen oder auch iterative Bewegungen der Geste (wie z. B. Tippen), tragen zu der Exponierung der verhandelten Option ebenfalls bei. Sequenziell kann das Beziehen einer Position, bei dem eine Option durch multimodale Ressourcen modalisierend beson-

⁶⁸ Es handelt sich insofern um ein Referenzieren auf einen abstrakten Referenten bzw. ein abstraktes Konzept, als dass der*die Sprecher*in nicht auf den Bungalowpark als konkreten Ort verweisen kann, sondern nur auf das abstrakte, auf dem Material schriftlich notierte Abbild bzw. Konzept dieses Ziels der Abschlussfahrt (vgl. u. a. Clark 2003).

⁶⁹ Vgl. zum Einführen eines neuen Referenten in Kombination mit dem Lokaladverb *hier* die Untersuchung von Mondada (2014a): Das turninitiale Lokaladverb, das gemeinsam mit einer deiktischen Geste geäußert wird, ermöglicht Interagierenden in Beschreibungen und Erklärungen eine frühzeitige Orientierung auf ein Zeigeziel hin, das anschließend benannt und für die Interaktion relevant gemacht wird.

ders hervorgehoben wird, sowohl initiiierend, z. B. einer Begründung vorauslaufend, als auch resümierend, z. B. eine Begründung lokal oder eine längere gemeinsame Aushandlung global abschließend, auftreten.

Hervorheben von Optionen im Beziehen von Positionen	
Beschreibung der multimodalen Gestalt	Benennen der Position (z. B. durch Stellungnahme, Bewertung) + Benennen der Option (X) + deiktische Geste auf das Material

Tabelle 8: Beschreibung der multimodalen Gestalt „Hervorheben von Optionen im Beziehen von Positionen“

Die nachfolgenden drei Beispiele illustrieren das Initiieren einer bewertenden Position mit anschließender Begründung (vgl. Beispiel 1), das lokale Resümieren einer Position, das auf eine Begründungssequenz folgt (vgl. Beispiel 2) sowie das globale Resümieren einer (gemeinsamen) Position (vgl. Beispiel 3). Das letzte Beispiel diskutiert ferner die Frage, inwiefern zeitlich vorauslaufende Gesten als eine Variante der multimodalen Gestalt des Hervorhebens von Optionen in (resümierenden) Positionen verstanden werden kann.

In dem ersten Beispiel nimmt die Sprecherin initiativ eine Position ein, die sie anschließend durch eine Begründung plausibilisiert. Sie markiert in ihrer bewertenden Position eine deutliche Dispräferenz gegenüber dem vom Sohn favorisierten Ziel des Bungalowparks. Während die Sprecherin Position bezieht, hebt sie das Bezugsobjekt ihrer Bewertung mittels einer Geste als Teil eines größeren Ressourcenverbundes auf das Material sichtbar hervor.

Beispiel 1 – Hervorheben von Optionen beim Beziehen von Positionen



158 MU: also |mit dem |BUN |galow|park das find ICH total
 mu-gz >>blickt auf Blatt----->>
 mu-gs |.....|tippt mit rKlFi auf Blatt|,,,,,|
 so-gz >>blickt auf Blatt----->>
 langweilig.
 159 da habt_a ja nur (ho) |nur (.) NUR freizeit;
 160 da macht ihr ja irgendwie GAR nichts [ge]scheites.
 161 SO: [mja.]
 162 (---)
 163 das IS (da ja [das gute)]

Die Sprecherin initiiert eine negative Bewertung (Z. 158), die zugleich durch das fokusakzentuierte Personalpronomen in der bewertungsmarkierenden formelhaften Phrase „find ICH“ (Z. 158; vgl. auch Hrnal 2018, 2020) als Opposition erkennbar wird; sie richtet sich gegen das vom Sohn favorisierte Ziel der Abschlussfahrt, den Bungalowpark. Die Bewertung wird turnfinal mit stark fallender Tonhöhenbewegung durch das Prädikativkomplement (Z. 158: langweilig, vgl. Hrnal 2020; vgl. auch Grundler et al. 2020) expliziert und, die Ablehnung der Sprecherin verstärkend, durch das Adjektiv *total* zusätzlich gesteigert. Dabei kontextualisieren der verwendete Indikativ (Z. 158: find ich) sowie das Ausbleiben

epistemisch abschwächender Modalisierungen die Äußerung als eine feststellende, informierende Äußerung – die Position wird nicht als verhandelbar eingebracht (vgl. zu *informings* Heritage 1984). Innerhalb der negativen Bewertung wird die Option, auf die die Sprecherin sich bezieht, dann sowohl durch die Stellung im Vorfeld (Z. 158: also mit dem BUNgalowpark) als auch durch die Fokusakzentuierung sowie durch die gestische Bezugnahme turninitial prominent gesetzt.

Während die Sprecherin nun Position bezieht, zeigt sich eine feinkörnige Koordinierung (vgl. Mondada 2014b) der multimodalen Ressourcen: Der Blick der Interagierenden ist bereits auf das Material ausgerichtet, während die Sprecherin ihre Position einleitet (Z. 158) und mit Beginn der Präpositionalphrase auch die Vorbereitung der referenzherstellenden Geste einsetzt. Der Gestenhöhepunkt der auf das Material verweisenden Geste wird dann mit dem fokusakzentuierten Kopf der eingebetteten Nominalphrase (Z. 158: BUNgalowpark; vgl. #1) aligniert. An dieser Stelle treffen die Ressourcen in einem „multimodalen Verdichtungsraum“ (Stukenbrock 2015: 32) zusammen, so dass durch die zeitliche Alignierung der multimodalen Ressourcen die Komponenten mit dem höchsten Informationsgehalt in den Fokus gestellt wird (vgl. Stukenbrock 2018: 40). Dabei stellen die sprachliche und gestische Ressource eine Bezugnahme auf das *gleiche* Objekt her. Die Ressourcen *ergänzen* sich hier nicht gegenseitig in der übermittelten Information, sondern sie weisen eine semantische Redundanz auf (vgl. Kendon 2004) und verstärken sich somit gegenseitig (vgl. Colletta et al. 2010: 569). Dieser Befund schließt auch an Studien wie Enfield et al. (2007) sowie Enfield (2009) an, die beobachten, dass deiktisch-referenzierende Gesten in ihrem Form- und Variantenreichtum unterschiedliche, auch pragmatische, Funktionen erfüllen können: Werden sie mit wenig Aufwand, d. h. mit weniger ausladenden Bewegungen, produziert und wird zugleich der Referent eindeutig benannt, so können sie als „backgrounded modifier“ (Enfield et al. 2007: 1738) die Herstellung einer Referenz unterstützen; sie sind aber dann nicht unbedingt für Bedeutungskonstitution der Äußerung notwendig.⁷⁰ Auch in den vorliegenden Daten zeigt sich, dass die Gesten aufgrund der konkret benannten Referenten nicht als notwendiges Mittel für die Bedeutungskonstitution fungieren, sondern v. a. pragmatisch relevant werden, indem die Interagierenden die Option, auf die sie sich in ihrer Position beziehen, für alle Beteiligten auch gestisch sichtbar hervorheben. Durch das wiederholte Tippen kommt der multimodalen Gestalt in Kombination mit *total* eine modalisierende Funktion, denn sie hebt verstärkt den Referenten der Positionsbekundung hervor, der hier abgelehnt wird. Indem zugleich der fokusakzentuierte Referent aus dem umgebenden Äußerungskontext hervorgehoben wird, übernimmt die multimodale Gestalt auch eine äusserungsstrukturierende Funktion.

Das zweite Beispiel legt dar, wie eine Sprecherin innerhalb eines größeren Äußerungspaketes eine bereits einleitend benannte Position im Anschluss an eine Begründung wiederholend benennt und dadurch die Begründungssequenz resümierend abschließt. Dabei verweist die Sprecherin mit Benennung des Referenten gestisch auf das Material als semiotische Ressource (vgl. Goodwin 2000) und hebt die Option des Segeltörns für beide Interagierenden sichtbar hervor.

Beispiel 2 – Hervorheben von Optionen beim Beziehen von Positionen

174 MU: also ich GLAU:be, (.)
 175 dass das ne ganz (.) tolle SAche is. (.)
 176 um: (.) um mal zu sagen das is n besonderes erLEBnis==
 177 =ERSmal (.) hast du (m/n) macht das ja auch nich JEder IMmer,
 178 (0.3)

⁷⁰ Enfield et al. (2007: 1733) sowie Enfield (2009: 104) beschreiben ebenfalls, dass die weniger aufwendig produzierten Gesten mit deiktischer Funktion unterstützende Informationen liefern können, wenn der*die Sprecher*in annimmt, dass der benannte Referent von dem*der Rezipient*in identifiziert werden kann – jedoch nicht gänzlich sicher ist. Der Einsatz der weniger aufwendigen, kleineren deiktischen Gesten kann somit, falls nötig, zur Identifizierung des Referenten beitragen und zugleich wird das Risiko des „over-telling“, d. h. „being more explicit than was actually necessary“, minimiert (Enfield 2009: 104).

179 ne,=guck ma ich denke so HAM:burch oder RO:M-
 180 oder CENTER parc;
 181 kann schon der ein oder andere HIN,
 182 [oder] is [da schonmal geWesen-]
 183 SO: [mhmh]
 184 [der EIN oder Andere;]



185 MU: |aber |so_n SEgel |törn |find ich is schon (1.0) TOLle
 mu-gs |.....|fährt mit rIF über Option|,,,,,
 mu-gz >>blickt auf Blatt (>Z. ?)
 so-gz >>blickt auf Blatt----->>
 sache.
 186 SO: hm ja GUT;=
 187 =das seh ich AUCH so.

In der vorliegenden Sequenz schließt die Position der Sprecherin (Z. 185) eine global strukturierte Begründung resümierend ab (Z. 174-182). Dabei markiert der initiale Adversativjunktorkor *aber* eine Abgrenzung zwischen den in der Begründung besprochenen Optionen (Hamburg, Rom, Center Parc) und dem von der Sprecherin favorisierten Segeltörn. Die Bewertung erfolgt hier nun in einer anderen morphosyntaktischen Realisierung als im ersten Beispiel, da sie in Form eines abhängigen Hauptsatzes realisiert wird, bei dem die formelhafte Phrase *find ich* eingeschoben eine anschließende Evaluation (Z. 185: *is schon TOLle sache.*) projiziert. Auch in dieser (resümierenden) Position treffen die sich entfaltenden Ressourcen zeitlich koordiniert an einem spezifischen Moment aufeinander: Während beide Interagierende bereits mit ihrem Blick auf das Material hin orientiert sind, ‚unterstreicht‘ die Sprecherin mit der Benennung des fokusakzentuierten Kopfes der Nominalphrase „so_n SEgeltörn“ (Z. 185) mit dem Indexfinger die abgebildete Option (vgl. #1, #2). Auch hier wird die Option, auf die die Sprecherin sich in ihrer *resümierenden* Position bezieht, durch die Verwendung semantisch redundanter Ressourcen gemeinsam mit der prosodischen Markierung des Fokusakzents besonders hervorgehoben.

Anhand des dritten Beispiels wird nun illustriert, wie Sprecher*innen durch aufeinanderfolgende Stellungnahmen zum einen Zustimmung signalisieren und zum anderen den Kernjob Begründen (vgl. Kap. 2.2.1) resümierend abschließen können. Im Unterschied zu den bislang untersuchten Sequenzen läuft im vorliegenden Beispiel die gestische Ressource der sprachlichen Benennung des Bezugsobjektes voraus. Dies wirft die Frage auf, ob und inwiefern dadurch die hervorhebende Funktion innerhalb der multimodalen Gestalt erhalten bleibt.

Beispiel 3 – Hervorheben von Optionen beim Beziehen von Positionen

193 VA: [hier] müssen alle mit ins RAD packen;
 194 SO: aber [wenn sich dann EIN]zelne dagegen stellen dann funktioniert das ja nich.
 195 VA: [jeder muss ()]
 196 (--)
 197 VA: DIE muss die gruppe halt mitneh|men.
 va-gz >>blickt auf Blatt-----|blickt zu SO->
 so-gz >>blickt zu VA----->

198 (0.2) | (0.3) | (0.2)
 so-gz ---->|auf Blatt----->>
 va-gz ----->|auf Blatt->>
 199 SO: |ja. |
 so-he |nickt|
 200 joa dann |(-) wär ich AUCH |für den segeltörn.
 so-gs |tippt mehrmals m. rKlFi|
 201 (--)
 202 VA: ne?
 203 ein_n tach MEHR.=
 va-gz >>blickt auf Blatt->>
 so-gz >>blickt auf Blatt->>
 204 =guck mal das ROM wär nur drei tAge,
 205 hamburg VIER tage.
 206 (1.0)
 207 segeltörn FÜNF tage. (-)
 208 oKAY.
 209 wir entscheiden uns also für den |SEgel |törn;
 va-gs |tippt silb. m. Stift|
 210 FÜNF tage;
 ((...))
 212 SO: ja.

Nach einer ersten Opposition gegenüber der Begründung des Vaters für den Segeltörn (Z. 194), stimmt der Sohn diesem nun zu (Z. 199: ja.) und bringt eine abgeschwächt modalisierte, resümierende Position für den Segeltörn ein (Z. 200: joa dann (-) wär ich AUCH für den segeltörn.). Dass diese als Folge aus der vorangegangenen Aushandlung resultiert, wird durch das Adverb *dann* markiert.

In Bezug auf die bislang beobachtete zeitliche Alignierung der gestischen und sprachlichen Ressourcen, mit denen eine Option hervorgehoben wird, zeigt sich in der zeitlichen Entfaltung der Ressourcen des vorliegenden Beispiels eine Divergenz: Die referenzherstellende Geste wird nicht auf dem Kernelement der Position (Z. 200: segeltörn) geäußert, sondern mit der Zustimmung ausdrückenden fokusakzentuierten Partikel *auch* koordiniert – die Geste wird noch vor Benennung der Option in eine Ruheposition zurückgeführt. Inwiefern kann nun dieser Ressourcenverbund in seiner zeitlichen Koordinierung die gleiche Funktion wie die multimodale Gestalt der bereits angeführten Beispiele erfüllen? Berücksichtigt man nun weitere Sequenzen im Datenkorpus, die eine solche zerdehnte zeitliche Entfaltung aufweisen, so zeigt sich, dass in diesen Sequenzen die gestische Ressource konsequent dann geäußert wird, wenn bereits sprachlich durch bewertungseinleitende Phrasen (*finde/fände ich*) oder die Explikation einer individuellen Entscheidung (*wäre ich für*) eine Position projiziert worden ist. Die referenzherstellende Geste auf das Material wird dann i. d. R. in der Nähe zum Kopf der Nominalphrase geäußert; sie tritt immer vor der Benennung der Option – nie nachher – auf. Die Geste ermöglicht dem Gegenüber also zwar, wie es typisch für Gesten mit referenzherstellender Funktion ist, eine frühzeitige Lokalisierung und Identifikation des Bezugselements, das von dem*der Sprecher*in in der Position anvisiert wird. Und dennoch dient die Geste nicht einer Zeigehandlung, deren Funktion es ist, „einen gemeinsamen perzeptorischen und kognitiven Aufmerksamkeitsfokus zwischen Sprecher und Adressatin (auf etwas Drittes) herzustellen“ (Stukenbrock 2015: 34f.), denn sie wird nicht im Verbund mit einem deiktischen Ausdruck (wie z. B. *fände ich hier den segeltörn schön*) geäußert, der einen sprachlichen Hinweis auf eine Zeigehandlung enthält und „[k]onstitutiv für deiktische Zeigehandlungen“ (Stukenbrock 2015: 65) ist. Sequenziell ist die zeitlich zerdehnte Entfaltung sprachlicher und gestischer Ressourcen v. a. während des resümierenden Beziehen einer Position zu beobachten, in denen die Option i. d. R. nicht als neues oder unbekanntes Element das erste Mal eingeführt, sondern dem Gegenüber bereits bekannt ist oder aufgrund der vorangegangenen Interaktion inferiert werden kann. Indem der*die Sprecher*in anschließend die bereits projizierte Option auch sprachlich expliziert, schließt der*die Sprecher*in nun die

gesamte Gestalt der argumentativen Handlung des Hervorhebens von Optionen im Beziehen einer Position ab (vgl. zum *Gestaltschluss* Stukenbrock 2018: 44). Insgesamt kann also festgehalten werden, dass die referenzherstellende Geste auf das Material zwar eine frühzeitige Lokalisierung und Identifizierung des Referenten ermöglicht, allerdings nicht, wie in Zeigehandlungen üblich (vgl. Stukenbrock 2015), sprachlich auf sie verwiesen wird; da die multimodale Gestalt sich aber weder in ihrer sprachlichen Ausgestaltung noch den weiteren beteiligten Ressourcen verändert und der Referent weiterhin gestisch sowie sprachlich eine doppelte Bezugnahme aufweist, wird angenommen, dass trotz der temporalen Zerdehnung die semantische Redundanz und somit hervorhebende Funktion nicht aufgegeben wird und es sich um eine Variante der multimodalen Gestalt des Hervorhebens einer Option im Beziehen von Positionen handelt.

Die o. g. Sequenz wird dann durch eine global abschließende Position (Z. 209) des Vaters beendet, in der er ebenfalls mit der Benennung der Option diese durch eine silbisch tippende Geste auf das Material hervorhebt. Die resümierende Position des Vaters schließt hier, affirmiert durch den Sohn (Z. 212), den Kernjob Begründen ab.

Zusammenfassend wird anhand der angeführten Beispiele Folgendes sichtbar: Die multimodale Gestalt des Hervorhebens einer Option im Beziehen von Positionen kann in Form von Stellungnahmen (*ich bin für X*) oder Bewertungen (*ich finde X schön*) eingebracht werden. In diesen drücken Interagierende ihre Dis-/Präferenz gegenüber einer von ihnen benannten Option (*X*) aus. Sequenziell können Positionen initiiierend (z. B. mit anschließend plausibilisierender Begründung) oder resümierend (z. B. im Anschluss an eine Begründung) geäußert werden. Die Untersuchung zeigt, dass Interagierende dabei während des Beziehens einer Position explizit die Option, auf die sie sich beziehen, benennen und dabei zeitlich koordiniert auf ebenjene Option gestisch verweisen. Dadurch wird das Material als semiotische Ressourcen (vgl. Goodwin 2000) für die argumentative Handlung relevant gemacht. Sind die multimodalen Ressourcen zeitlich so miteinander koordiniert, dass sie in einem multimodalen Verdichtungsraum zusammenlaufen, können sie gemeinsam die Komponenten mit dem höchsten Informationsgehalt in den Fokus stellen (vgl. Stukenbrock 2018). Ferner konnte eine Variante der multimodalen Gestalt rekonstruiert werden, in der die referenzielle gestische Ressource der sprachlichen Benennung des Referenten vorausläuft – diese zeigt sich vornehmlich in resümierenden Positionsbekunden. Sowohl die sprachliche als auch die gestische Bezugnahme auf die Option der eingenommenen Position führen zu einer semantischen Redundanz (vgl. Kendon 2004: 176). Dadurch wird die Option, auf die die Interagierenden sich beziehen, gemeinsam mit weiteren Ressourcen, wie prosodischen Markierungen, verstärkt hervorgehoben. Indem die multimodalen Ressourcen die verhandelte Option exponieren, tragen sie sowohl zur Modalisierung Positionsbekundung sowie zur Strukturierung der Äußerung bei.

8.1.2 Position beziehen durch Ausschließen einer Option

Interagierende können in argumentativen Entscheidungsfindungen auch Position beziehen, indem sie das Verwerfen bzw. Ausschließen einer Option anzeigen. Dies kann beispielsweise sprachlich durch Explikationen wie *ich würde sagen X auf keinen fall* erfolgen. Das Ausschließen einer Option kann allerdings nicht nur sprachlich benannt werden, sondern auch durch den Einsatz körperlich-sichtbarer Ressourcen markiert werden. Interagierende nutzen dafür in der vorliegenden Untersuchung *Throwing Away* Gesten, bei denen die Hand aus dem Handgelenk heraus abknickend fallengelassen wird (vgl. Bressemer und Müller 2014a: 1581), bzw. *Holding Away* Gesten, bei denen Sprecher*innen die flache Hand abgewandt vom Körper halten „as if holding or pushing away an object“ (Bressemer und Müller 2014a: 1582). Aufgrund ihrer Bewegungsrichtung, bei der sich die Hand grundsätzlich weg vom Körper des*der Sprecher*in bewegt, werden beide Gesten der Gestenfamilie der *Away* Gesten zugeordnet (vgl. Bressemer und Müller 2014a, 2014b; vgl. auch Kap. 3.2.2), deren pragmatische Funktion darin besteht,

„[g]etting rid of, removing and dismissing annoying topics of talk“ (Bressemer und Müller 2014a: 1581).⁷¹ Die Gesten zeigen also eine Sprechhandlung sichtbar an, wie auch aus der nachfolgenden Übersicht 9 zur Beschreibung der multimodalen Gestalt hervorgeht.

Position beziehen durch Ausschließen einer Option	
Beschreibung der multimodalen Gestalt	Benennen der Option + <i>Throwing</i> bzw. <i>Holding Away</i> Geste (optional: sprachliche Explikationen des Ausschlusses)

Tabelle 9: Beschreibung der multimodalen Gestalt „Position beziehen durch Ausschließen einer Option“

Anhand der nachfolgenden Beispiele wird dargelegt, wie Interagierende durch den Einsatz von *Throwing Away* Gesten (Beispiel 1) und *Holding Away* Gesten (Beispiel 2) eine Option als verworfen markieren. Da beide Gestenformen die gleiche performative pragmatische Funktion übernehmen, indem sie das Ausschließen einer Option anzeigen, werden sie in dem vorliegenden Kapitel gemeinsam behandelt.

In dem folgenden Beispiel wird die Option ‚Bungalowpark‘ von der Sprecherin durch eine *Throwing Away* Geste aus der weiteren Verhandlung ausgeschlossen.

Beispiel 1 – Position beziehen durch Ausschließen einer Option

```

068   TO:   ich würd ma sagen V:IERte möglichkeit auf keinen [fall,]
069   MU:                                     [also ]
        |das hab ich mir schon [(geDA:/geNAU:) gedacht;]
mu-gz  |blickt auf Blatt (>Z. 073)
070   TO:   [((lacht)) ]

```



071	MU:	VIERte is sch	auf <<pp>(keinen)>	]
	mu-gs	1H Throwing Away	,,,,,,,,,,,,,,,,,,,,,	
072	TO:			[(räuspert sich)]
073	MU:	ZWEIte auch nich,		
	mu-gz	-----> blickt zu TO->>		

In einer ersten Präferenzäußerung bringt die Tochter initiiierend den Ausschluss des Bungalowparks ein. Die dabei eingesetzte Modalisierung (Z. 068: auf keinen fall) wird mit Pomerantz (1986) als sogenannte *Extreme Case Formulation*. Die initiiierende Bewertung der Tochter wird mit absoluter epistemischer Gewissheit vorgebracht. In ihrer interaktiven Konsequenz können *Extreme Case Formulations* dann dazu führen, dass andere Beteiligte die vorangegangene Äußerung nicht allein durch Rezeptionsbekundungen oder Affirmationspartikeln (zustimmend) zur Kenntnis nehmen, sondern eigene Beiträge, z. B. in Form von Bewertungen anschließen. Zweite, gleichlaufende Bewertungen werden dabei i. d. R. ver-

⁷¹ Die von den Autorinnen gewählte Beschreibung der Funktion der Geste zeigt dabei deutliche Bezugspunkte zu einem praxeologischen Verständnis (vgl. insbesondere Streeck 2009, 2015), das eine Nähe von manuellen Handlungen (wie dem Fernhalten spezifischer Objekte vom eigenen Körper) sowie der Bedeutung und Funktion von Gesten annimmt (vgl. auch Teßendorf 2014).

stärkend eingebracht (vgl. Auer und Uhmann 1982). In der abgebildeten Sequenz verwirft nun die Mutter (Z. 071), konform mit der Präferenzäußerung der Tochter, die Option des Bungalowparks in einer eigenen argumentativen Handlung: Sie wiederholt dazu das prosodisch hervorgehobene Bezugselement ihrer Position (Z. 071: VIERte), hebt mit der Fortsetzung der Äußerung (Z. 071: is) die linke Hand mit der Handfläche nach unten orientiert an (vgl. #1) und lässt sie, mit dem sprachlichen Abbruch (Z. 071: sch), in einer *Throwing Away* Geste fallen (vgl. #2; vgl. Bressemer und Müller 2014a). Die Geste zeigt hier zu einem sehr frühen Zeitpunkt der Äußerung der Mutter, nämlich unmittelbar nach Benennung der Option, das Ausschließen ebenjener an. Auf sprachlicher Ebene hingegen verbleibt die argumentative Handlung durch die sehr leise, teils unverständliche Wiederholung der epistemischen Modalisierung „auf keinen (fall)“ (Z. 068, Z. 071) äußerst vage. Möglicherweise liegt dies darin begründet, dass die Position der Tochter für die Mutter eine ‚erwartbare‘ Position bildet, wie in Zeile 069 (das hab ich mir schon [(geDA:/geNAU:) gedacht;]) erkennbar wird. Wenngleich im vorliegenden Beispiel die Präferenzäußerung der Tochter auf sprachlicher Ebene nicht gesteigert wird, zeigt sich aber in der gestischen Ressource ein verstärkendes Element, denn die Mutter schließt durch die *Throwing Away* Geste die benannte Option demonstrativ aus dem weiteren Diskurs aus. In dem vorliegenden Beispiel wird also frühzeitig das Verwerfen einer Option und die somit bezogene Position durch die *Throwing Away* Geste mit performativer Funktion erkennbar.

Innerhalb des zweiten Beispiels wird das Ausschließen einer Option von der Sprecherin durch eine *Holding Away* Geste angezeigt.

Beispiel 2 – Position beziehen durch Ausschließen einer Option

041 SO: <<dim, lächelnd>wir gehen zum VIERten;>
 ((...))
 045 MU: AL:so;
 046 VIER tage ham:burg mit mUsicalbesuch.



047 WA|rum |<<p>nicht,> (1.0) |
 mu-gz |...|blickt zu SO-----|
 so-gz >>blickt auf Blatt----->



048 |find |ich AU_nich |gut;=
 mu-gz |.....->
 mu-bo |richtet ObKö auf--|
 mu-he |dreht Kopf nach re----->
 mu-gs |Holding Away|
 so-gz ----->|.....|blickt zur Seite->


```

049      =würd |ich jetzt AU_nich vor|schla[gen.]
mu-gz    .....|blickt zur Seite-----|blickt auf Blatt->>
mu-he    ----->|
so-gz    ----->|blickt auf Blatt->>

```

Nachdem ein globaler Zugzwang der Mutter, mit dem sie ihren Sohn in Bezug auf die Option ‚Hamburg‘ begründungspflichtig macht, unbearbeitet bleibt (Z. 047), verändert sich die gesamte Körperpositur der Mutter (Z. 048, vgl. #1, #2): Sie richtet den Oberkörper auf, löst den Blick vom Sohn und beginnt den Kopf zu drehen (vgl. u. a. Scheflen 1964). Sie übernimmt außerdem nun selbst erneut das Rederecht. Die Mutter bringt eine eigene Position in Form einer Evaluation ein, in der sie die benannte Option (Z. 046: VIER tage ham:burg mit mUsicalbesuch.) ablehnt. Sie stellt hier durch das fokusakzentuierte Adverb *auch* (Z. 048: find ich AU_nich gut;) eine Verbindung zu einer vorangegangenen Präferenzäußerung des Sohnes her (Z. 041), die den Ausschluss aller anderen Optionen, und somit auch der Option ‚Hamburg‘, beinhaltet. Die Bewertung wird nun zeitlich aligniert mit dem Öffnen der linken Hand zu einer kleinräumigen *Holding Away* Geste (vgl. #2), mit der die Sprecherin die Option an sich zurückweist und dadurch die multimodale Gestalt des Beziehens einer Position durch Ausschließen einer Option realisiert.

Fasst man diese Befunde zusammen, so zeigt sich, dass Sprecher*innen Optionen, auf die sie deiktisch und/oder sprachlich Bezug nehmen, im Verbund mit *Throwing Away* sowie *Holding Away* Gesten aus dem weiteren Diskurs ausschließen können. Das Beziehen einer Position durch Ausschließen einer Option kann somit in syntaktisch unvollständigen *Turns*, sofern die Option bekannt ist, auf die sich die Sprecher*innen beziehen, erfolgreich für das Gegenüber angezeigt werden. Die Gesten übernehmen hier in ihrem Ressourcenverbund eine performative, pragmatische Funktion.

8.1.3 Zusammenfassung

Zusammenfassend konnten die folgenden multimodalen Gestalten, die wiederkehrend in der argumentativen Handlung des Beziehens einer Position auftreten, anhand des vorliegenden Kapitels 8.1 rekonstruiert werden:

- Interagierende bringen innerhalb von argumentativen Entscheidungsaufgaben Positionen initierend sowie resümierend ein und können dabei durch referenzherstellende Gesten auf das Material, das als semiotische Ressource (vgl. Goodwin 2000) genutzt wird, das Bezugselement ihrer Position hervorheben (vgl. Kap. 8.1.1). Der Referent wird dabei sowohl sprachlich als auch gestisch, in semantischer Redundanz (vgl. Kendon 2004: 176), eingeführt. Da sich in den ermittelten Sequenzen allerdings keine deiktischen Ausdrücke finden, die gemeinsam mit der Geste als „[k]onstitutiv für die deiktische Zeigehandlung“ gelten (Stukenbrock 2015: 65), kann angenommen werden, dass die Geste nicht primär einer *Zeigehandlung* dient, in der Sprecher*innen bewusst die Aufmerksamkeit des Gegenübers auf etwas Drittes lenken wollen. Vielmehr dient die Herstellung gestischer Referenz gemeinsam mit der sprachlichen Benennung der Option dazu, das Kernelement einer Position besonders hervorzuheben und dadurch epistemische Gewissheit zu markieren – dabei ist die Geste nicht notwendigerweise erforderlich für die Bedeutungskonstitution (vgl. 8.1.1, Beispiel 1).⁷² Es zeigt sich in der zeitlichen Koordination der an der argumentativen Handlung beteiligten multimodalen Ressourcen, dass diese

⁷² Dass eine Gestenäußerung zeitgleich mehrere Funktionen übernehmen kann, darauf weist Kendon (2004: 203f.) beispielsweise in der Beschreibung von Zeigehandlungen hin, die mit einer geöffneten Handfläche realisiert werden: Indem der*die Sprecher*in mit der geöffneten Handfläche zeigt, kann einerseits das Zeigziel identifiziert werden (deiktische Funktion) und zugleich der Referent bzw. das Bezugsobjekt auf der geöffneten Handfläche präsentiert werden (pragmatische Funktion) (vgl. aber auch McNeill 1992).

in einem multimodalen Verdichtungsraum (vgl. Stukenbrock 2015) zusammenlaufen können, sodass die Komponenten mit dem höchsten Informationsgehalt, d. h. die Kernelemente der Positionen (gemeint sind die jeweiligen Optionen), im Fokus stehen. Schließlich können die initiierenden Positionsbekundungen Ausgangspunkt für ein größeres Äußerungspaket bilden, in dem die Position plausibilisierend begründet wird, oder auch konklusiv Begründungssequenzen resümierend abschließen.

- Ferner konnte gezeigt werden (vgl. Kap. 8.1.2), wie Sprecher*innen durch das gestische Ausschließen einer Option Position beziehen: So können einmal eingeführte Optionen durch *Throwing Away* sowie *Holding Away Gesten* als ausgeschlossen markiert werden, auch wenn dies sprachlich nicht oder nur implizit benannt wird. Dabei übernehmen die Gesten innerhalb der multimodalen Gestalt eine performative pragmatische Funktion, denn sie zeigen den Abschluss der Option für alle Beteiligten sichtbar an.

8.2 Begründen

Interagierende können durch Begründungshandlungen z. B. eingebrachte Positionen (vgl. Kap. 8.1) für ihr Gegenüber plausibilisieren oder, ohne dass eine Position explizit benannt wird, erste Präferenzen des*der Sprecher*in hinsichtlich der zur Auswahl stehenden Optionen anzeigen. Diese können dann zu größeren Äußerungspaketen aufgebaut werden, indem Interagierende beispielsweise im Rahmen der argumentativen Entscheidungsfindung, die ein hypothetisches Ereignis betrifft, imaginierte Szenarios aufbauen oder veranschaulichende Beispiele einbringen. Doch Begründungen können auch, i. S. von nicht elaborierten Begründungen (vgl. Kap. 2.2.2), lokal strukturiert sein und z. B. allein eine bewertende Proposition beinhalten.

Dabei zeigen sich in der Realisierung der argumentativen Handlung des Begründens wiederkehrende multimodale Gestalten. So können Interagierende beispielsweise mittels einer gehaltenen *Palm Up* Geste (komplexere) Begründungen sowie einzelne, für die Begründung relevante Lexeme präsentieren (vgl. Kap. 8.2.1). Darüber hinaus halten Interagierende in der Begründung berücksichtigte Gesprächsgegenstände durch deiktische Gesten auf das Material für alle Beteiligten präsent (vgl. Kap. 8.2.2), führen mittels *Palm Up* Gesten Turnfortsetzungen ein (vgl. Kap. 8.2.3) oder zeigen durch Gestenwiederholungen multimodale Kohäsion und Kohärenz an (vgl. Kap. 8.2.4). Durch Gesten der Familie der *Away* Gesten können Interagierende aber auch das Verwerfen von Optionen in Begründungen anzeigen (vgl. Kap. 8.2.5).

8.2.1 Präsentieren während des Begründens

In argumentativen Entscheidungsfindungen können Begründungen von den Interagierenden eingebracht, zur gemeinsamen Betrachtung auf der geöffneten Handfläche präsentiert und zur gemeinsamen Ansicht angeboten werden (vgl. Kendon 2004; Müller 2004; Heller 2018a; Schönfelder und Heller 2019). Das abstrakte diskursive Objekt wird dabei nun als „concrete manipulable entity“ angeboten, „to join the perspective offered on the speaker’s open hand“ (Heller 2018a: 274). Mit einem turnfinalen Halten der Geste kann der präsentierte Gesprächsgegenstand auch für die Fortsetzung der Interaktion und somit für die Reaktion der anderen Beteiligten relevant gemacht werden (vgl. Streeck 2007; Schönfelder und Heller 2019) – erst wenn das diskursive Objekt bearbeitet worden ist, können die Beteiligten sich weiteren Gesprächsgegenständen zuwenden.

In den vorliegenden argumentativen Entscheidungsfindungen nutzen Interagierende zum Aufbau der argumentativen Handlung des Begründens unterschiedliche Konfigurationen ihrer multimodalen, d. h. sprachlichen wie körperlich-sichtbaren Ressourcen. So können je nach Konfiguration Begründungen oder Begründungssegmente auf der nach oben geöffneten Handfläche zur gemeinsamen Betrachtung angeboten werden (vgl. Kap. 8.2.1.1) oder aber einzelne Lexeme als relevante Elemente der Begründung (vgl. Kap. 8.2.1.2) *kurzzeitig, punktuell* präsentiert und somit innerhalb der Äußerung sichtbar hervorgehoben werden. Die *Palm Up* Geste übernimmt dabei in ihrem Ressourcenverbund sowohl performative, äußerungsstrukturierende als auch interaktionsorganisatorische Funktionen.

8.2.1.1 Präsentieren von (komplexeren) Begründungen

Interagierende können Begründungen bzw. Anteile von Begründungen, i. S. von Begründungssegmenten, als abstrakte diskursive Objekte (vgl. Müller 2004) auf der geöffneten Hand präsentieren. Die gemeinsam mit dem Begründungssegment bzw. der Begründung geäußerte *Palm Up* Geste wird dabei nicht direkt nach dem Gestenhöhepunkt (*Stroke*) wieder in ihre Ruheposition (*Recovery*) zurückgeführt, sondern in einem *Post-stroke Hold* (Kendon 2004: 112) gehalten. Somit wird das Element der Begründung bzw. die Begründung selbst als präsentiert kontextualisiert. Eine Übersicht über die rekonstruierte multimodale Gestalt des Präsentierens von komplexeren Begründungen gibt die nachfolgende Tabelle 10.

Präsentieren von (komplexeren) Begründungen	
Beschreibung der multimodalen Gestalt	Begründung bzw. Begründungssegment(e) + <i>Palm Up</i> Geste + <i>Post-stroke Hold</i> (+ turnfinales Halten)

Tabelle 10: Beschreibung der multimodalen Gestalt „Präsentieren von (komplexeren) Begründungen“

In dem vorliegenden Kapitel wird dargelegt, dass sowohl global strukturierte (vgl. Beispiel 1, 2) als auch lokal strukturierte (vgl. Beispiel 3) Begründungen bzw. Begründungssegmente auf der geöffneten Handfläche präsentiert werden können. Das turnfinale Halten der *Palm Up* Geste kann ferner eine Reaktion des Gegenübers elizitieren (vgl. Beispiel 2).

In dem ersten Beispiel präsentiert die Sprecherin ihre resümierende Begründung für die Ablehnung der Option ‚Rom‘ auf der geöffneten Handfläche.

Beispiel 1 – Präsentieren von (komplexeren) Begründungen

```

081  SO:  ROM is schon [mal nich schlecht,]
082  MU:  [VAtika:n- h°          ]
083  SO:  nur kultURprogramm is (.) na:j:a;
084      (-)
085  MU:  und SIND auch nur drei tage;=
086      =ne,
087      muss man [DAS ma sehen;          ]
088  SO:  [<<nickend>und es SIND] auch nur drei tage (ja);>
        ((...))

```



```

101  MU:  |aber|[sind nur drei TAge;                ]
      mu-gz |....|blickt zu SO (>Z. 105)
      mu-gs |öffnet rH Palm Up Ri SO (>Z. 105)
102  SO:  [(aber mit_m bus hin/aber xxx müssen;)]
103  MU:  (un) |wenn (man) |DREI                |tage mit_m [bus nach] rOm,=|
      mu-bo |.....|Beat mit ObKö & Kopf-|ObKö bleibt n. vorn gebeugt|
104  SO:  [ja_JA; ]
105  MU:  =is| na|türlich: äh (.) GANZ schön lang.
      mu-gz -->|
      mu-gs -->|,,,|
106  (--)

```

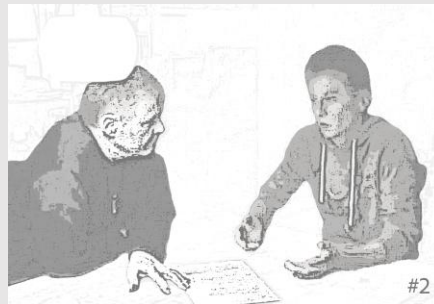
Das ausgewählte Transkript zeigt, dass die Begründungshandlung der Mutter gemeinsam mit einer *Palm Up* Geste realisiert wird (Z. 101ff.). Die Begründungssequenz schließt dabei an eine vorangegangene Begründung der Mutter an (Z. 085), die sie hier in einer Art Rahmenbildung wieder aufnimmt (Z. 101) und dadurch auf das eingangs erwähnte Ausschlusskriterium der Kürze der Abschlussfahrt zurückkommt. Diese anfänglich satzförmige Proposition, i. S. eines Statements (vgl. Goetz und Shatz 1999; Kyratzis et al. 2010), die die Sprecherin mit dem Diskursmarker *aber* einleitet, der als Markierung von Diskontinuität (vgl. Quasthoff et al. 2019: 99) dient, baut die Sprecherin anschließend weiter zu einer Begründung in Form einer *wenn(-dann)*-Konstruktion (Z. 103-105) aus. Dabei öffnet sie bereits in der Proposition turninitial (Z. 101) die rechte Hand zu einer *Palm Up* Geste und präsentiert diese ihrem Gegenüber als abstraktes diskursives Objekt (vgl. Müller 2004). Die Geste überdauert die erste Turnkonstruktionseinheit (Z. 101), die aufgrund ihrer fallenden Intonationsbewegung auch als turnübergaberelevant interpretiert werden könnte – durch das Halten der Geste wird allerdings eine Fortsetzung des Turns projiziert (vgl. zsf. Bohle 2014b). Das Halten der Geste mit dem Gestenhöhepunkt in einem *Post-stroke Hold* (Kendon 2004: 112) bindet nun die Turnkonstruktionseinheiten äußerungsstrukturierend zu einer „semantically complete phrase of speech“ (Kendon 2004: 112) zusammen. Zugleich wird der Sohn durch die Ausrichtung der Handfläche sowie durch die Blickzuwendung und das Vorbeugen des Oberkörpers direkt adressiert. Diese Formen der Involvierung des Gegenübers u. a. durch gestische Ressourcen können anzeigen, „that something they [the speaker, Anm. N. S.] just said is the source of what the current speaker is just now saying [...]“ (Kendon 2004: 273f.) – im vorliegenden Fall nimmt die Mutter Bezug auf die vorangegangene Sequenz, in der beide Interagierenden die Anzahl der Tage als zu kurz befunden haben (nicht im Transkript enthalten). Mit der Ratifizierung der Begründung durch den Sohn (Z. 104) schließt die Sprecherin die Begründung mit einer Konklusion ab (Z. 105) und führt die *Palm Up* Geste in eine Ruheposition zurück (Z. 105).

Die Geste markiert also durch ihre Kontinuität die präsentierten Elemente – Proposition und elaborierende Begründung – als zusammengehörig. Das Halten der *Palm Up* Geste kann somit Anteile größerer Äußerungspakete, i. S. von Begründungssegmenten, zum einen präsentieren sowie zum anderen diese (zugleich) in einer Gesteneinheit (*Gesture Unit*) als zusammengehörig markieren – sie übernimmt somit sowohl performative als auch äußerungsstrukturierende Funktionen (vgl. Kap. 3.2.2.1).

Das nachfolgende Transkript zeigt, wie global strukturierte Begründungssegmente innerhalb eines größeren Äußerungspakets nicht nur präsentiert werden können, sondern durch das turnfinale Halten der *Palm Up* Geste eine Reaktion vom Gegenüber nachdrücklich elizitiert werden kann (vgl. Streeck 2007). Die turnfinal gehaltene Geste übernimmt hier also auch eine interaktionsorganisatorische Funktion (vgl. Kap. 3.2.2.1). Innerhalb eines größeren Äußerungspakets begründet der Sohn in einem *Multi-Unit Turn* seine Ablehnung, die Klassenfahrt nach Rom zu unternehmen. Dabei produziert er innerhalb einer Gesteneinheit mehrere Gestenphrasen (vgl. Kap. 3.1.1), deren beständige Konstituente in einer beidhändigen *Palm Up* Geste besteht, mit der die elaborierende Begründung präsentiert wird.

Beispiel 2 – Präsentieren von (komplexeren) Begründungen

239 SO: ich meine ROM an sich is geil:,
 ((...))
 243 aber (.) das kulTURprogramm werden die meisten absolut ablehnen.
 ((...))
 246 =vielleicht (-) obWOHL kann man so nich sagen;
 247 weil es gibt ja immer so: (.) wieder so: (.) LEUTE die da ja halt da
 rumrennen un:d- (1.0) jede (.) jedes zweite haus: fotografieren.
 248 darum kann auch sein dass die das: DENen das gefällt.
 249 aber die r (.) die: zum beispiel- (0.9)
 250 ich kann mir (au) VORstellen==



251 | =das: (0.5) | (0.2) | pu kul|TUR |
 so-gs |dreht bHde auf, HFn Ri SO|.....|führt lH Palm Up nach
 so-gz |blickt zu VA|
 va-gz |blickt zu SO (>Z. 253)



| (.) |kulTURprogramm |würde ja heißen-
 so-gs außen|....|führt lH Palm Up nach außen|.....->
 so-gz |blickt i. d. Raum----->



```

252      dass wir jetzt zum beispiel ne |FÜH                                |rung|
so-gs      .....->|führt 1H Palm Up nach außen-|Beat|
so-gz      ----->|blickt zu VA----->
           |machen |oder so;
so-gs      |.....->
so-gz      -----|

```



253		da	würd sich (ja) eine neunte klasse zu	TO		da
	so-gs		bewegt bHde lateral	wackelt m.		
	so-gz		blickt zu VA (>Z. 255)			
	va-gz	-----	>			
		langweilen.				
	so-gs	lH lateral vor & zurück				
254		(1.7)		(-)		
	so-gs	hält bHde auf Handkanten aufgestellt	,,,->			
255	VA:	oKAY,				
	so-gs	,,,,,				
	so-gz	---->				
256	SO:	also ich WEISS s nich.				

Nachdem die vorangegangene einlenkende Sequenz abgeschlossen ist (Z. 248), beginnt der Sprecher ein neues, durch den Adversativjunktorkonkurrenz *aber* abgegrenztes Begründungssegment, das anfänglich durch einen Abbruch (Z. 249) und Einsetzen von Häsitationen in Form von Längungen sowie Pausen (vgl. Z. 250ff.) geprägt ist. Zur Überbrückung der Pause (Z. 251) öffnet der Sprecher beide Hände zu einer *Palm Up* Geste (vgl. #1), die hier markiert, dass der *Turn* nicht abgebrochen, sondern präsentierend fortgesetzt wird. Mit dem Aufbau der Begründung hält der Sprecher nun die rechte Hand mit der Handfläche nach oben weiterhin präsentierend geöffnet⁷³ – während die linke Hand u. a. in Reparaturen kurzzeitig das Reparans präsentiert (vgl. Z. 251: kultURprogramm, vgl. #3) oder für die Begründung relevante Bedeutungselemente punktuell mit einer seitlichen Bewegung der Hand präsentierend hervorhebt (vgl. Z. 252: FÜHrung).⁷⁴ Mit dem Hervorheben des Kernelements des eingebrachten Beispiels (Z. 252: FÜHrung, vgl. #4) sowie des Kernelements des Resümeees (Z. 253: TOda, vgl. #5) wird zugleich auch die

⁷³ Die Standbilder zeigen, dass die rechte Hand eine *Palm Up* Geste formt, bei der die Finger leicht geflext sind. Streeck (2009b: 188, Hervorh. i. O.) beobachtet diese Form der Gestenausführung in epistemisch abgeschwächt modalisierten Äußerungen und beschreibt die Funktion der Geste daher als „*weak offerings*“.

⁷⁴ Dass auch die Geste der rechten Hand mit der Vorbereitung weiterer Gestenphrasen der linken Hand aligniert ist, zeigt sich in Standbild vier: Aus der Vorbereitung der lateralen Geste resultierend, in der der Sohn die Finger der rechten Hand enger zusammenführt, wird nun eine Orientierung der Handfläche der rechten Hand zum Körper des Sohnes ersichtlich.

rechte Hand in der lateralen Bewegung weiter aufgedreht und die präsentierende Funktion der beidhändigen *Palm Up* Geste offensichtlich. Der Sohn wendet zum Ende des Äußerungspakets den Blick seinem Vater zu (Z. 253) – der kurz darauf seinerseits den Blick vom Sohn abwendet – und hält die beidhändige *Palm Up* Geste auch in der anschließenden Pause noch 1.7 Sekunden aufrecht. Eine Rückmeldung des Vaters wird durch die Blickzuwendung und das Halten der Geste deutlich elizitiert, denn die Geste „displays the unfulfilled conditional relevance of the turn and solicits a contiguous response“ (Schönfelder und Heller 2019: 84; vgl. auch Bavelas et al. 1995; Streeck 2007). Hier übernimmt die Geste also auch eine interaktionsorganisatorische Funktion. Die Gesteneinheit wird dann durch das Überführen der Hände in ihre Ruheposition abgeschlossen, bevor der Vater das Äußerungspaket des Sohnes mit leicht steigender Intonationsbewegung ratifiziert (Z. 255: oKAY,; vgl. Betz und Deppermann 2021).

Insgesamt weist das Präsentieren von (komplexeren) Begründungen eine durchaus flexible multimodale Gestalt auf, denn Interagierende können sowohl ausgebaute satzübergreifende Äußerungspakete (vgl. Beispiel 1, 2) als auch einfache, satzförmige Propositionen als Begründungen innerhalb einer Gesteneinheit mithilfe eines *Post-stroke Holds* (vgl. Kendon 2004) präsentieren. Im Folgenden wird nun anhand einer lokalen, satzförmig strukturierten Begründung ein Beispiel des anderen Endes des Global-Lokal-Spektrums illustriert. Das nachfolgende Transkript zeigt, wie die Tochter eine evaluierende Proposition (Z. 078: ich mein musicalbesuch wäre interesSANT.) als Begründung auf der geöffneten Handfläche präsentiert.

Beispiel 3 – Präsentieren von (komplexeren) Begründungen

077 TO: aber ich überLEge zwischen: (.) eins und zwei.



078 ich mein musi|cal|besuch wäre |intere|sSANT.|
to-gs |...|dreht rH Palm Up auf--|,,,,,,|
to-he |neigt Kopf nach re-|,,,,,,,,,,,,|

079 MU: |JA.= |
mu-he |nickt|
mu-gz |blickt zu TO (>Z. 085)

080 TO: =(na)TÜrlich ne,
081 MU: |geNAU.|
mu-he |nickt-|

082 TO: aber ROM (.) <<p>mit kulturprogramm find ich eigentlich au_noch->
083 MU: ja find ich [AUCH.]
084 TO: [SEHR] interessant.
085 MU: SEHR interessant.|
mu-gz ----->|

Nachdem die Tochter zwei favorisierte Optionen benennt (Z. 077), bringt sie im Folgenden eine Begründung in Form einer nicht weiter elaborierten Proposition (Z. 078: ich mein musicalbesuch wäre interesSANT) für die zweite Option („Hamburg“) ein. Diese leitet sie mit der turninitialen und prosodisch in die Turnkonstruktionseinheit integrierten Phrase *ich mein* ein (vgl. Günthner und Imo 2003; Imo 2006), die als Diskursmarker „einen Wechsel der Aktivität“ (Günthner und Imo 2003: 12) von der Benennung der Favoriten hin zu einer Begründung markiert. Dabei dreht die Tochter mit dem Lexem „musicalbesuch“ (Z. 078) die rechte Hand zu einer *Palm Up* Geste mit gekrümmten Fingern (vgl. hierzu

Streeck 2009b: 188) auf und präsentiert innerhalb einer Gestenphrase die Proposition, bevor die Geste am Ende des *Turns*, das durch eine stark fallende Intonation markiert wird, in die Ruheposition zurückgeführt wird. Das Neigen des Kopfes bildet hier ein langsames Nicken in Richtung der Mutter, mit dem die Sprecherin ihre Begründung insgesamt verstärkt. Diese wird im direkten Anschluss von der Mutter ratifiziert (Z. 079), bevor die Interagierenden sich dann anderen Optionen zuwenden (vgl. Z. 082ff.). Insgesamt wird in diesem Beispiel also die lokal organisierte Begründung durch das Aufdrehen der Hand zu einer *Palm Up* Geste präsentiert und zur Disposition gestellt. Die Gesteneinheit enthält in ihrer Konfiguration ferner Informationen über die Verstärkung der evaluierenden Proposition.

Zusammenfassend konnte anhand der Daten gezeigt werden, dass eine Begründung bzw. Segmente einer komplexeren Begründung, die global in einem *Multi-Unit Turn* oder lokal in einer Turnkonstruktionseinheit organisiert und sprachlich als eigenständige Begründungssegmente eingeleitet werden (vgl. *aber* (vgl. Beispiel 1, 2), *ich mein* (vgl. Beispiel 3)), durch das Einsetzen der *Palm Up* Geste als abstrakte diskursive Objekte präsentiert werden können. Die Analysen zeigen ferner, dass die *Palm Up* Geste über eine gesamte Gesteneinheit im *Post-stroke Hold* gehalten wird und dass der Geste dadurch zusätzlich zu der performativen Funktion des Präsentierens auch äußerungsstrukturierende Funktionen zukommen. Denn die Geste markiert spezifische Segmente einer Äußerung als zusammengehörig. In Koordination mit weiteren Ressourcen können ferner einzelne Kernelemente der Begründung hervorgehoben werden (vgl. Beispiel 2) sowie enthaltene Evaluationen verstärkt werden (vgl. Beispiel 3).

8.2.1.2 Kurzzeitiges Präsentieren und Hervorheben relevanter Bedeutungselemente der Begründung

Das vorangegangene Kapitel hat dargelegt, inwiefern Begründungen und Begründungssegmente als abstrakte, diskursive Objekte mittels der *Palm Up* Gesten in einer Gesteneinheit als präsentiert kontextualisiert werden können. Die Untersuchung zeigt allerdings auch, dass Interagierende während des Begründens einzelne, häufig fokusakzentuierte Lexeme *punktuell* durch *Palm Up* Gesten präsentierend hervorheben können (vgl. Tab. 11). Diese sind zeitlich so koordiniert, dass das i. d. R. fokusakzentuierte Lexem mit dem Gestenhöhepunkt der *Palm Up* Geste zusammenfällt, sodass das Lexem durch prosodische sowie „visual prominence markers“ (Ambrazaitis und House 2017: 101) als relevantes Bedeutungselement aus dem Äußerungskontext hervorgehoben wird. In den vorliegenden mündlichen Argumentationen treten diese multimodalen Gestalten v. a. innerhalb der argumentativen Handlung des Begründens auf.

Kurzzeitiges Präsentieren und Hervorheben relevanter Bedeutungselemente der Begründung	
Beschreibung der multimodalen Gestalt	• einzelnes Lexem (i. d. R. + Fokusakzent) + <i>Palm Up</i> Geste (optional: <i>Beat</i>)

Tabelle 11: Beschreibung der multimodalen Gestalt „Kurzzeitiges Präsentieren und Hervorheben relevanter Bedeutungselemente der Begründung“

Wie nun Interagierende das Hervorheben relevanter Begründungselemente der Begründung identifizierbar machen, wird im Folgenden illustriert (vgl. Beispiel 1 & 2). Weitere Ressourcen wie beispielsweise ein *Beat*, der mit der *Palm Up* Geste erfolgt, oder besonders starke Fokusakzentuierungen, können der Präsentation der relevanten Bedeutungselemente einen besonderen Nachdruck verleihen (vgl. Beispiel 2).

In der abgebildeten Sequenz bezieht der Sohn initiiierend Position (vgl. Kap. 8.1.1), die er anschließend erst mit einer nicht weiter elaborierten Proposition und einem Vergleich zu begründen sucht. Dabei hebt er zwei Lexeme (Z. 049, 050: SCHULE, SHOPping) sowohl prosodisch als auch gestisch als relevante Elemente der Begründung hervor.

Beispiel 1 – Kurzzeitiges Präsentieren und Hervorheben relevanter Bedeutungselemente der Begründung

- 046 SO: also ICH würd ja sagen-
so-gz >>blickt auf Blatt (> Z.049)
- 047 ähm-
048 wir nehmen |den: (.)|BUNgalowpark,=
so-gs |.....|tippt geräuschvoll mit IF auf Option-->
- 
- 049 =denn |es |is ja (.)|von der |SCHU
so-gs -----|.....|öffnet rH Palm Up|laterale Bewegung, krümmt MF,
so-gz ----->|blickt zu VA----->|blickt auf Blatt----->
|le her,|
RiFi, klFi|,,,,,|
so-gz ----->
- 050 und ähm (1.0) |FREI |zeitmöglichkeiten mit
so-gs |tippt mit rIF auf Blatt|
so-gz ----->
- 
- |SHOP |ping| und |s schwimmen;=|
so-gs |öffnet rH Palm Up|,,,,|
so-gz ----->|blickt zu VA--|
- 051 =is Eher was-
052 als ähm (-) auf_nem SEgelschiff (.) da (-) fünf tage zu sein;
053 | (3.0)
so-gz |blickt zu VA-->>
- 054 VA: <<p>ich find das ALles irgendwie spannend.>

Die erste Begründung, bestehend aus einer nicht weiter elaborierten Proposition (Z. 049: denn es is ja von SCHUle her,), führt der Sprecher mit der kausalen Konjunktion *denn* ein und hebt das hier als Kernelement der Begründung fungierende Lexem „SCHUle“ (Z. 049) hervor: Neben dem Fokusakzent markiert auch der Gestenhöhepunkt der lateralen Bewegung der *Palm Up* Geste das Lexem „SCHUle“ kurzzeitig (vgl. #1, #2).⁷⁵ Das Lexem wird dadurch inhaltlich das Kernelement der Begründung. Die Begründung selbst verbleibt allerdings inhaltlich implizit, denn warum gerade ein Bezug zur Schule die Wahl des Bungalowparks plausibilisieren soll, wird nicht weiter elaboriert. Allerdings tragen sowohl die Modalisierung durch die Modalpartikel *ja* (vgl. Reineke 2016, 2018) als auch die laterale Bewegung

⁷⁵ Die Finger der geöffneten Hand beginnen, sich bereits mit Einsetzen des Fokusakzentes zu krümmen; jedoch ist der Gestenhöhepunkt der lateralen Bewegung der Geste, der sich in der Rekonstruktion der multimodalen Gestalt des *Kurzzeitigen Präsentierens und Hervorhebens relevanter Bedeutungselemente der Begründung* als bedeutsam erweist, noch nicht erreicht.

der *Palm Up* Geste (vgl. Kendon 2004) dazu bei, die Begründung als offensichtlich bzw. evident und somit als ausreichend plausibilisiert zu markieren – das kurzzeitig hervorgehobene Lexem „SCHULE“ (Z. 049) am Ende der Intonationsphrase erhält somit einen Argumentstatus, das keiner weiteren Plausibilisierung bedarf. Das Äußerungspaket wird fortgesetzt (Z. 050-052), indem ein neues Inhaltselement (Z. 050: FREIzeitmöglichkeiten) deiktisch eingeführt und zwei Beispiele für die Ausgestaltung der Freizeitmöglichkeiten benannt werden: „SHOPping“ und „schwImmen“ (Z. 050). Das Lexem „SHOPping“ (Z. 050) wird hier sowohl prosodisch als auch gestisch präsentierend hervorgehoben: Die rechte Hand öffnet sich koordiniert mit dem Fokusakzent des Lexems kurzzeitig zu einer *Palm Up* Geste, die aufgrund der interpersonellen Ausrichtung den Vater als Adressaten erkennbar macht (vgl. #3; vgl. *Palm Addressed*, Kendon 2004: 271ff.). Unmittelbar nach dem Fokusakzent wird die Geste in eine Ruheposition überführt.

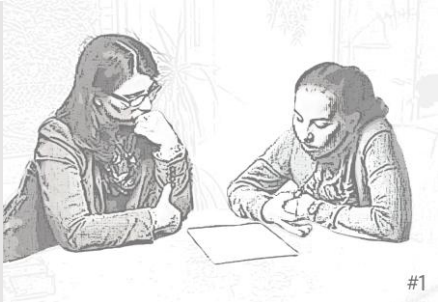
Das zweite Beispiel illustriert, dass das kurzzeitige Präsentieren von Lexemen auch mit Nachdruck erfolgen kann. Im vorliegenden Beispiel wird dazu die *Palm Up* Geste in der Bewegung einer *Beat* Geste nach unten geführt, um das Kernelement eines Beispiels turnfinal zu präsentieren. Damit schließt die Sprecherin eine Liste bestehend aus drei Argumenten ab.

Beispiel 2 – Kurzzeitiges Präsentieren und Hervorheben relevanter Bedeutungselemente der Begründung

```

295  MU:  °h ja;
296      jetzt müssen wa nur dann (.) uns eine entSCHEIdung- (--)
297      oder für etwas entSCHEIden;
298      und dann sagen waRUM;
299      (1.7)
300  TO:  m:-
301      (3.7)
302      äh: vier tage HAMBurg mit musicalbesuch würd ich nehmen,
303      weil:-
304      (1.1)
305      m erstens sind_s ähm es sind vier TAge;=
306      =das (sind/würde) vollkommen (.) is vollkommen AUSreichend;
307      äh: (.)ZWEItens;
308      (0.7)
309      ähm (2.2) man:: (.) kann da (.) geNAUso;
310      (---)
311      <<p, all>wie du geSACHT has;>
312      so: (-) SHOPpen und SCHWIMmen;
313      und sowas machen,
314      (1.1)
315      und DRITten::s,
316      (1.1)
317      äh (1.9) ähm man: (1.2) KANN auch-
318      halt AUSser |diesen musical          |besuch,          |
    to-gs          |zeigt mit rIF auf Blatt|
    mu-gz          |blickt auf Blatt-|blickt zu TO->>
319      (0.5) |(0.5)
    to-gs          |rH Palm Up Ri Körper->
320      mit der klASSE |was zuSAMmen          |machen;
    to-gs          ----->|führt rH Palm Up nach vorn-|hält-->

```

#1

321 zum bei|spiel| ins: (-) |!DUN! |geon| gehen;
to-gs ----->|,,,,,| |öffnet rH Palm Up, Beat|,,,,,|
322 (1.1)
323 MU: ZUM beispiel;
324 (1.0)
325 TO: [ja;]
326 MU: [°h ja:-]

In der abgebildeten Sequenz hebt die Sprecherin innerhalb einer Begründung, die sie listenförmig (vgl. Selting 2004) aufbaut, das Lexem „!DUN!geon“ (Z. 321) besonders hervor: Sie führt es innerhalb einer Beispielnennung (vgl. Brünner und Gülich 2002b; Kinalzik 2024) an, mit der sie ihre Proposition (Z. 317, 320: [...] man: KANN auch- / [...] / mit der klASSE was zuSAMmen machen;“) veranschaulichend stützt. Die Relevanz der Beispielnennung hervorhebend präsentiert sie das Lexem als relevantes Bedeutungselement mit einer *Palm Up* Geste, die eine Orientierung zur Sprecherin enthält, führt die Hand in einem *Beat* mit sichtbarer Bewegung nach unten und hebt das Lexem zugleich prosodisch durch einen besonders starken Fokusakzent kurzzeitig hervor. Die dabei entstehende Nachdrücklichkeit wird verstärkt durch die dem Lexem voranstehende und es dadurch vom Rest der Äußerung abtrennende Pause sowie das anschließende Beenden des gesamten Äußerungspakets, das mit einer fallenden Intonation abschließt. Nach einer Pause erfolgt dann ein schwaches Zustimmung in Form einer Bestätigung seitens der Mutter in Form einer lexikalischen Wiederholung (Z. 321: zum beispiel / Z. 323: ZUM beispiel), an die jedoch keine weitere Affirmation anschließt (vgl. Schegloff 1996).

Insgesamt kann festgehalten werden, dass Interagierende durch den Einsatz der *Palm Up* Geste sowie i. d. R. auch prosodische Markierungen einzelne Bedeutungselemente sprichwörtlich hervorheben: Die Elemente werden dabei sichtbar kurzzeitig auf der geöffneten Handfläche einer *Palm Up* Geste präsentiert und aus dem Äußerungskontext i. S. einer äußerungsstrukturierenden Funktion hervorgehoben. In ihrer Gesamtgestalt bilden sie das Kurzzeitige Präsentieren und Hervorheben relevanter Bedeutungselemente der Begründung.

8.2.2 Präsenthalten eines verhandelten Gesprächsgegenstandes innerhalb von Begründungen

Interagierende können in Gesprächen nicht nur Inhalte durch Gesten präsentierend hervorheben (vgl. Kap. 8.2.1), sondern durch das Halten von Gesten auch aktuell verhandelte Gesprächsgegenstände für alle Beteiligten (über einen längeren Zeitraum) sichtbar präsent halten. Dabei wird auf die hohe Funktionalität gehaltener Gesten bereits bei Mondada (2007: 220) verwiesen. Sie rekonstruiert, wie Interagierende, während sie unterbrochen werden, begonnene Gesten nicht in eine Ruheposition zurückführen, sondern aufrechterhalten und dadurch anzeigen, weiterhin Anspruch auf das Rederecht zu erheben. Das Halten der Geste während der Unterbrechung hat also Konsequenzen für die Organisation der Interaktion sowie den dadurch hochgradig erwartbar werdenden Anschluss bzw. die Fortsetzung der gerade unterbrochenen Interaktion: Eine Aktivität wird weiterhin durch körperliche Ressourcen präsent gehalten, während eine weitere bearbeitet wird (vgl. hierzu auch Schegloff 1998). Ebenfalls können Inter-

gierende durch das turnfinale Halten ihrer Geste eine Reaktion des Gegenübers in Bezug auf ein präsentiertes diskursives Objekt elizitieren (vgl. Streeck 2007; vgl. 8.2.1.1, Beispiel 2), wodurch die konditionelle Relevanz (vgl. Schegloff 1968) einer erwarteten Rückmeldung verstärkt wird (vgl. Schönfelder und Heller 2019). Ferner übernimmt nicht nur das Halten, sondern auch der Rückzug der Geste wichtige interaktionsorganisatorische Funktionen: Indem die Geste z. B. in eine Ruheposition überführt wird, wird auch die etablierte konditionelle Relevanz zurückgenommen. Dies kann beispielsweise eintreten, wenn das Gegenüber auf den präsent gehaltenen Gesprächsgegenstand Bezug nimmt.

In den vorliegenden Daten übernimmt das Halten von Gesten insgesamt drei voneinander zu unterscheidende Funktionen: Durch das Halten einer Geste kann ein etablierter Gesprächsgegenstand, z. B. ein spezifisches Ziel der Abschlussfahrt oder aber für die Entscheidung relevante Kriterien, für alle Beteiligten in der Fortsetzung des *Turns* sowie über Sprecher*innenwechsel hinaus visuell präsent gehalten werden (vgl. Kap. 8.2.2.1). Dies ermöglicht den Interagierenden, sich auf den aktuell verhandelten Gesprächsgegenstand hin (zurück) orientieren zu können und zugleich, durch das Halten der Geste, den etablierten Gesprächsgegenstand auch für die Rückmeldung des Gegenübers relevant zu machen. Das Halten von Gesten kann ebenfalls dem Projizieren einer Turnfortsetzung in überlappenden Sequenzen dienen, wenn eine*r der Sprecher*innen das Rederecht weiterhin für sich reklamiert, und somit auch zur Durchsetzung des Rederechts beitragen (vgl. Kap. 8.2.2.2). Das Halten von deiktisch-referenzierenden Gesten wird darüber hinaus auch von Rezipient*innen eingesetzt, um durch das Zeigen auf eine Option, während der*die Andere gerade eine Option verhandelt, das *eigene* Verstehen anzuzeigen und dabei zugleich den Gesprächsgegenstand (rezipient*innenseitig) präsent zu halten. Sie zeigen eine aktive Beteiligung bzw. Partizipation (vgl. Goodwin 2007) an der Begründungshandlung der aktuellen Sprecher*innen an (vgl. Kap. 8.2.2.3).

8.2.2.1 Etablieren und Präsenthalten eines verhandelten Gesprächsgegenstandes

In mündlichen Argumentationen können Interagierende aktuelle diskursive Objekte bzw. Gesprächsgegenstände nicht nur sprachlich benennen, sondern auch durch deiktisch-referenzierende sowie anhand von Zählgesten für alle Interagierenden visuell sichtbar machen und über einen längeren Zeitraum durch das Halten der Geste präsent halten. Dabei kann ein Gesprächsgegenstand während des Begründens innerhalb des eigenen *Turns* eines*einer Interagierenden (vgl. Beispiel 1) sowie über turnübergaberelevante Stellen hinweg (vgl. Beispiel 2) präsent gehalten werden. Wie die nachfolgende Übersicht 12 darlegt, können nicht nur deiktisch-referenzierende Gesten einen Gesprächsgegenstand präsent halten, sondern auch Zählgesten können in dieser Funktion verwendet werden und z. B. eingeführte Entscheidungskriterien äusserungsübergreifend bzw. über Sprecher*innenwechsel hinweg sichtbar anzeigen (vgl. Beispiel 3).

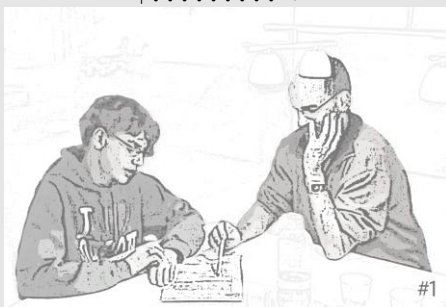
Etablieren und Präsenthalten eines verhandelten Gesprächsgegenstandes	
Beschreibung der multimodalen Gestalt	• Benennen des Gesprächsgegenstandes (z. B. Option X) + deiktisch-referenzierende Geste (auf Material) + Halten der Geste äusserungsintern und/oder über Sprecher*innenwechsel hinweg • Benennen des Gesprächsgegenstandes (z. B. entscheidungsrelevante Kriterien) + Zählgeste + Halten der Geste äusserungsintern und/oder über Sprecher*innenwechsel hinweg

Tabelle 12: Beschreibung der multimodalen Gestalt „Etablieren und Präsenthalten eines verhandelten Gesprächsgegenstandes“

Der erste Transkriptauszug zeigt, wie Interagierende mit einer deiktisch-referenzierenden Geste auf das Material Bezug auf einen Gesprächsgegenstand nehmen und diese Referenz innerhalb eines größeren Äußerungspakets aufrechterhalten. In der nachfolgenden Sequenz führt der Vater das Bezugsobjekt ‚Hamburg‘ zu Beginn einer Begründung ein und hält es während der Begründung durch eine deiktisch-referenzierende Geste mit einem Stift sichtbar präsent.

Beispiel 1 – Etablieren und Präsenthalten eines verhandelten Gesprächsgegenstandes

125 VA: also ich |SACH mal.
va-gz >>blickt auf Blatt (>Z. 129)
va-rh |.....->



126 VIER |tage HAM|burg. |(|.)|
va-rh|senkt Stift auf Option ab|...|
so-gz |blickt auf Blatt (>Z. 128)
127 |EIn_n |tach irgendwie auf nen abend MUsical;
va-rh |tippt-|hält (>Z. 136)
128 was macht man den |REST der zeit |in hamburg.
va-gz |blickt in mittlere Distanz (>Z. 135)
so-gz ----->|blickt in mittlere Distanz (>Z. 131)
129 (-)
130 |°hh |
va-lh |.....|
va-he |hebt Kopf an|



131 SO: |FISCHmar|kt.
so-gz ----->|blickt zu VA----->>
va-lh |lArm aufgestellt, öffnet lH Palm Up (>Z. 133)
132 VA: ja:|=zum BEIspiel;
133 (geht man) auf den |FISCH|ma:rkt-
va-lh ----->|Beat-|hält Palm Up->
134 geht man ir|gendwie: (.) zum |SHOP |pen,|
va-gz ----->|blickt auf Blatt (>Z. 138)
va-lh ----->|Beat-|hält|
135 |GEHT ma|n-
va-lh |,,,,,,|legt lIF auf Option->
136 <<kopfschüttelnd>was |weiß ICH, |
va-lh ----->|
va-rh ----->|fährt mit Stift in rH über Option|
137 |einmal> |Hafenrundfahrt muss sei:n-
va-gs |hebt rH mit Stift & lH Palm Down an|
138 dann geht man irgendwie: in die SPEI|cherstadt-
va-gz ----->|blickt zu SO (>Z. 142)
139 hamburg DUN:geon-
140 geht man in: MIniwelt-

141		oder SONST irgendwas.
142		(1.0)
	va-gz	blickt auf Blatt
143		(GUT/GOT) . =

In dem global strukturierten Äußerungspaket (Z. 125-142) nimmt der Vater Bezug auf die Option ‚Hamburg‘ (Z. 126) und hält diese über mehrere Turnkonstruktionseinheiten hinweg für beide Interagierenden visuell präsent (Z. 126-136): In Vorbereitung auf die anschließende selbstbeantwortete (Informations-) Frage (Z. 128: was macht man den REST der zeit in hamburg.; vgl. Bücken 2015), die eine Beantwortung durch das Gegenüber (eigentlich) nicht erwartbar macht, referenziert der Sprecher mittels einer deiktischen Geste auf die Option ‚Hamburg‘ (Z. 126; vgl. #1). Der Blick des Sohnes folgt der Geste bereits in ihrer Vorbereitung. Mit der Wiederholung der vorgegebenen Aktivität des Musicalbesuchs (Z. 127) wird die deiktisch-referenzierende Geste, die hier durch das Zeiginstrument eines Stiftes geäußert wird (vgl. Stukenbrock 2015: 37), auf der Option gehalten.

Im Anschluss an die Frage (Z. 128) setzt nun eine Pause ein: In dieser Pause hebt der Vater den Kopf an und lenkt den Blick in die mittlere Distanz, während zugleich eine *Palm Up* Geste (vgl. Müller 2004; Kendon 2004) vorbereitet wird (vgl. #2). Zeitgleich mit der Entfaltung der multimodalen Ressourcen des Vaters, d. h. dem Blick sowie der Vorbereitung der Geste, die gemeinsam ein *thinking display* (vgl. Heller 2021b) aufbauen, benennt nun der Sohn eine erste Beispielaktivität (Z. 131: FISCHmarkt.; vgl. Brünner und Gülich 2002b). Der Vater hält die Freizeitaktivität des Musicalbesuchs weiterhin mit dem Stift präsent, während er das Beispiel des Sohnes durch eine minimale Form der Affirmation (Z. 132: ja,=zum BEispiel;) sowie eine elaborierte Wiederholung (Z. 133: (geht man) auf den FISCHmarkt-) ratifiziert und diesem Beispiel ein weiteres hinzufügt (Z. 134: geht man irgendwie: (.) zum SHOPpen.). Beide Listenelemente sammelt er, den linken Unterarm aufgestellt, auf der geöffneten Handfläche der linken Hand. Dabei hält der Vater während der Beispielnennung des Sohnes sowie der ersten zwei von insgesamt sechs Listenelementen (vgl. Z. 133-140; vgl. Selting 2004) durch die deiktisch-referenzierende Geste die Freizeitaktivität des Musicalbesuchs als Bezugspunkt der aufgezählten Freizeitaktivitäten über mehrere Turnkonstruktionseinheiten für beide Interagierenden visuell sichtbar präsent. Das Halten der Geste wird erst mit einer Anzeige von epistemischer Unsicherheit (Z. 136) aufgelöst. Zusammenfassend übernimmt die deiktische Geste *in* diesem und *für* diesen sequenziellen Kontext also nicht nur eine referenzierende Funktion (vgl. Kap. 3.2.2), sondern das Halten der deiktischen Geste kann in längeren Äußerungspaketen des Begründens helfen, die verhandelte Option in für die Interagierenden visuell präsent zu halten, ohne wiederholt sprachliche Referenz herstellen zu müssen. Dadurch übernimmt das Halten der deiktischen Geste zum einen auch eine äußerungsstrukturierende Funktion, denn der Bezugspunkt der Begründung wird konstant für alle Beteiligten sichtbar präsent gehalten (vgl. auch Mondada 2007). Zum anderen ermöglicht das multimodale Fokussieren auf einen ausgewählten Referenten, diesen über einen längeren Zeitraum zu verhandeln und so eine vertiefende Begründung aufzubauen.

Indem eine deiktisch-referenzierende Geste über die Äußerung des*der aktuellen Sprecher*in hinaus gehalten wird, kann das Referenzobjekt auch für die Äußerung des Gegenübers relevant gemacht werden. In dem nachfolgenden Beispiel referenziert der Sohn nacheinander auf zwei Optionen, die er als Alternativoptionen vorstellt und mit dem Halten der zweiten deiktischen Geste die Option auch über den Sprecher*innenwechsel hinaus präsent hält.

Beispiel 2 – Etablieren und Präsenhalten eines verhandelten Gesprächsgegenstandes

214 SO: ich MEine- (-)
 215 ich BIN ja schon als schülervertreter da;
 216 dann sollte ich ja auch zumindest so ein PAAR meinungen vertreten.
 217 VA: hm_hm,
 218 (2.0)
 219 oKAY,
 220 was dEnkst du denn was DEI[ne-]
 221 SO: | [u:n:d] (.) |dann: (.) |
 so-rh ||rH Palm Up, Fi. gekr., Ri Blatt|
 so-gz >>auf Blatt (>Z. 224)
 va-gz >>auf Blatt->>
 |würd ich sa|gen-|
 so-rh |legt rH auf Blatt->>
 so-lh |....|



222 |DAS: |wär |dann eher (.) so (.) drinne;
 so-lh |legt lIF auf Blatt-|tippt 2x-|hält----->
 223 und wenn ich |GAR |nich schü |ler|ver |treter wär,
 so-lh ----->|.....|legt lMF neben lIF-|hält----->
 so-he |schüttelt Kopf|



224 dann |wär |DAS: (.) wahrscheinlich drin|ne.
 so-lh ---->|tippt mit lMF-|hält (>Z. 226)
 so-gz |zu VA (>Z. 226)
 225 weil dann hätt ich nichts zu SAgen un:d- (--)
 226 |ja. | (0.8) | (.)
 so-gz ----->|blickt auf Blatt->
 so-gs |tippt m. lIF 3x-----|hält lH auf Blatt->
 so-he |nickt seitlich|



#3

227 VA: (aber/nur) ich denk du wirst |nich |der EINzige sein |
 so-lh -----|nimmt Fi vom Blatt|
 so-gz ----->
 in der klasse |dem da: (.) |schlEcht wird.
 so-gz ----->|blickt Ri lH d. VA-|blickt zu VA->>

In der vorliegenden Sequenz stellt der Sprecher innerhalb eines veranschaulichenden Beispiels (vgl. Kinalzik 2024) zwei Optionen (Bungalowpark, Segeltörn) als Alternativen einander gegenüber und weist hier auf einen Interessenskonflikt hin: Denn er selbst bevorzugt eine andere Option als er für seine Mitschüler*innen annimmt, deren Meinung er aber als Schüler*innenvertreter vertreten sollte (vgl. Z. 214-216). Mit Benennung beider Optionen, in Form fokusakzentuierter, gedehnter Demonstrativpro-nomen (Z. 222, 224) referenziert er auch gestisch auf diese (vgl. #1, #2). Nachdem der Sprecher nun auf die Option des zweiten Alternativbeispiels (Z. 223f.: und wenn ich GAR nicht schülervvertreter wär, / dann wär DAS: (.) wahrscheinlich drinne.) mit dem linken Mittelfinger verweist, folgt eine Expansion des *Turns* in Form einer Begründung (Z. 225f.: weil dann hätt ich nichts zu Sagen un:d- (--) / ja.), die anschließend abgebrochen wird. Die referenzierende Geste mit dem linken Mittelfinger hält den ‚Segeltörn‘ über den Abbruch und den anschließenden Sprecher*innenwechsel hinaus präsent (vgl. Z. 227, vgl. #3). Das Präsenthalten der Option hat nun Konsequenzen für die Fortsetzung der Interaktion: Der Vater nimmt in dem sich nun anschließenden Einwand nicht auf die zuerst benannte Option des Bungalowparks Bezug, sondern auf den zuletzt benannten Segeltörn (Z. 227), den der Sohn mittels der deiktischen Geste präsent hält. Mit Beginn des Einwandes wird die gehaltene Geste nun aufgelöst: Die Rückführung der Geste beginnt noch bevor innerhalb der Äußerung des Vaters explizit ein Verweis auf den Segeltörn benannt wird. Da der Vater aber bereits zuvor den Segeltörn ausgeschlossen und dabei zum Kern der Begründung die Seekrankheit des Sohnes gemacht hat, projiziert im vorliegenden Äußerungsbeginn der Bezug auf den Sohn (Z. 227: (aber/nur) ich denk du wirst), dass sein Einwand sich auf die vom Sohn weiterhin präsent gehaltene Option des Segeltörns beziehen wird. Mit Einlösen der konditionellen Relevanz (vgl. Schegloff 1984), die durch die Geste aufgebaut worden ist, wird die deiktische Geste nun in eine Ruheposition überführt (vgl. Z. 227).

In der Studie zeigt sich ferner, dass nicht nur deiktisch-referenzierende Gesten, sondern auch Zählgesten gehalten werden, mittels derer die Interagierenden beispielsweise listenförmig eingebrachte Argumente oder auch Kriterien zur Entscheidungsfindung für alle Beteiligten präsent halten. Im nachfolgenden Beispiel führt die Mutter zu Beginn der Interaktion drei Kriterien zur Entscheidungsfindung ein, visualisiert die Kriterien durch eine Zählgeste und setzt sie damit turnübergreifend für die Fortsetzung der Interaktion relevant. Der Geste kommen in der nachfolgenden Sequenz somit sowohl referenzielle und abbildende als auch pragmatische Funktionen zu.

Beispiel 3 – Etablieren und Präsenhalten eines verhandelten Gesprächsgegenstandes

120 MU: ((schnalzt)) auf der EInen seite is die [FRAge;]
 (>Z. 122)
 mu-gz >>blickt auf Blatt (>Z. 138)
 to-gz >>blickt auf Blatt (>Z. 134)
 121 TO: [(viele)][(leute)]
 122 MU: [welches] ZIEL wählt
 man aus,
 123 (1.6)
 124 dAnn hast_e das problem mit den TAgen,
 125 und DANN- (-)
 126 wie LAnge man weg ist,
 127 und dann (-) hast_e noch unterschiedliche FREIzeitmöglichkeiten.
 128 nach DEnen man auswählen kann.



129 |hast_e |also drei verschiedene kriTErien;
 mu-gs |dreht LH zu Zählgeste „drei“ auf-|hält----->
 130 TO: ja;|
 mu-gs -->|



131 TO: |°h |also:-
 mu-gs |senkt Zählgeste leicht ab-|hält (>Z. 135)
 132 TO: (--)
 133 TO: den bungalowpark würde ich (hier/eh) NICH nehmen;
 134 (.) glaub das is so was ähnliches wie movie|park;
 to-gz ----->|blickt zu MU (>Z. 138)
 135 TO: also-|
 mu-gs ---->|
 136 TO: (was weiß ich/weiß nicht-)
 137 (-)
 138 |UNter dem gesichtspunkt |dass: (0.7) man nicht |weit WEG|fährt,|
 to-gz |blickt auf Blatt-----|blickt zu MU-----|blickt
 n. u.->>
 mu-gz ----->|blickt zu TO---|
 139 dass ALle auch zu hause machen können,
 140 (--)
 141 un:d-
 142 (1.4)
 143 MU: °h naja;=
 144 =aber es geht ja AUCh darum-
 145 dass die KLASsengemeinschaft nicht mehr da is;
 146 es wäre ja SCHÖN,
 147 wenn man zum schluss noch mal was gEMEINsames [machen könnte;]

Zu Beginn der Interaktion benennt die Mutter insgesamt drei Kriterien, die durch den Fokusakzent in der Auflistung hervorgehoben (Z. 122, 126, 127: ZIEL, wie LANge, FREIzeitmöglichkeiten) und mit steigender Intonation aneinandergereiht werden (vgl. Selting 2004). Diese Kriterien sollen zur Entscheidungsfindung beitragen, wie der anschließende Metakommentar expliziert (Z. 129: *hast_e* also drei verschiedene kriTERien;). Die Mutter dreht mit Beginn der resümierenden Turnkonstruktionseinheit die linke Hand in Form einer Zählgeste, bei der der Daumen sowie der Zeige- und Mittelfinger ausgestreckt sind, nach oben (vgl. #1) und visualisiert dadurch die von ihr etablierten Kriterien.⁷⁶ Die Tochter markiert durch die Affirmationspartikel *ja* (Z. 130) eine minimale Zustimmung gegenüber den Kriterien, woraufhin die aufgedrehte Zählgeste leicht abgesenkt (Z. 131, vgl. #2), aber weiter gehalten wird. Mit Sikveland und Ogden (2012: 178) kann das Absenken bzw. das Entspannen der Hand als eine Markierung des „shared understanding“ verstanden werden. Das Entspannen bei Gesten äußernden Interagierenden setzt ein, sobald das Gegenüber ein gemeinsames Verständnis markiert. In dem vorliegenden Datum setzt eine Entspannung der Geste nach kindseitigen Affirmation der Kriterien ein (Z. 130f.), wenngleich die abgesenkte Zählgeste die Kriterien in der anschließenden Begründung der Tochter weiterhin präsent hält und die Geste nicht, wie bei Sikveland und Ogden (2012), in eine Ruheposition überführt wird. In der Turnübernahme bearbeitet die Tochter nun das erste Kriterium der Mutter, indem sie das Ziel (vgl. oben, Kriterium eins) des Bungalowparks auf Grundlage eines Vergleichs ausschließt (Z. 131-134) und nach einigen Verzögerungsmarkierungen (Z. 136f.) weiter elaboriert (Z. 138f.). Die Zählgeste der Mutter, die die etablierten Kriterien der Entscheidungsfindung präsent hält, wird bis zum Ende des Vergleichs (Z. 134) gehalten und dann in eine Ruheposition überführt (Z. 135). Mit dem Rückzug der Geste werden nun die bislang präsent gehaltenen Kriterien aus dem Wahrnehmungsraum der Interagierenden entfernt – an dieser Stelle verbleibt vorerst unklar, ob die Geste zurückgezogen wird, weil die Kriterien von der Tochter *genügend* oder aber *unzureichend* bearbeitet werden. Der anschließende Einwand (Z. 143ff.) lässt vermuten, dass die Mutter die Bearbeitung der Kriterien seitens der Tochter als unzureichend bewertet.⁷⁷

Zusammenfassend zeigt sich, dass Interagierende sowohl durch deiktisch-referenzierende Gesten auf das Material als auch durch Zählgesten Gesprächsgegenstände etablieren und durch das Halten der Geste den Bezugspunkt ihrer Begründung in äußerungsstrukturierender Funktion sichtbar präsent halten können. Werden die Gesten außerdem über den Sprecher*innenwechsel hinweg gehalten, kommt den gestischen Ressourcen auch eine interaktive bzw. interpersonelle Funktion zu, da somit die entsprechenden Gesprächsgegenstände bzw. etablierten Kriterien einer gemeinsamen Entscheidungsfindung auch für die Fortsetzung der Interaktion und somit für das Gegenüber relevant gemacht werden können.

8.2.2.2 *Etablieren und Präsenthalten eines Gesprächsgegenstandes bei konkurrierender Turnübernahme*

Interagierende können mithilfe von deiktisch-referenzierenden Gesten sowie Zählgesten Gesprächsgegenstände einführend etablieren und sowohl äußerungsintern als auch äußerungsübergreifend für alle Beteiligten sichtbar präsent halten (vgl. Kap. 8.2.2.1). Beide Gestenformen können auch eingesetzt werden, um eine Turnübernahme des Gegenübers zu verhindern: In überlappenden, d. h. simultan produzierten Äußerungen (vgl. Sacks et al. 1974), in denen beide Sprecher*innen den *Turn* übernehmen und/oder jeweils andere Gesprächsgegenstände etablieren möchten, kann das Präsenthalten der Geste

⁷⁶ In den vorliegenden Daten wird ersichtlich, dass sich die Zählgeste auch sukzessive in der Interaktion entfalten kann: Die inhaltliche ‚Belegung‘ einzelner Finger kann also auch einzeln und nacheinander erfolgen.

⁷⁷ Ganz auflösen lässt sich die Frage jedoch nicht, denn die Mutter nimmt an dieser Stelle selbst keinerlei Bezug mehr auf die von ihr eingeführten Kriterien, sondern führt nun ein neues Kriterium ein, das der Klassengemeinschaft (Z. 144-147).

sowie das sprachliche Fortsetzen des *Turns* anzeigen, dass das Rederecht (vgl. Sacks et al. 1974) bei einem*einer der Sprecher*innen verbleiben und dessen etablierter Gesprächsgegenstand weiter verhandelt werden soll (vgl. Mondada 2007; Oloff 2013). Das Aufrechterhalten der Geste im Verbund mit der Turnfortsetzung erhält also eine weitere interaktionsorganisatorische i. S. einer sprecher*innenwechselorganisierenden Funktion (vgl. Tab. 13).⁷⁸

Etablieren und Präsenhalten eines verhandelten Gesprächsgegenstandes bei konkurrierender Turnübernahme	
Beschreibung der multimodalen Gestalt	Benennen des Gesprächsgegenstandes (z. B. Option X) + Halten der ausgeführten Geste in konkurrierender Turnübernahme

Tabelle 13: Beschreibung der multimodalen Gestalt „Etablieren und Präsenhalten eines verhandelten Gesprächsgegenstandes bei konkurrierender Turnübernahme“

Anhand der nachfolgenden Beispiele wird dargelegt, wie das Halten unterschiedlicher Gestenformen wie deiktische (vgl. Beispiel 1) oder Zählgesten (vgl. Beispiel 2) bei konkurrierender Turnübernahme anzeigen können, dass der*die Sprecher*in weiterhin das Rederecht beansprucht.

In dem folgenden Transkriptausschnitt verhandeln die Interagierenden parallel und dabei überlappend eine Option. Durch das Halten einer deiktischen Geste auf das Material zeigt der Vater sichtbar für beide Beteiligten an, den *Turn* fortzusetzen und somit das Rederecht zu beanspruchen.

Beispiel 1 – Etablieren und Präsenhalten eines verhandelten Gesprächsgegenstandes bei konkurrierender Turnübernahme

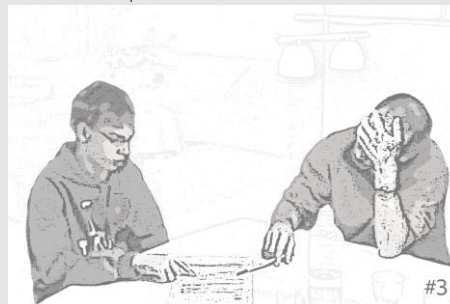


596 SO: au[ßerdem] is die fahrt nicht so |lang |wie nach |RO:M,
 so-gs |.....|zeigt mit rIF
 auf Option (>Z. 599)
 so-gz >>blickt auf Blatt (>Z. 605)
 va-gz ----->|auf Blatt
 597 VA: [oKAY;]
 598 VA: |man HAT (.) |<<nickend>entsprechend | [mehr |ZEIT,>]|
 va-fa |zieht AB hoch|
 va-gs |.....|zeigt mit Stift auf Option-|tippt-|hält---|
 so-gs ----->|hebt Hand & rIF an----->
 599 SO: [O:der]
 600 VA: |klar,=| (.)
 va-gs |tippt-|hält->
 so-gs ----->
 601 SO: |[NACH::-]
 so-gs |senkt rIF langsam auf Option (Segeltörn) ab-->

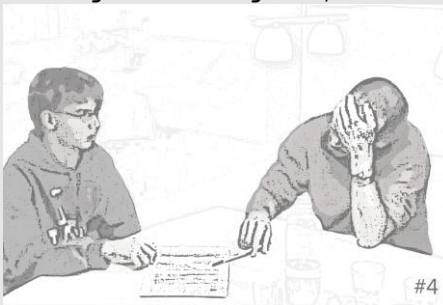
⁷⁸ Vgl. zur Funktion von Handgesten für die Sprecher*innenwechselorganisation im *Universitätskontext* Schmitt (2005), für *Arbeitsbesprechungen* Mondada (2007) und für *gemeinsames Erzählen* Zima (2017). Vgl. außerdem zum Zusammenhang von Gestenphasen, Beendigungen von Turnkonstruktionseinheiten und Sprecher*innenwechseln Kendrick et al. (2023).



602 VA: |[nja |das is] RICHTig-|
 va-gz ----->|blickt zu SO (>Z. 615)
 va-gs |zeigt auf andere Option (Rom) (>Z. 614)
 so-gs ----->|



603 VA: |drei |tage ROM,
 so-gs |tippt mit rIF auf Option-|hält (>Z. 605)
 604 VA: heißt ja einen tag HIN,



605 VA: einen |tag zuRÜCK,
 so-gs ,,,,,,|
 so-gz ----->|blickt zu VA (>Z. 617)
 606 SO: einen tag D[A;]
 607 VA: [einen] tag DA;;
 608 is eigentlich VÖLLig ineffek[tiv;]
 609 SO: [AU]Sser=
 610 SO: =es is |einen |tag HIN,|
 so-gs |tippt m. rIF auf Tisch-|tippt---|
 611 SO: drei tage DA:,
 612 SO: |einen| tag zuRÜCK.
 so-gs |tippt|
 613 (1.6)
 614 SO: aber das [GLAUB] ich |[nich;]
 va-gs ----->|
 615 VA: |[°h]
 va-gz ----->|auf Blatt (>Z. 617)
 616 VA: [ja] aber dann is |man auch von der FAHRT-
 617 is man irgendwie kaPU|TT,
 va-gz ----->|zu SO----->>
 so-gz ----->|

Nach einer anfänglichen ko-konstruierten Begründung (vgl. Kreuz 2021) für die Option ‚Hamburg‘ (Z. 596f.), in der die Interagierenden die Länge der Fahrtzeit zu einer anderen Destination, nämlich ‚Rom‘, vergleichen, zeigt sich eine Wendung von einer ko-konstruierten zu einer *parallelen, konkurrierenden* Aushandlung: Während der Vater nun die Option ‚Rom‘ weiter bearbeitet (Z. 602ff.), setzt der Sohn seinen Vergleich fort (Z. 599, 601). Beide Sprecher überlappen in ihren *Turns* (Z. 601, 602) und verweisen, parallel zueinander, mit deiktischen Gesten auf die von ihnen bearbeiteten Optionen ‚Rom‘ und ‚Segeltörn‘ (vgl. #1). Während der Sohn weiterhin auf das Material schaut, adressiert der Vater ihn nun mit seinem Blick (Z. 602) und verstärkt durch die Blickzuwendung sowie den überlappenden Redezug, das Rederecht behalten zu können. Mit Fortsetzung des *Turns* (vgl. Z. 603-605) hält der Vater seine Geste auf dem von ihm etablierten und besprochenen Gesprächsgegenstand ‚Rom‘ (Z. 602-614), während der Sohn nun seine projizierte Turnfortsetzung abbricht, indem er seine Geste zurücknimmt (Z. 605)⁷⁹ und den Blick vom Material löst. Anschließend beteiligt er sich durch eine syntaktische Konstruktion an der Begründung des Vaters (Z. 606; vgl. Günthner 2015; Kreuz 2021). In der überlappenden Sequenz zwischen Vater und Sohn erwirbt also der Vater, der die Geste *konstant* auf das Material und den Blick auf den Sohn gerichtet hält und zugleich seinen *Turn* fortsetzt, das Rederecht und kann seine Begründung fortsetzen.

Dass das Etablieren und Präsenthalten von Gesprächsgegenständen bei konkurrierender Turnübernahme nicht an deiktische Gesten gebunden ist, illustriert das nächste Beispiel, in dem der Vater durch das Halten einer Zählgeste sowie die anschließende Turnfortsetzung sein Rederecht in einer Überlappung absichert und seine begonnene Aufzählung fortsetzt.

Beispiel 2 – Etablieren und Präsenthalten eines verhandelten Gesprächsgegenstandes bei konkurrierender Turnübernahme

581 VA: dann LASS uns noch ein paar argumente dafür sammeln-
 582 wie wir ANDere sonst davon überzeugen-
 583 dass |DAS |tatsächlich das beste is.

va-gs |tippt mit Stift auf Option|

584 VA: ((schiebt Aufgabenblatt Ri SO))

585 (--)

586 VIER tage hamburg mit musicalbesuch;

va-gz >>blickt zu SO (>Z. 595)

so-gz >>blickt auf Blatt (>Z. 589)

587 ERS[tens,]

588 SO: [die] [F:AHRT,]



589 VA: |[weil wir das] MU|sical- (1.0)

va-gs |zählt mit rDaumen----->

so-gz ----->|blickt zu VA (>Z. 595)

590 2äh |DU |kennst könig der löwen,=|

va-gs --->|zeigt mit Zählgeste Ri SO, Beat-|hält-----|

591 SO: =ja:-

⁷⁹ Der Sohn projiziert sowohl sprachlich (vgl. Z. 601) als auch durch die lange Vorbereitung der deiktischen Geste (vgl. Z. 598-602) sowie das anschließende Halten dieser (vgl. Z. 603-605) eine Fortsetzung seines *Turns* und somit das Beanspruchen des Rederechts (vgl. Mondada 2007; Oloff 2013). Erst mit Auflösen der Geste wird dem Vater das Rederecht zugestanden.

592 VA: |ICH kenne |könig der löwen und |TAR
 va-gs |zeigt m. Zählgeste & Beat kurz Ri VA| |lässt rH
 |zan,
 va-gs geschl. auf Tisch fallen|
 593 und BEI[des]
 594 SO: [BEIde] sind gu:t,
 595 VA: sind |!RICH!tig |TOLle musicals;
 va-gz ---->|blickt nach unten----|blickt zu SO->>
 so-gz ----->|blickt nach unten----->>

Mit dem Adverb „ERStens“ (Z. 587) wird der Beginn einer Aufzählung (vgl. Selting 2004) projiziert, deren erstes Listenelement in Form einer Begründung mit der kausalen Konjunktion *weil* sowie einer Zählgeste (Z. 589, vgl. #1) eingeführt wird. Dabei überlappt die Einführung des ersten Listenelements (Z. 587/589) mit einer versuchten Turnübernahme des Sohnes (Z. 588). Beide Sprecher brechen daraufhin ihre Äußerungen ab. Der Vater hält allerdings die Zählgeste und somit das eingebrachte Listenelement über die 1.0-sekündige Pause hinweg präsent und projiziert das weitere Beanspruchen des Rederechts sowie eine Turnfortsetzung der Begründung (vgl. Mondada 2007; Oloff 2013). Die anschließende Turnfortsetzung weist dann keinerlei Bezug zur Äußerung des Sohnes auf, denn der Vater nimmt das von ihm selbst mit der Zählgeste eingebrachte Element des Musicalbesuchs wieder auf. Die Zählgeste wird währenddessen weiter gehalten und zwischenzeitlich in deiktischer Funktion auf den Sohn (Z. 590) sowie auf den Vater selbst (Z. 592) gerichtet. Wie im ersten Beispiel bricht auch hier der Sohn seine Äußerung ab, wendet dann den Blick dem Vater zu (Z. 589) und markiert somit, dass das Rederecht dem Vater zugestanden wird. Anschließend ko-konstruiert er syntaktisch die abschließende Bewertung (Z. 594).

Zusammenfassend zeigen beide Beispiele, dass das Halten von Gesten in konkurrierend überlappenden Begründungssequenzen zusammen mit der Turnfortsetzung seitens der Sprecher*innen dazu beitragen können, das potenziell gefährdete Rederecht für sich durchzusetzen. Gerade in mündlichen Argumentationen erweist sich diese Funktion multimodaler Ressourcen, dass Interagierende ihre Begründung z. B. für oder gegen die Auswahl einer Option in der Interaktion platzieren können, als hochgradig relevant. Die gestische Ressource markiert dabei visuell wahrnehmbar das Präsenthalten des etablierten Gesprächsgegenstandes, sodass auch über Pausen hinweg (vgl. Beispiel 2) eine Fortsetzung des *Turns* projiziert wird. Insgesamt zeigt sich die multimodale Gestalt des Etablierens und Präsenthalten eines Gesprächsgegenstandes bei konkurrierender Turnfortsetzung als relevant für die Organisation des Sprecher*innenwechsels (vgl. Sacks et al. 1974): Sie übernimmt innerhalb der argumentativen Handlung des Begründens eine relevante Funktion für die sequenzielle Organisation der weiteren Interaktion.

8.2.2.3 Referenzieren und Präsenthalten eines verhandelten Gesprächsgegenstandes als Verstehensdokumentation (rezipient*innenseitig)

Innerhalb des Begründens zeigen Interagierende sich nicht nur an, dass Gesprächsgegenstände, die sprecher*innenseitig eingebracht worden sind von ihnen selbst präsent gehalten werden (vgl. Kap. 8.2.2.2), sondern Interagierende nutzen deiktische Gesten auf das Material auch *rezipient*innenseitig* zum Präsenthalten des Gesprächsgegenstandes sowie zur multimodalen Verstehensdokumentation (vgl. Deppermann und Schmitt 2008; Mondada 2011; Schmitt und Knöbl 2013): Indem die Rezipient*innen auf die Option verweisen, die von dem*der aktuellen Sprecher*in eingeführt worden ist, markieren sie ihrerseits, das sie erfasst haben, welcher Gesprächsgegenstand vom Gegenüber gerade verhandelt wird.⁸⁰ Der

⁸⁰ In Kapitel 3.4 wurden aus einer multimodalen konversationsanalytischen Perspektive die Begriffe „Sprecher*in“ und „Rezipient*in“ problematisiert. In dem vorliegenden Kapitel wird im Wissen dieser Problematik zur Be-

Begriff des *Verstehens* (i. S. der Verstehensdokumentation) wird dabei aus einer konversationsanalytischen Perspektive nicht als ein zu untersuchendes „psychisches Phänomen“ aufgefasst, wie Deppermann und Schmitt (2008: 222) ausführen, sondern vielmehr als ein interaktiver Prozess, in dem sich Interagierende gegenseitig z. B. ihre Interpretation der Vorgängeräußerung anzeigen. Auch die rezipient*innenseitige Referenzierung auf das Material innerhalb der Begründung des Gegenübers kann als solche Verstehensdokumentation verstanden werden. Damit wird ein Verstehen der *inhaltlichen* Begründung allerdings weder für die Beteiligten der Interaktion noch für Konversationsanalytiker*innen ableitbar (vgl. Mondada 2011: 543). Durch das Referenzieren auf die verhandelte(n) Option(en) demonstrieren Rezipient*innen also einerseits das Erfassen der verhandelten Bezugspunkte, zugleich zeigen sie sich aber auch als aktiv involviert in die Begründungssequenz des Gegenübers. Die multimodale Gestalt, wie sie in der nachfolgenden Übersicht (Tab. 14) beschrieben ist, übernimmt interaktionsorganisatorische Funktionen.

Referenzieren und Präsenthalten eines verhandelten Gesprächsgegenstandes als Verstehensdokumentation (rezipient*innenseitig)	
Beschreibung der multimodalen Gestalt	deiktische Geste auf Blatt während der Begründung des Gegenübers

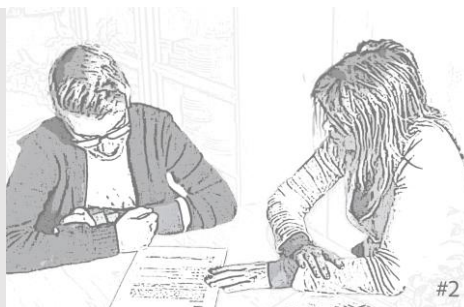
Tabelle 14: Beschreibung der multimodalen Gestalt „Referenzieren und Präsenthalten eines verhandelten Gesprächsgegenstandes als Verstehensdokumentation (rezipient*innenseitig)“

Im nachfolgenden Transkriptausschnitt legt die Mutter innerhalb einer sich im Aufbau befindenden Begründung des Sohnes zwei Finger auf die von ihm benannten Optionen und belässt sie dort bis zur eigenen Turnübernahme.

Beispiel 1 – Referenzieren und Präsenthalten eines verhandelten Gesprächsgegenstandes als Verstehensdokumentation (rezipient*innenseitig)

162	SO:	dann ham_wa nur noch EINS und dr ei.	...
	so-gs		
	so-gz	>>blickt auf Blatt (>Z. 169)	
	mu-gz	>>blickt auf Blatt->>	
163		(1.7)	
	so-gs	streicht m rIF und Daumen das Kinn, bewegt Daumen über Kinn (>Z. 167)	
164		(hm hm,)	
			
165		(0.7)	
	mu-gs	legt rH auf Blatt, IF & MF auf Optionen	
166	SO:	ja GU:T.=	
	mu-gs	 (>Z. 168)	
167		=vielleicht leute die: wirk lich SPASS dran haben==	
	so-gs	----->	

schreibung der argumentativen Handlung des Begründens allerdings auf sie zurückgegriffen, um zu verdeutlichen, dass hier die Geste aufseiten der Person erfolgt, die nicht an der sprachlichen Realisierung der Begründung beteiligt ist.



168		=auch vielleicht n bisschen Action da drin zu haben,
	mu-gs	 legt rH auf Blatt, IF & MF auf Optionen (>Z. 175)
169		also ich mein vielleicht n musi calbesuch da SITZte ja auch nur rum,=
	so-gz	-----> zu MU----- auf Blatt (>Z. 173)
170		=wär da vielleicht wirklich n bisschen (-) SPANnenderer (.) FÜNf tage
		lang (-) diesen segeltörn zu machen.
171		(2.6)
172		ich glaub das is unsere EINigung;=
173		=Oder?
	so-gz	blickt zu MU->
174	MU:	also ich GLAU:be, (.)
	mu-he	senkt Kopf nach vorn ab
	so-gz	----> auf Blatt-->>
175		dass das ne ganz (.) tolle SAcHe is. (.)
	mu-gs	----> tippt mit rIF auf dritte Option

Der Sohn benennt zu Beginn des Transkriptausschnitts resümierend die noch verbliebenen und zur Disposition stehenden Optionen (Z. 162) und markiert durch das Streichen über das Kinn (Z. 163-167) und die gefüllte Pause (Z. 164) eine Fortsetzung des *Turns*. Die Mutter legt nun die Hand auf das Material (Z. 165, #1), Index- und Mittelfinger verweisen jeweils auf die vom Sohn benannten Optionen – trotz der Gestenäußerung an einer potenziell turnübergaberelevanten Stelle zeigt sie keine Beanspruchung der Turnübernahme an. Vielmehr dient die Geste hier zur Anzeige der *online* Verstehensdokumentation, denn indem die Mutter auf die vom Sohn benannten Optionen verweist, markiert sie die Bezugsgegenstände der Äußerung des Sohnes als ‚nachvollzogen‘ bzw. ‚erfasst‘. Nachdem sie die Geste noch etwas weiter auf die entsprechenden Optionen ausgerichtet hat (vgl. 166-168), verbleibt die deiktische Geste bis zur eigenen Turnübernahme (Z. 174f.) auf den Optionen und hält die vom Sohn eingebrachten Optionen somit für beide Interagierenden sichtbar präsent. Dadurch zeigt sie sich selbst in der Rolle Rezipientin noch involviert in die sich im Aufbau befindende Begründung des Gegenübers (Z. 166-170). Die gestische Referenzherstellung fungiert allerdings nicht nur als Verstehensdokumentation, sondern sie dient auch dem Sohn als (*monitoring*) *Display* (vgl. Schmitt und Deppermann 2007; Oloff 2013): Die explizite deiktische Referenz auf die zwei Optionen ermöglicht dem Sohn die Passung der von der Mutter hergestellten Referenz zu überprüfen, d. h., inwiefern sie die von ihm eingebrachten Optionen ‚korrekt‘ nachvollzieht.

Zusammenfassend kann festgehalten werden, dass rezipient*innenseitig eingebrachte deiktische Gesten auf das Material als Verstehensdokumentation fungieren können. Sie machen sichtbar, inwiefern der*die Rezipient*in die aktuell verhandelten Gesprächsgegenstände nachvollzieht und zeigen zugleich seine*ihre Involvierung in die aktuelle Begründung des*der Sprecher*in an. Die Gesten in dieser multimodalen Gestalt übernehmen dementsprechend eine interaktive Funktion. Anders als beispielsweise bei Schegloff (1984: 271) beschrieben, bilden Gesten also kein „speaker’s phenomenon“, sondern können – wie multimodale Ressourcen insgesamt – von Interagierenden unabhängig von ihrer Sprecher*innenrolle funktional innerhalb kopräsender Interaktionen eingesetzt werden.

8.2.3 Multimodales Einführen einer Turnfortsetzung

In mündlichen Argumentationen können unterschiedliche Ressourcen zur Strukturierung einer Begründung eingesetzt werden. So können beispielsweise Konklusionen durch die Verwendung von Adverbkonnektoren wie *deshalb* eingeleitet, Themenwechsel mit einer Kopfbewegung in die entgegengesetzte Richtung der bisherigen Positur markiert (vgl. Reig Alamillo et al. 2012; vgl. auch Kendon 1972b) oder Übergänge zu neuen Abschnitten gestisch angezeigt werden (vgl. McNeill 1992; Kendon 2004; Graziano 2014b; vgl. auch Rönqvist und Lindström 2021; vgl. auch Kap. 4.2). *Palm Up* Gesten sind dabei beobachtbar

in association with passages in the verbal discourse which serve as an introduction to something the speaker is about to say, or serve as an explanation, comment, or clarification of something the speaker has just said. (Kendon 2004: 266; vgl. auch Graziano 2014b)

In mündlichen Argumentationen und unter Berücksichtigung sprachlicher wie auch körperlich-sichtbarer Ressourcen zeigt sich nun, dass *Palm Up* Gesten gemeinsam mit Konnektoren oder formelhaften Phrasen koordiniert werden, um eine Turnfortsetzung in Begründungen einzuführen (vgl. Tab. 15).⁸¹

Multimodales Einführen einer Turnfortsetzung	
Beschreibung der multimodalen Gestalt	sprachliches Strukturelement (z. B. formelhafte Phrase oder Konnektor) + <i>Palm Up</i> Geste

Tabelle 15: Beschreibung der multimodalen Gestalt „Multimodales Einführen einer Turnfortsetzung“

Das nachfolgende Beispiel 1 illustriert, wie sowohl formelhafte Phrasen als auch Konnektoren gemeinsam mit *Palm Up* Gesten Turnfortsetzungen einführen können. In Beispiel 2 wird außerdem eine Sequenz dargelegt, in der sich durch das Halten der *Palm Up* Geste in einem *Post-stroke Hold* nicht nur die Funktion der Geste, sondern die multimodale Gestalt selbst verändert. Denn durch das Halten der *Palm Up* Geste wird dann die gesamte Turnfortsetzung zur gemeinsamen Ansicht präsentierend angeboten (vgl. auch Kap. 8.2.1.1).

Die Sprecherin des ersten Beispiels baut in der vorliegenden Sequenz eine global strukturierte Begründung auf (Z. 185-191). Vermittels einer kurzzeitigen *Palm Up* Geste sowie einer zeitlich alignierten formelhaften Phrase (Z. 186) bzw. einer Konjunktion (Z. 189) führt die Sprecherin in neue Turnkonstruktionseinheiten ein und markiert somit die Fortsetzung ihres *Turns*.

Beispiel 1 – Multimodales Einführen einer Turnfortsetzung

183 TO: ROM,
 184 <<pp>ja was IS denn mit rom;>
 185 MU: <<nickend>ist auch am wei|testen entFERNT;=ne,>
 mu-gz ----->|blickt zu TO (>Z. 189)

⁸¹ Das multimodale Einführen einer Turnfortsetzung kommt in den vorliegenden Daten bis auf einen Fall ausschließlich in Begründungen vor. Aufgrund des frequenten Auftretens in dieser argumentativen Handlung wird die multimodale Gestalt in dem vorliegenden Kapitel vorgestellt.



#1

186 **die FR|Age [is-]**
 mu-gs |rH Palm Up auf Aufgabenblatt-->
 187 **TO: [ja;]**
 188 **MU: was HAT |man| von diesen drei tagen;=**
 mu-gs ----->|,,,|



#2

189 **|=UND |°h ähm:-**
 mu-gs |rH Palm Up auf Blatt|
 mu-gz ----->|auf Blatt ->>
 190 **man sitzt halt VIEL im bus drin=**
 191 **=und m im GRUNde genommen hat die geMEInSchaft (.) die KLASse (.)**
 nichts daVON; (-)

Auf die *denn*-Frage (vgl. Deppermann 2009) der Tochter (Z. 184) führt die Mutter eine Begründung an, die gegen die Auswahl der Option ‚Rom‘ (Z. 185) gerichtet ist. Die Sprecherin setzt nachfolgend ihren *Turn* durch zwei weitere Turnkonstruktionseinheiten fort (Z. 186-188, Z. 189-191), die sie jeweils mit einer *Palm Up* Geste sowie einem sprachlichen Strukturelement einführt: Der erste Anschluss erfolgt im Verbund einer formelhaften Phrase „die FRage is-“ (Z. 186) und einer *Palm Up* Geste (vgl. #1), die gemeinsam das durch die Phrase projizierte Fragenelement (Z. 188: was HAT man von diesen drei tagen;=) multimodal einführen. Im Anschluss an die syntaktisch vollständige Turnkonstruktionseinheit (Z. 186, 188) wird der *Turn* erneut fortgesetzt (Z. 189). Dabei bindet die Konjunktion *und* die Fortsetzung syntaktisch an die vorherige Turnkonstruktionseinheit zurück und führt gemeinsam mit einer erneuten *Palm Up* Geste (vgl. #2) ein weiteres Argument ein (Z. 190f.). Somit markiert die Sprecherin die Fortsetzung des *Turns* nicht nur sprachlich, sondern zeigt sie auch visuell wahrnehmbar für alle Beteiligten an. Durch die sequenzielle Position der Geste, ihre zeitliche Alignierung mit einem Konnektor und die dann anschließende neue Aussage übernimmt die Geste eine äußerungsstrukturierende Funktion.

Das zweite Beispiel zeigt, wie eine Sprecherin die satzförmige Ablehnung ihres Sohnes gegenüber der Option ‚Hamburg‘ elaboriert und dann in eine Konklusion überleitet. Diese führt sie mit einem Konnektor sowie einer *Palm Up* Geste ein. Durch das anschließende Halten der *Palm Up* Geste verändert sich die multimodale Gestalt, sodass die Konklusion auf der geöffneten Handfläche präsentiert wird.

Beispiel 2 – Multimodales Einführen einer Turnfortsetzung

064 **SO: da war ich ECHT schon SO [!OFT!-]**
 ((...))
 069 **MU: ICH würde sagen-**
 070 **wenn man nach HAMurg fährt==**


```

071      =das kann man |auch mit den ELtern |mit der faMilie machen;=|
      mu-gs      |.....|
      mu-gz      |blickt zu SO (>Z. 075)

```



072		=deshalb WARST du schon so oft- öffnet rH Palm Up----- ,,,,,,
073		weil wir so oft in [hamburg] WAREn;
074	SO:	[GENau;]
075	SO:	also: sch:ließen wir HAMBurg [schon mal aus.]
	mu-gz	----->
076	MU:	[HAMBurg schließen wir] aus.

Nach der Elaborierung einer Begründung gegen die Auswahl der Option ‚Hamburg‘ (Z. 069-071), schließt die Sprecherin dann mit dem Adverbkonnektor *deshalb* eine elliptische Konklusion an (Z. 072: deshalb WARST du schon so oft-), die sie abschließend mit einer weiteren Begründung stützt (Z. 073). Zeitlich koordiniert und bereits in Zeile 071 vorbereitet führt die Sprecherin mit einer *Palm Up* Geste sowie dem Adverbkonnektor *deshalb* in die Turnfortsetzung ein (Z. 072, vgl. #1). Indem die Sprecherin die Geste nun in einem *Post-stroke Hold* fortgesetzt hält, verändert sich die Funktion der multimodalen Gestalt vom multimodalen Einführen einer Turnfortsetzung zum Präsentieren von (komplexeren) Begründungen (vgl. auch Kap. 8.2.1.1). Es findet also eine Transformation der multimodalen Gestalt sowie ihrer damit verbundenen Funktion statt. Mit dem Ende der Intonationsphrase wird die Geste dann in eine Ruheposition überführt.

Zusammenfassend zeigt das vorliegende Kapitel, dass Interagierende in mündlichen Argumentationen in enger Koordination sprachlicher und gestischer Ressourcen eine Turnfortsetzung multimodal einleiten können: Auf sprachlicher Ebene wird die Äußerung durch Konnektoren oder formelhafte Phrasen angeschlossen und zeitlich mit einer kurzzeitigen *Palm Up* Geste koordiniert; gemeinsam führen sprachliche sowie körperlich-sichtbare Ressourcen die Turnfortsetzung der Begründung ein. Die zeitliche Alignierung von der *Palm Up* Geste und dem Konnektor, die gemeinsam eine neue Aussage einleiten, führen zu einer Äußerungsstrukturierung des Beitrags der Sprecher*innen. Anhand des zweiten Beispiels konnte dann außerdem eine Veränderung der multimodalen Gestalt durch das Halten der Geste in einem *Post-stroke Hold* illustriert werden.

8.2.4 Multimodale Kohäsion und Kohärenz

In Begründungen, in denen Sprecher*innen auch satzübergreifende Äußerungspakete aufbauen, stellen Interagierende auf vielfältige Weise Kohäsion (vgl. Halliday et al. 1980: 318ff.), d. h. lokale Beziehungen zwischen einzelnen Segmenten her: So können beispielsweise lexikalische und syntaktische Wiederholungen (vgl. u. a. Bublitz 2001) aufgrund ihrer Ähnlichkeit als zusammengehörige Elemente identifiziert werden; oder auf Ebene der Prosodie Intonationsphrasen aufgrund des „als kohäsiv wahrgenommenen Tönhöhenverlauf[s]“ (Selting et al. 2009: 309; vgl. auch Auer und Selting 2001) als zusammengehöriges Ganzes erkannt werden. Und auch Gestenwiederholungen können dazu beitragen, Äußerungssegmente als zusammengehörig zu markieren (vgl. Kendon 1980). McNeill (1992: 16) bezeichnet Gesten in dieser Funktion als *Cohesives* bzw. Kohäsionsgesten (vgl. Kap. 3.2.1). Dabei können Gesten

unabhängig von ihrer weiteren Funktion mit dem Ziel wiederholt werden, „to tie together thematically related but temporally separated parts of the discourse“ (McNeill 1992: 16), und somit Kohäsion innerhalb einer Äußerung anzeigen.^{82,83}

Bislang werden Gesten, die zur Kohäsionsmarkierung beitragen, vornehmlich in erwerbsorientierten Studien untersucht (vgl. u. a. Colletta und Pellenq 2009; Colletta et al. 2010; Reig Alamillo et al. 2012; Cristilli 2014; Colletta et al. 2015; vgl. auch Kap. 4.2.). Unter dem terminologischen Dach der kohäsionsmarkierenden Gesten (*Cohesive*, *Discursive* oder auch (*Deictic*) *Anaphoric Gestures*) werden dann nicht nur Gestenwiederaufnahmen gefasst, sondern auch Gesten, die z. B. eine lexikalische Kohäsionsmarkierung hervorheben. Aufgrund der quantitativen Ergebnispräsentation der Studien bleiben dann qualitative Einblicke aus, die darlegen, wie Gestenwiederaufnahmen zur Kohäsionsmarkierung beitragen können (vgl. aber Cristilli 2014). Dezidiert qualitative Studien zu Gestenwiederholungen fokussieren hingegen auf kohärenzstiftende Funktionen. Hier tragen die Gestenwiederaufnahmen u. a. zur Markierung von Anfang und Ende einer Erzählung bei (vgl. Heller 2018a), sodass sie insgesamt globale Zusammenhänge hervorheben (vgl. u. a. Levy und McNeill 2013; Ohlhus 2016; Heller 2018b).

Die Daten der vorliegenden Studie zeigen, dass in mündlichen Argumentationen Gestenwiederholungen bzw. Gestenwiederaufnahmen, i. S. der o. g. Definition McNeills (1992: 16) von *Cohesives*, besonders häufig in Begründungen genutzt werden, um sowohl Kohäsion als auch Kohärenz in größeren Äußerungspaketen zu markieren (vgl. auch Jacquin 2017).⁸⁴ Es zeigt sich, dass kohäsionsstiftende Gesten Verbindungen zwischen einzelnen aufeinander folgenden Äußerungsbestandteilen innerhalb größerer Äußerungspakete sichtbar markieren. Dabei werden die Gesten in einer deutlichen zeitlichen Nähe zueinander wiederaufgenommen – i. d. R. *in einer* oder *in* sequenziell *aneinander anschließenden* Intonationsphrase(n) bzw. Turnkonstruktionseinheit(en). Die wiederaufgenommenen Gesten ähneln sich in ihrer Form, Bewegungsrichtung sowie ihrer Ausrichtung. Kohärenzstiftende Gesten hingegen treten in größeren Äußerungspaketen in zeitlicher Distanz zueinander auf. Sie können beispielsweise in einzelne Äußerungssegmente eines Äußerungspakets einleiten und abschließend wiederaufgenommen werden. Dadurch wird es möglich, ein Äußerungssegment als zusammengehörige Einheit zu identifizieren (vgl. auch Ferré 2012; Ohlhus 2016; Heller 2018b). Da kohärenzstiftende Gesten allerdings in deutlicher zeitlicher Entfernung zueinander auftreten können (z. B. zu Beginn und am Ende eines Äußerungssegments), werden sie nicht allein durch die Ähnlichkeit der Form, Bewegungsrichtung sowie Ausrichtung der gestischen Ressource identifizierbar, sondern es bedarf u. a. auch einer Berücksichtigung thematischer bzw. inhaltlicher Wiederaufnahmen, in deren Kontext die Gesten wiederholt geäußert werden. Die folgende Übersicht 16 liefert eine zusammenfassende Beschreibung der multimodalen Gestalt.

⁸² In späteren Arbeiten von McNeill (2005) sowie Levy und McNeill (2013, 2015) wird der Begriff der *Cohesives* durch den Begriff der *Catchments* abgelöst, der in einer breiteren Definition die Wiederholung von „one or more gesture feature [...] in at least two (not necessarily consecutive) gestures“ (McNeill 2005: 116; vgl. auch Levy und McNeill 2015) bezeichnet

⁸³ Levy und McNeill (2013) sprechen selbst von Kohäsionsmarkierungen. In ihren Daten zeigt sich aber, dass Gestenwiederaufnahmen sowohl lokal die Verbindung zweier Äußerungen, i. S. von Kohäsion, markieren (vgl. Levy und McNeill 2013: 557f.) als auch zur Markierung von globaler Verbindungen, i. S. von Kohärenz, eingesetzt werden können, durch die zwei zeitlich getrennte Äußerungen als miteinander verbunden angezeigt werden können (vgl. Levy und McNeill 2013: 559ff.; vgl. auch Levy und McNeill 2015).

⁸⁴ Das Markieren multimodaler Kohäsion und Kohärenz tritt in den vorliegenden Daten grundsätzlich in größeren Äußerungspaketen wie dem Begründen und Einwenden auf. Zwar unterscheidet sich die argumentative Handlung, nicht aber die Form der multimodalen Gestalt, sodass sie an dieser Stelle für das Begründen vorgestellt wird.

Multimodale Kohäsion und Kohärenz	
Beschreibung der multimodalen Gestalt	• Gestenwiederaufnahmen, die sich in Form und/oder Bewegung ähneln + – ähnliche lexikalische/morpho-syntaktische Elemente – Wiederaufnahme inhaltlicher bzw. thematischer Elemente (obligatorisch für das Markieren von Kohärenz) an relevanter sequenzieller Position

Tabelle 16: Beschreibung der multimodalen Gestalt „Multimodale Kohäsion und Kohärenz“

Gestenwiederaufnahmen können gemeinsam im Verbund mit Konjunktionen (vgl. Beispiel 1) oder im Kontext fokusakzentuierter Lexeme (vgl. Beispiel 2) die Zusammengehörigkeit der Bestandteile einer Begründung kohäsiv markieren. Zeitlich getrennte Äußerungsbestandteile eines begründenden Äußerungspakets können wiederum durch lexikalische, gestische sowie inhaltliche Wiederaufnahmen i. S. der Kohärenz miteinander in Beziehung gesetzt werden (vgl. Beispiel 3).

Im Folgenden Beispiel 1 werden durch die Gestenwiederaufnahmen einer auf das Material ausgerichteten *Palm Up* Geste drei Bestandteile einer Begründung miteinander verbunden.

Beispiel 1 – Multimodale Kohäsion und Kohärenz

183 TO: ROM,
184 <<pp>ja was IS denn mit rom;>
185 MU: <<nickend>ist auch am wei|testen entFERNT;=ne,>
mu-gz ----->|blickt zu TO (>Z. 189)



186 die FR|Age [is-]
mu-gs |rH Palm Up auf Aufgabenblatt->
187 TO: [ja;]
188 MU: was HAT |man| von diesen drei tagen;=
mu-gs ----->|,,,|



189 |=UND |°h ähm:-
mu-gs |rH Palm Up auf Blatt|
mu-gz ----->|auf Blatt ->>
190 man sitzt halt VIEL im bus drin=-



191 |
 =und m |im | GRUNde genommen hat die geMEINschaft (.)
 mu-gs |rH Palm Up auf Blatt|
 die KLASse |(.)| nich |ts| daVON; (-)
 mu-gs |dreht rH vertikal auf-,,|
 mu-he |knickt Kopf nach re ab|
 mu-bo |Shoulder Shrug-----|
 192 °h ALso-
 193 oKAY.

In der vorliegenden Sequenz bringt die Mutter eine Begründung gegen die Auswahl der Option ‚Rom‘ ein. Dazu reiht sie ein Gegenargument, das metakommunikativ als W-Frage (vgl. u. a. Koshik 2003) gerahmt wird (Z. 186, 188), sowie eine das Gegenargument inhaltlich elaborierende, aber syntaktisch eigenständige Proposition (Z. 190f.) und die daraus resultierende Konklusion (Z. 191) aneinander, die sie jeweils zu Beginn mit einer auf das Aufgabenblatt ausgerichteten *Palm Up* Geste einführt (vgl. Kap. 8.2.3). In der Entfaltung der Äußerungen ist nun einerseits eine Segmentierung sowie andererseits die Herstellung von Kohäsion zu beobachten: Sowohl die metakommunikative formelhafte Phrase (Z. 186: die FRAGE ist) als auch die parataktischen Konjunktionen (Z. 189, 191: UND / und) markieren drei syntaktisch voneinander abgrenzbare, eigenständige Segmente, die zeitlich koordiniert mit drei Gesten auch visuell sichtbar als jeweils einzelne Bestandteile des Äußerungspakets erkennbar werden. Zugleich werden die einzelnen Bestandteile der Begründung syntaktisch sowohl durch die Konjunktion *und* sowie die Wiederaufnahme der Gesten, die in ihrer Form sowie der Ausrichtung gleich sind, miteinander verbunden: Die drei Bestandteile der Äußerung werden hier durch die gleichen Ressourcen, die sie *segmentieren*, gleichzeitig als *zusammengehörig* markiert. Die multimodale Gestalt übernimmt also eine pragmatische Funktion, da durch sie zwischen den einzelnen Äußerungsbestandteilen der Sprecherin Kohäsion erzeugt wird.

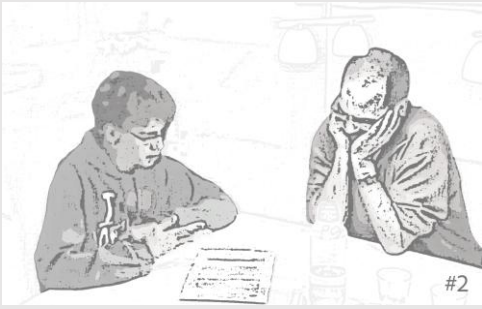
Innerhalb des zweiten Beispiels werden die Gestenwiederaufnahmen in der Umgebung fokusakzentuierter Lexeme und jeweils eher am Ende der jeweiligen Intonationsphrase produziert.

Beispiel 2 – Multimodale Kohäsion und Kohärenz

046 SO: also ICH würd ja sagen-
 so-gz >>blickt auf Blatt (>Z. 049)
 va-gz >>blickt auf Blatt->>
 047 ähm-
 048 wir nehmen den: (.) BUNgalowpark,=



049 =denn |es |is ja (.)|von der |SCHU
 so-gs |.....|öffnet rH Palm Up|laterale Bewegung, krümmt MF,
 so-gz ----->|blickt zu VA----->|blickt auf Blatt----->
 |le her,|
 so-gs RiFi, klFi|,,,,,,|
 so-gz ----->



050 und ähm (1.0) FREIzeitmöglichkeiten mit |SHOP |pingl|
 so-gs |öffnet rH Palm Up|,,,,,|
 so-gz ----->
 und |s schwImmen;=|
 so-gz --->|blickt zu VA-|

051 =is Eher was-
 052 als ähm (-) auf_nem SEgelschiff (.) da (-) fünf tage zu sein;
 053 | (3.0)
 so-gz |blickt zu VA->>

In dem vorliegenden Transkriptausschnitt elaboriert der Sohn seine Position (Z. 046-048) anhand einer anschließenden Begründung (Z. 049-052). Diese besteht aus insgesamt zwei Argumenten, wovon das erste (Z. 049: denn es is ja (.) von der SCHUle her,) nicht weiter erläutert wird und somit inhaltlich opak bleibt, und innerhalb des zweiten Arguments (Z. 050-052) ein Vergleich zwischen zwei Optionen ausgeführt wird. Dabei wird zwischen dem ersten und dem zweiten Argument der Begründung nun durch mehrere Ressourcen Kohäsion hergestellt: Das erste Argument endet mit einer steigenden Tonhöhenbewegung, die eine Fortsetzung den *Turns* erwartbar macht. Außerdem werden beide Argumente durch eine Konjunktion (Z. 050: und) aneinander angeschlossen, sodass auch syntaktisch die Verbundenheit der Bestandteile der Äußerung markiert wird. Ferner trägt auch der Einsatz der gestischen Ressourcen dazu bei: Zwei *Palm Up* Gesten, die jeweils am Ende der Intonationsphrasen (Z. 049, 050) in der Umgebung fokusakzentuierter Lexeme produziert werden, markieren aufgrund ihrer Ähnlichkeit der Form sowie der Gestenausrichtung (vgl. #1, #2) wahrnehmbar die Zusammengehörigkeit beider Äußerungsbestandteile. Die zweite Geste bildet – retrospektiv – eine Gestenwiederholung zur ersten *Palm Up* Geste.

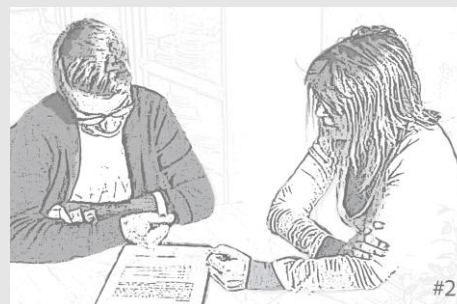
In dem letzten Beispiel wird nun innerhalb eines längeren Äußerungspakets durch eine Wiederaufnahme lexikalischer, gestischer sowie inhaltlicher Elemente ein *kohärenzbildender* Rahmen um einen Teil des Äußerungspakets gespannt.

Beispiel 3 – Multimodale Kohäsion und Kohärenz

196 MU: und warum wärste da AUCH eher für?
 197 SO: weil es ja auch so: (.) ACTION is;=ne,
 198 weil du: KANNST dich ja bewEgen-
 199 du: (.) MACHST da ja wat;



200 |du sitzt da ja wie gesacht |nich |nur |in_nem
 so-gz |blickt auf Blatt (>Z. 205)
 so-gs |.....|Beat m. Faust ü. Blatt|
 |SESSel-
 mu-gz |blickt auf Blatt->>
 201 un:d- (0.8)|
 mu-gs ----->|
 202 JA.=
 203 =ich glaub dat würde MEHR interessieren als n_Musicalbesuch |zum
 beispiel.
 so-gz ----->|blickt
 zu MU--->
 204 und das würde das auch zuSAMMEN|schmeißen;=
 so-gz ----->|
 205 =WEIL man ja auch als GRUP|pe zusammenarbeiten mus|s.
 so-gz ----->|blickt zu MU (>Z. 207)
 mu-gs |...->
 206 |(0.5) |
 mu-he |nickt-----|
 mu-gz |blickt zu SO->
 mu-gs|



207 |und nich |zum bei |spiel |nu:r |(0.3) |
 so-gz |blickt auf Blatt----->
 so-gs |.....|Beat m. Faust ü. Blatt|
 mu-gs |öffnet rH Palm Up m. gekr. Fi|
 mu-he |nickt----|
 mu-gz ----->|blickt auf Blatt-----|
 |(0.5) in den SITZ|en sitzt.
 so-gz ----->|blickt zu MU->
 mu-gz |blickt zu SO----->>
 208 MU: genau.|
 so-gz ----->|
 209 SO: hm_hm,|

In der vorliegenden Begründung (Z. 197-207), die der Sohn in Folge eines globalen Zugzwangs der Mutter (Z. 196) ausführt, markiert die Gestenwiederaufnahme einer *Beat* Geste (vgl. #1, #2) über dem Material einen Zusammenhang zwischen einzelnen Segmenten des Äußerungspakets: Nachdem die ersten Argumente der Begründung eingebracht worden sind, die sich auf den ‚Segeltörn‘ beziehen (Z. 197-199), stellt der Sohn nun (Z. 200) einen Bezug zu dem von ihm ursprünglich favorisierten Ziel ‚Hamburg‘ samt Musicalbesuch her. Dabei kontrastiert er den Aspekt der „Action“ (Z. 197), den er für den Segeltörn positiv herausstellt, mit der sitzenden Aktivität während des Musicalbesuchs (Z. 200: du sitzt da ja wie gesacht nich nur in_nem SEssel-), sodass eine Ablehnung der Option ‚Hamburg‘ erkennbar wird. Das Adverb *nur* verweist hier auf die eingeschränkten Aktivitätsmöglichkeiten, die sich aus dem Besuch des Musicals ergeben. Es wird mit einer *Beat* Geste, die auf das Material hin orientiert ist, besonders hervorgehoben. Nach weiteren kontrastierenden Ausführungen zu den Optionen des Segeltörns und des Musicalbesuchs (Z. 201-205) kommt der Sprecher am Ende des Begründungssegments erneut auf die sitzende Aktivität zurück (Z. 207). Dabei werden folgende Wiederaufnahmen zwischen der Zeile 200 und der Zeile 207 sichtbar: lexikalische sowie partielle syntaktische Wiederaufnahmen (Z. 200: sitzt [...] nich nur / Z. 207: nich [...] nu:r [...] sitzt), eine semantische Wiederaufnahme von Lexemen, die auf Sitzmöbel verweisen und die jeweils prosodisch durch Fokusakzente hervorgehoben werden (Z. 200: SEssel / Z. 207: SITZen), sowie die Wiederaufnahme des *Beats*, der über dem Material und aligniert mit dem Adverb *nur* geäußert wird. Das einleitende Argument (Z. 200) wird hier als Konklusion der Begründung (Z. 207) rahmenspannend wiederaufgenommen. Die Kohärenz zwischen Anfang und Ende des Begründungssegmentes wird durch die gestische Wiederaufnahme für die Beteiligten auch sichtbar angezeigt.

Zusammenfassend lässt sich für das vorliegende Kapitel festhalten, dass die Wiederaufnahme von Gesten, gemeinsam mit weiteren Ressourcen, dazu beiträgt, sequenziell aufeinander folgende, einzelne Äußerungsbestandteile (segmentierend) erkennbar zu machen und zugleich durch die gleiche Form, Bewegungsrichtung und Ausrichtung der Geste die Kohäsion der Bestandteile zu markieren (vgl. Beispiel 1-3). Dabei können die Gesten sowohl mit Strukturelementen der Äußerung, wie Konjunktionen, und durch deren Position in der Turnkonstruktionseinheit (vgl. Beispiel 1) oder mit fokusakzentuierten Bedeutungselementen (vgl. Beispiel 2) aligniert sein. Trotz der variantenreichen Realisierung der Markierung von Kohäsion zeigt sich, dass die Gesten in ihrer Form, Bewegungsrichtung sowie Ausrichtung sich entsprechend äquivalent zueinander verhalten. Diese Kongruenz sowie die zeitliche, d. h. auch sequenzielle Nähe der geäußerten Gesten tragen entscheidend zur Rekonstruktion und Identifikation dieser Funktion von Gestenwiederholungen bei. In längeren Äußerungssegmenten, in denen die Gestenwiederaufnahme nicht in unmittelbarer zeitlicher Nähe stattfindet und durch diese identifiziert werden kann, bedarf es zusätzlich der Berücksichtigung des sequenziellen Kontextes (z. B. Einleiten in ein neues Begründungssegment) sowie inhaltlicher bzw. thematischer Wiederaufnahmen, die zur Identifizierung der multimodalen Gestalt beitragen. Kohärenzstiftende multimodale Gestalten können dabei z. B. neu eingeführte Argumente einer Begründung rahmenbildend umschließen (vgl. Beispiel 3). Insgesamt tragen Gestenwiederaufnahmen, eingebettet in die gesamte multimodale Gestalt, sowohl in ihrer kohäsions- als auch kohärenzstiftenden Funktion zur Strukturierung der Begründungen bei.

8.2.5 Verwerfen in Begründungen

Innerhalb von mündlichen Argumentationen können Sprecher*innen nicht nur sichtbar Gesprächsgegenstände in Begründungen etablieren und präsent halten (vgl. Kap. 8.2.2), sondern Interagierende können auch das Verwerfen beispielsweise von (antizipierten Gegen-)Argumenten oder konkreten Optionen multimodal anzeigen. Besonders funktional erweisen sich dafür Gesten der Familie der *Away* Gesten, mit denen Interagierende Gesprächsgegenstände, Themen etc. aus der weiteren Interaktion ausschließen

können (vgl. Bressemer und Müller 2014a, 2014b). In den vorliegenden Daten verwenden die Sprecher*innen die *Brushing* und *Throwing Away* Gesten, die die Funktion teilen, „to clear the immediate surrounding of disturbing and useless objects“ und „to mark arguments, ideas and actions as uninteresting and void“ (Bressemer und Müller 2014b: 1599). Interagierende nutzen die Gesten in den vorliegenden Daten, um bestimmte Gesprächsgegenstände aus der weiteren Interaktion auszuschließen bzw. zu verwerfen (vgl. Tab. 17). Diese Ablehnung wird i. d. R. auch sprachlich expliziert.

Verwerfen in Begründungen	
Beschreibung der multimodalen Gestalt	• <i>Throwing Away</i> bzw. <i>Brushing Away</i> Gesten + Benennung ausgeschlossener Gesprächsgegenstände

Tabelle 17: Beschreibung der multimodalen Gestalt „Verwerfen in Begründungen“

Die nachfolgenden Ausführungen zeigen nun, wie das Verwerfen in Begründungen multimodal realisiert wird. Beispiel 1 illustriert das Verwerfen eines antizipierten Gegenargumentes während des Aufbaus einer Begründung. Beispiel 2 legt dar, dass eine Option projizierend aus der weiteren Interaktion ausgeschlossen wird.

Das erste Beispiel illustriert, wie die Mutter dem Ausschluss einer Option, den der Sohn vorgeschlagen hat, zustimmt und in einer anschließenden Begründungssequenz ein *antizipiertes Gegenargument* verwirft.

Beispiel 1 – Verwerfen in Begründungen

081 SO: SO:;
 082 ((liest sich die Aufgabenstellung durch, 4.6 Sekunden))
 083 ich würd auf jeden fall SA:gen,
 084 dass äh::m das ZWEITE würd ich AUSSchließen;
 085 (1.4)
 086 mit diesem DREI tage rom mit kulturprogramm,
 087 WEIL (.) das is nichts für (.) äh welche (-) in dem ALter;
 088 MU: °h (.) also ich hätte jetzt rOm AUCh ausgeschlossen;
 089 aber aus nem anderen GRUND-
 090 |weil ich FINde;
 so-gz |blickt zu MU-->
 091 dass_s für die drEi tage rom sehr weit ent|FE:RNT is;=
 so-gz ----->|
 092 SO: =ja DAS stimmt;=
 093 MU: =ne,
 094 UND ähm (-) °h wahrSCHEINlich is jetzt so_n (.) |selbst so_n|
 mu-gs |.....|



|FLUG | würde jetzt ja viel zeit |in anspruch nehmen;=
 mu-gs |LH Throwing Away|
 so-gz |blickt zu MU----->>>
 095 MU: [und] das_s ja dann auch viel zu TEUer für so_ne fahrt;
 096 SO: [ja;]
 097 MU: °h <<nickend>>DARum hätt ich das jetzt (.) a:uch nich> in erwägung
 gezogen;
 098 SO: ja das stimmt AUCh;

099 °h (-) äh:m:-
 100 für den:: (--) VIER tage hamburg mit musicalbesu:ch,
 101 bin ich AU_nicht wirklich;

In der vorliegenden Sequenz schließt sich die Mutter dem Ausschluss der Option ‚Rom‘ an, den der Sohn zuvor vorgeschlagen hat (Z. 083ff.), erklärt allerdings in einem metakommunikativen Kommentar einen anderen Grund als Ursache (Z. 088f.). Diesen führt sie sodann, eingeleitet durch die Konjunktion *weil* sowie das modalisierende *verbum putandi* (*ich*) *finde*, an (Z. 090f.: *weil* ich FINde; / dass_s für die drEi tage rom sehr weit entFE:RNT is;) und antizipiert dabei in der Fortsetzung ihrer Begründung nun einen potenziellen Einwand Dritter (Z. 094: und ähm (-) °h wahrSCHEINlich is jetzt so_n (.) selbst so_n FLUG würde jetzt ja viel zeit in anspruch nehmen;). Den antizipierten Einwand entkräftet sie unmittelbar mit einem Verweis darauf, dass auch der Flug viel Zeit (Z. 094) und Geld (Z. 095) kosten würde. Zeitlich koordiniert mit dem fokusakzentuierten Kernelement des antizipierten Gegenarguments (Z. 094: FLUG) zeigt nun die *Throwing Away* Geste (vgl. #1, #2) den bereits inhaltlich entkräfteten Einwand als verworfen an. Die Geste wird auch vom Gegenüber wahrgenommen (vgl. Z. 094, #1). Im Anschluss an eine metadiskursive, rahmenbildenden Konklusion der Mutter (Z. 097: °h <<nickend>>DARum hätt ich das jetzt (.) a:uch nich> in erwägung gezogen;), mit der sie an ihre eingangs erbrachte Zustimmung zum Ausschluss der Option anknüpft (Z. 088: °h (.) also ich hätte jetzt rOm AUCH ausgeschlossen;), lassen die Interagierenden nun die gesamte Option fallen und wenden sich der Aushandlung einer weiteren Option zu.

Innerhalb des zweiten Beispiels wird das Verwerfen einer *Option* unmittelbar nach Benennen der Option gestisch projizierend angezeigt und die Begründung anschließend fortgesetzt.

Beispiel 2 – Verwerfen in Begründungen

181 SO: am ende dann noch mal (--) zuSAMmenhalt beweisen; (.)
 182 meinste;
 183 VA: joa und n dann könnt ihr auch noch mal bevor ihr auseinandergeht was
 zuSAMmen unternehmen;=|ne,=
 va-gs |...->
 va-gz >>blickt auf Blatt----->
 184 =hier beim (.) |bei diesem |°h| (-) äh bei diesem BUNgalowpark
 va-gs ----->|tippt mit Stift|,,|
 va-gz ----->



|
 |(.)| da |(.)| da ist ja nichts zusammen.|
 va-gs |rH Brushing Away|,,|
 va-gz --->|blickt zu SO----->|
 185 SO: [ja.]
 186 VA: [s kein] is ja kein echter AB[schluss.]=
 187 SO: [stimmt.]
 188 VA: =wenn |da JE|der- |
 va-gs |....|bewegt bHde Palm Up auseinander|
 189 (1.1)
 190 VA: und in diesem |BUN |galowparks (.) könnt ihr halt auch nicht
 va-gs |dreht bHde auf|
 zusammen Abhängen,
 191 alle zusammen zumindest.

192 und HIER müsst_a wirklich- (---)
 193 SO: [jo;]
 194 VA: [hier] müssen alle mit ins RAD packen;

Die Interagierenden haben zuvor (nicht im Transkript enthalten) ko-konstruierend eine Begründung für die Option des Segeltörns aufgebaut. Kontrastierend hat dabei der Vater auch Argumente gegen den vom Sohn eigentlich favorisierten Bungalowpark eingebracht. Diesen Kontrast greift der Sprecher nach einer Rückfrage des Sohnes (Z. 181) erneut auf und initiiert weitere Argumente gegen den Bungalowpark. Nach der Benennung der Option, der einige problemanzeigende Signale (vgl. Birkner et al. 2020: 376) in Form von gefüllten und ungefüllten Pausen sowie Reparaturen (Z. 184) vorauslaufen, produziert der Vater mit dem anaphorisch verwendeten Lokaladverb *da* (Z. 184) eine *Brushing Away* Geste, mit der das Verwerfen der Option *projizierend* angezeigt wird. Nach einer Reparatur (Z. 184: (.)) da (.)) da wird der Ausschluss der Option kontrastierend zum Segeltörn dann auch begründet (Z. 183: joa und n dann könnt ihr auch noch mal bevor ihr auseinandergeht was zuSAMmen unternehmen;=ne, / 184: bei diesem BUNgalowpark (.)) da (.)) da ist ja nichts zuSAMmen.). Anschließend lässt der Sprecher die Option fallen und wendet sich dem ‚Segeltörn‘ zu (Z. 194).

Zusammenfassend zeigt sich, dass *Throwing* und *Brushing Away* Gesten dazu beitragen, in Begründungen das Verwerfen antizipierter Gegenargumente oder einzelner Optionen zu markieren – entweder gleichzeitig mit dem Aufbau einer Begründungssequenz (vgl. Beispiel 1) oder aber dieser vorauslaufend (vgl. Beispiel 2). I. d. R. wenden die Interagierenden sich anschließend anderen Gesprächsgegenständen zu.

8.2.6 Zusammenfassung

Zusammenfassend konnten in dem vorliegenden Kapitel die nachfolgenden multimodalen Gestalten, die an der Realisierung der argumentativen Handlung des Begründens beteiligt sind, rekonstruiert werden.

- In Kapitel 8.2.1 wurde gezeigt, dass sowohl lokal als auch global strukturierte Begründungssegmente, die u. a. mit Diskursmarkern oder Phrasen eingeleitet werden, im Verbund mit einer gehaltenen *Palm Up* Geste in mündlichen Argumentationen präsentiert werden können (vgl. Kap. 8.2.1.1). Auch einzelne Bedeutungselemente können durch *Palm Up* Gesten als besonders relevant für die Begründung kurzzeitig präsentierend hervorgehoben werden (vgl. Kap. 8.2.1.2). Dabei übernehmen die multimodalen Gestalten jeweils eine performative Funktion, da durch sie kontextualisiert wird, dass eine Begründung oder ein Begründungssegment präsentiert wird. Zugleich kommt der *Palm Up* Geste auch eine äusserungsstrukturierende Funktion zu, da sie einen Teil des Äußerungspakets als abgrenzbares Segment markiert bzw. ein einzelnes Element der Äußerung vom Rest der Äußerung abhebt.
- In Kap. 8.2.2 wurde gezeigt, dass das Halten von deiktischen sowie (repräsentierenden) Zählgesten in Abhängigkeit von dem sequenziellen Kontext einen Gesprächsgegenstand etablieren und äusserungsintern sowie -übergreifend präsent halten kann. Wenn die Geste außerdem turnfinal gehalten wird, entsteht für den*die anschließende*n Sprecher*in eine konditionelle Relevanz (vgl. Schegloff 1984), die präsent gehaltene Option bzw. die präsent gehaltenen Kriterien weiter zu verhandeln (vgl. Kap. 8.2.2.1). Präsent gehaltene Gesten können auch zur Durchsetzung des Rederechts bei simultaner, konkurrierender Turnfortsetzung beitragen. Denn der Ressourceneinsatz unterstützt die Turnübernahme und das Durchsetzen des Rederechts insofern, als dass er anzeigt, dass der*die Sprecher*in nicht von der Fortsetzung des *Turns* zurücktritt. Der Bezugsgegenstand der Äußerung verbleibt dabei satzübergreifend präsent (vgl. Kap. 8.2.2.2). Das Referenzieren auf einen vom Gegenüber eingeführten Gesprächsgegenstand

sowie das anschließende Präsenthalten kann aber auch der rezipient*innenseitigen Verstehensdokumentation dienen, die zugleich eine aktive Involviertheit an der Begründung des Gegenübers markiert. Durch die für alle Beteiligten sichtbare Referenzherstellung ist dem*der aktuellen Sprecher*in außerdem möglich zu überprüfen, ob die richtigen Bezugspunkte von der*dem Rezipient*in identifiziert worden sind (vgl. Kap. 8.2.2.3).

- Interagierende können Turnfortsetzungen, die z. B. mit Konnektoren oder formelhaften Phrasen, d. h. sprachlichen Strukturelementen eingeleitet werden, vermittelt einer kurzzeitigen *Palm Up* Gesten multimodal einführen (vgl. Kap. 8.2.3). Wird die Geste anschließend in einem *Post-stroke Hold* gehalten, kann sich ihre Funktion sowie die multimodale Gestalt insofern verändern, als dass dann die gesamte Turnfortsetzung präsentiert wird (vgl. dazu auch Kap. 8.2.1.1).
- Zur Äußerungsstrukturierung tragen auch Gestenwiederaufnahmen bei, die aufgrund ihrer Form und/oder Bewegung sowie sequenziellen Nähe zueinander kohäsionsstiftende Eigenschaften übernehmen können. Gestenwiederaufnahmen können aber auch zu einer satzübergreifenden Kohärenz beitragen und beispielsweise thematisch abgrenzbare Argumente rahmenbildend umspannen (vgl. Kap. 8.2.4).
- Die Analysen aus Kapitel 8.2.5 zeigten zuletzt, dass *Throwing* und *Brushing Away* Gesten dazu beitragen, in Begründungen das Verwerfen von Gesprächsgegenständen, z. B. von antizipierten Gegenargumenten oder Optionen, zu markieren. In der interaktiven Konsequenz zeigt sich, dass diese dann von den Interagierenden (vorerst) fallengelassen werden.

8.3 Zustimmungen

In argumentativen Entscheidungsfindungen können Interagierende sich durch gegenseitiges Zustimmung zu eingebrachten Positionen oder Begründungen einen Konsens anzeigen. Dabei sind Zustimmungen immer als reagierende Handlungen auf eine Vorgängeräußerung zu verstehen, die an unterschiedlichen sequenziellen Positionen beispielsweise entweder lokale Aushandlungssequenzen oder aber, global, den Kernjob des Begründens abschließen können (vgl. Heller 2012: 91). Dabei können zustimmende Handlungen entweder mit Nachdruck eingebracht werden oder aber als sogenanntes „weak agreement“ (Pomerantz 1984: 64) bzw. schwaches Zustimmung in Konzessionen einem Einwand oder Widerspruch vorauslaufen (vgl. Kap. 2.2.2; vgl. Pomerantz 1984; Couper-Kuhlen und Thompson 2000; Heller 2012: 91; Grundler et al. 2020).

Das vorliegende Kapitel zeigt, wie körperlich-sichtbare und sprachliche Ressourcen gemeinsam unterschiedliche multimodale Gestalten des Zustimmens konstituieren, die wiederkehrend vorkommen: In ausgebauten zustimmenden Sequenzen werden deiktisch-tippende bzw. ostentativ-deiktische Gesten eingesetzt, um eine Zustimmung nachdrücklich zu verstärken (vgl. Kap. 8.3.1); das Zustimmung in Konzessionen (vgl. Kap. 8.3.2) sowie zu evidenten Begründungen und Positionen (vgl. Kap. 8.3.3) wird abgeschwächt realisiert.

8.3.1 Nachdrückliches Zustimmung

In den vorliegenden Daten werden Zustimmungen mit Nachdruck hervorgebracht, wenn der Konsens aus einer vorauslaufenden Opposition heraus entsteht. Diese Opposition nun (einlenkend) auflösend erfolgt die Zustimmung zum Gegenüber besonders eindringlich, indem der*die Sprecher*in die Affirmation als ausgebaute(re) Turnkonstruktionseinheit – und nicht mittels Form einfacher Affirmations- und Negationsmorpheme – gemeinsam mit deiktisch-tippenden bzw. ostentativ-deiktischen Gesten auf das Material hervorbringt (vgl. Tab. 18). Dadurch stellen die gestischen Ressourcen einerseits einen Bezug

zu der Option her, auf die sich das Beipflichten bezieht. Andererseits verleihen sie der Zustimmung durch die iterative Bewegung des Tippens sowie das ostentative Referenzieren auf die Option besonderen Nachdruck. Da mit der Zustimmung „eine vorangegangene Behauptung oder Begründung gestärkt [wird]“ (Heller 2012: 91), die im vorliegenden Kontext immer an eine vorangegangene Opposition anschließt, ist anzunehmen, dass das *nachdrückliche* Zustimmung erforderlich wird, um dem Gegenüber ‚glaubhaft‘ den Konsens zu versichern und dabei den eigenen epistemischen Access zu der entsprechenden Position zu markieren.

Nachdrückliches Zustimmung	
Beschreibung der multimodalen Gestalt	• (ausgebaute) sprachliche Zustimmung + iterative bzw. ostentative deiktische Geste auf das Material • anschließend an vorangegangene Opposition

Tabelle 18: Beschreibung der multimodalen Gestalt „Nachdrückliches Zustimmung“

Die nachfolgenden Beispiele legen dar, wie Interagierende ihrem Gegenüber sowohl in satzförmigen als auch satzübergreifenden Äußerungen zustimmen und durch deiktisch-tippende (vgl. Beispiel 1) bzw. ostentativ-deiktische Gesten auf das Material (vgl. Beispiel 2) ihre Zustimmung nachdrücklich hervorheben können.

In dem nachfolgenden Beispiel wendet der Sohn gegen Option ‚Hamburg‘ ein und stärkt anschließend seine Position für den ‚Bungalowpark‘. Die Mutter, die zuvor Einwände gegen den Bungalowpark vorgebracht hat, stimmt dem Sohn nun zu. Durch die mehrmals deiktisch-tippenden Gesten auf die Option wird dabei ihre Zustimmung nachdrücklich verstärkt.

Beispiel 1 – Nachdrückliches Zustimmung

```

153 SO: und wenn du in so_nem MUsicalbesuch bis oder sowas,
154 wo du einfach nur ZWEI: (.) bis zweieinhalb stunden da sitzt,
155 und einfach nach VORne guckst-
156 is jetzt vielleicht nicht SO: toll.
157 MU: ^h ja gut [ich mein es sind ja] vier TA:ge hamburg;=ne?
((...))
161 SO: ja: Aber h°-
162 (---)
163 TROTZdem.=
164 =was MACHSten da sonst so;
165 da kAnnste vielleicht in die STADT gehen,
166 und dA n_bisschen was SHOPpen.=
167 =aber was ist denn MIT, (-) |
so-bo |Shoulder Shrug|
168 wenn da jemand KEIN bock [drauf hat, ]
169 MU: [|NEE also ich]
mu-he |nickt----->
170 [|ich |seh AUCH.] |
mu-he ----->|
mu-gs |zeigt mit lIF auf Option-|tippt mehrmals|
171 SO: [da kannse_s ] (.) das KANNse dann auch (.) in
so_m äh: park machen,
172 aber GLEICHzeitig auch noch was anderes.
173 (is/doch) viel BESSer;
174 MU: also |ICH SEH |das auch.|
mu-he |nickt-----|
mu-gs |.....|
175 |ich würd das |AUCH machen.
mu-gs |tippt schnell mit lIF auf Blatt|.....->
176 nur |WIE gesagt-
mu-gs ....|tippt mit lIF auf Hamburg->>
177 ich wollt nur SAgen;
178 das is ja nicht nur das: (.) MUsical,
179 [sondern] das sind ja die: (.) [VIER tage;=ne?]

```

Nachdem der Sohn den eingebrachten Einwand der Mutter (Z. 157) entkräftet hat (Z. 164-168), erfolgen zwei Zustimmungssequenzen (Z. 169-170, Z. 174-175): Im Anschluss an die erste Zustimmung der Sprecherin (Z. 169-170), die sie durch die affirmierende Negation und Nicken, das fokusakzentuierte und Zustimmung markierende Adverb sowie durch iteratives Tippen anzeigt, erwirbt sie das Rederecht vorerst nicht. Es erfolgt sodann ein zweiter Versuch der Turnübernahme (Z. 174-175). Nachdem der Sohn in einem *Multi-Unit Turn* gegen die Mutter eingewandt hat, schließt er diesen mit einer evaluierenden Konklusion und Fürsprache für eine Abschlussfahrt in einen Bungalowpark ab (Z. 171-173), sodass eine turnübergaberelevante Stelle entsteht. Nach einer ersten, elliptisch verbleibenden Zustimmung (Z. 174: also ICH SEH das auch.) erfolgt dann eine syntaktisch vollständige Äußerung (Z. 175: ich würd das AUCH machen.), in der die Sprecherin ihrem Gegenüber durch die Verwendung des fokusakzentuierten Adverbs *auch* Zustimmung signalisiert. Riskant ist für die Sprecherin hierbei, dass sie ihre eigene oppositionelle Position vorerst aufgibt und die ihres Gegenübers einnimmt. Mit dem Einsetzen der Intonationsphrase beginnt sie nun, mehrmals schnell auf die Option des Bungalowparks auf dem Blatt zu tippen. Das riskante Umschwenken von einer ursprünglichen Opposition zu einem Konsens wird durch die deiktische-tippende Geste mediiert: Die Geste trägt nicht nur zur Desambiguierung des unterspezifizierten Lexems *das* bei, mit dem der Bungalowpark als Bezugsobjekt ihrer Zustimmung ersichtlich wird. Durch das mehrmalige Tippen auf das Material wird dem favorisierten Ziel des Sohnes und der Begründung für den Bungalowpark außerdem *mit Nachdruck* zugestimmt, denn „[r]ather than pointing *once* on the document, individuals *tap* on the document repeatedly, thereby *repeating the production of the point*“ (Hindmarsh und Heath 2000: 1868, Hervorh. N. S.). Indem die Position durch körperlich-sichtbare Ressourcen der iterativen Geste mit Nachdruck eingebracht wird, wird die ‚fremde‘ Position in eine ‚eigene‘ Position der Sprecherin überführt. Durch die Geste wird der epistemische Access (vgl. Stivers et al. 2011) der Sprecherin zu der nun eingenommenen Position verstärkt. Die Iteration der Geste führt somit nicht zuletzt dazu, dass der Konsens ‚glaubhaft‘ abgesichert wird.

Diese Form des nachdrücklichen Zustimmens kann auch mit einer ostentativ-deiktischen Geste erfolgen, wie das nachfolgende Beispiel zeigt. Hier nimmt die Mutter eine oppositionelle Haltung (Z.064-074) zu der Position des Sohnes ein, der als Ziel der Abschlussfahrt ‚Hamburg‘ favorisiert (nicht im Transkript enthalten), woraufhin der Sohn ihr (mit einer Einschränkung) zustimmt (Z. 076-079). Dabei äußert er eine deiktische Geste, die durch ihre lange Vorbereitung auffällt (Z. 077-078).

Beispiel 2 – Nachdrückliches Zustimmung

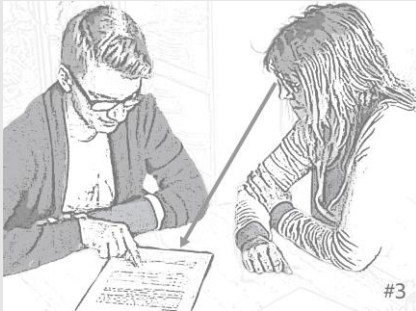
064 MU: meinste NICH,=
 065 =dass so_n SEgeltörn (.) mal ne GANZ tolle erfahrung is auf so_m
 schiff,
 066 um nochmal so die geMEInschaft (-) zuSAMmenzubringen und SPASS zu
 haben,=
 067 =oder °hh dieser BUNgalow|park.
 068 MU: ((hebt rH an & lässt sie auf den Tisch fallen))
 069 MU: aber DAS is dann auch wieder so ne sache-
 da gehen n_paar DA hin,
 071 und n_paar SCHWIMmen,
 072 und n_paar SHOPpen, °h
 073 aber für eure geMEInschaft und für eure LETZtes gemeinsames tolles
 erlEbnis,
 074 fände ich den SEgeltörn schöner.
 075 ?? °hh



076 SO: oKAY.=
so-gz >>blickt auf Blatt (>Z. 081)



077 =<<all>in EIner sa|che sind wir uns einig;
so-gs |.....->



078 und zwar der |sache |VIER,>= |
so-gs|zeigt mit rIF auf Option|
mu-gz >>blickt zu SO|blickt auf Blatt->>

079 =die würde ich dann nämlich (vielleicht) AUCH ausschließen-

080 MU: ((lässt LH auf Tisch fallen))

081 SO: |(-) <<p> |WIE du schon sachttest,>
so-gz |blickt zu MU|

082 SO: fünf tage SEgeltörn==

In Bezug auf die anfängliche Position des Sohnes für den Ausflug nach Hamburg wird anschließend von Seiten der Mutter ein Einwand formuliert (nicht im Transkriptausschnitt enthalten), innerhalb einer rhetorischen Frage (Z. 064-066) der Favorit des Segeltörns benannt und anschließend Gründe gegen den Bungalowpark (Z. 067-074) angeführt. Der Sohn pflichtet der Mutter (einschränkend) in einem satz-übergreifenden Turn bei (Z. 076-079), indem er ebenfalls den Bungalowpark als mögliches Reiseziel für die Abschlussfahrt ausschließt. Bereits zu Beginn seines Turns hebt er den rechten Unterarm vom Tisch ab und streckt den Indexfinger aus (Z. 077, #1). Über zwei Turnkonstruktionseinheiten hinweg wird die gestische Bezugnahme auf den Referenten vorbereitet: Indem der Sprecher den Fokusakzent in Zeile 076 auf „EIner“ legt und kurz darauf die Gestenvorbereitung beginnt, in der der Indexfinger bereits ausgestreckt ist, wird die referenzierende Funktion der Geste projiziert. Desambiguiert wird „EIner“ nun durch die Benennung der Option (Z. 078: VIER) sowie durch die deiktische Geste, der die Mutter mit ihrem Blick folgt. Der Ausschluss der Option wird dann anschließend expliziert (Z. 079). Die lange Vorbereitungsphase der Geste führt zu einer besonderen Hervorhebung des eingeführten Referenten:

Durch das ostentative Zeigen auf den Referenten wird die Zustimmung zum Ausschluss einer gestisch markierten Option besonders nachdrücklich eingebracht. Ähnlich wie in Beispiel 1 trägt hier das ostentative Zeigen dazu bei, in einer für den*die Sprecher*in risikobehafteten Situation, in der er*sie von einer ursprünglichen Position abrückt und einlenkt, den epistemischen Access zu der Positionsbekundung zu beanspruchen und dadurch die nun eingenommene Position ‚glaubhaft‘ zu vertreten.

Zusammenfassend konnte anhand der Analysen gezeigt werden, dass das nachdrückliche Zustimmung in sequenziellen Kontexten auftritt, in denen die Interagierenden zuvor (gegenseinander) Einwände vorgebracht oder die Begründungen und favorisierten Optionen hinterfragt haben. Stellen Interagierende in diesem Kontext nun Konsens her, indem sie (einlenkend) ihrem Gegenüber zustimmen, so erfolgt die Realisierung der Zustimmung mit besonderem Nachdruck: Die Sprecher*innen bringen ausgebaute(re) affirmierende Turnkonstruktionseinheiten ein und markieren verstärkt durch iterative, deiktisch-tippende und ostentativ-deiktische Gesten sowohl ihren eigenen epistemischen Access zu der eingenommenen Position als auch ihre Zustimmung, um im gemeinsamen Ressourcenverbund dem Gegenüber ‚glaubhaft‘ den Konsens zu versichern. Die Geste übernimmt in diesem Ressourcenverbund eine verstärkende Funktion.

8.3.2 Schwaches Zustimmung in Konzessionen

Neben dem nachdrücklichen Zustimmung (vgl. Kap. 8.3.1) realisieren Interagierende im Rahmen von Konzessionen auch abgeschwächte Formen des Zustimmens. In Konzessionen wird einer Behauptung (X) von Sprecher*in B partiell zugestimmt (X') und anschließend gegen sie Einwände vorgebracht oder dieser widersprochen (Y) (vgl. Pomerantz 1984; Couper-Kuhlen und Thompson 2000; Heller 2012; vgl. Kap. 2.2.2). Das schwache Zustimmung innerhalb von Konzessionen wird in den vorliegenden argumentativen Entscheidungsfindungen durch sprachliche sowie körperlich-sichtbare Ressourcen in einer gemeinsamen multimodalen Gestalt realisiert: Dabei werden lexikalische Affirmationen (wie z. B. die Kombination der Partikeln *ja gut*) mit der gestischen Ressource des Aufdrehens der Hand bzw. die Extension einzelner Finger sowie einem vom Gegenüber abgewandten Blick koordiniert. I. d. R. knicken die Interagierenden außerdem ihren Kopf in die entgegengesetzte Richtung des Gegenübers ab, wie die nachfolgende Übersicht 19 zeigt. Auf das schwache Zustimmung erfolgt dann im Rahmen der Konzession ein Einwand oder Widerspruch, wie das nachfolgende Beispiel 1 zeigt.

Schwachtes Zustimmung in Konzessionen	
Beschreibung der multimodalen Gestalt	- schwache sprachliche Zustimmung + Aufdrehen der Hand bzw. Extension einzelner Finger + Blick abgewandt + i. d. R. Abknicken des Kopfes weg vom Gegenüber - läuft sequenziell einem Einwand/Widerspruch voraus

Tabelle 19: Beschreibung der multimodalen Gestalt „Schwachtes Zustimmung in Konzessionen“

In der hier abgebildeten Sequenz erfolgt auf die negative Bewertung des ‚Segeltörns‘, die die Tochter äußert, ein schwaches Zustimmung seitens der Mutter (Z. 196-197), dem im Rahmen einer Konzession dann ein Einwand folgt.

Beispiel 1 – Schwaches Zustimmung in Konzessionen

190 du HAST auch gesagt-
 191 etwas wo man (.) noch nicht WAR,
 192 TO: ich glaub nich-
 193 ja gut==
 194 =aber es is (.) <<p>HOLLand->

195 es is nich (.) |<<pp>sowas> [COO|les;]|
 to-gs |lH & rH Air Quotes-----|
 to-gz |blickt zu MU (>Z. 197)



196 MU: [ja gut |e-] |
 mu-gz >>blickt auf Blatt->>
 mu-gs |dreht rH auf-----|
 mu-he |kippt Kopf zur Seite|



197 |es is |nich |HOLland,=
 mu-gs |Shoulder Shrug re---|
 mu-he |lehnt Kopf zur Seite|
 to-gz ----->blickt auf Blatt->>
 198 =aber:-
 199 |°h |
 mu-gs |zeigt mit rMF auf Blatt|
 to-fa |runzelt Stirn-----|

Überlappend mit der Bewertung der Tochter beginnt eine Konzession seitens der Mutter (Z. 196ff.: ja gut e- / es is nich HOLland,= / =aber:-), in der sowohl die schwach zustimmenden Partikeln *ja gut*, der Adversativjunktorkonkurrenz *aber* sowie Teile der Fortsetzung des *Turns* der Tochter recycelt werden (vgl. Goodwin und Goodwin 1987: 212; vgl. zur (schwachen) Affirmation durch lexikalische Wiederholung Pomerantz 1984; Clark und Bernicot 2008; Arendt und Zadunaisky Ehrlich 2020). Dadurch bleibt insgesamt die konzessive Struktur des Äußerungsaufbaus der Tochter erhalten – an der Äußerungsoberfläche entsteht eine Kohäsion zwischen den *Turns* beider Sprecherinnen.

Innerhalb der von der Mutter eingebrachten Konzession, wird die schwache Zustimmung (X') durch die Kombination der unakzentuierten Partikeln *ja gut* markiert (Z. 196). Die Sprecherin kippt zugleich den Kopf zur Seite – in die entgegengesetzte Richtung der Tochter – und dreht die rechte Hand auf, wobei ein Teil der Finger eingeklappt bleibt (vgl. #1, #2). In Anlehnung an Cienki (2021), der *Palm Up* Gesten als ein Kontinuum begreift, das von der Extension einzelner Finger bis zum Aufdrehen der gesamten Handfläche reicht, wird die Geste der Sprecherin als eine Variante der *Palm Up* Geste verstanden. Das schnelle Aufdrehen der Hand dient hier, in Kombination mit der abgeschwächten Affirmationspartikel *ja (gut)* und der Kopfgeste, einer partiellen Bestätigung der Vorgängeräußerung (vgl. hierzu auch Bavelas et al. 1995). Das seitliche Abknicken des Kopfes hingegen „displays mitigated affiliation“ (Debras

und Cienki 2012: 935)⁸⁵ und kann ferner eingesetzt werden, wenn Perspektivwechsel eingeleitet werden (vgl. Maury-Ruan 2011). Indem die Mutter der Bewertung der Tochter schwach zustimmt – entgegengesetzt ihrer Position, die in dem anschließenden Einwand deutlich wird – zeigt sich hier gewissermaßen ein zumindest kurzzeitiger Perspektivwechsel der Sprecherin. Durch die Bewegungsrichtung des Kopfes in die *entgegengesetzte* Richtung des Gegenübers schafft die Sprecherin zugleich eine räumliche Distanzierung zwischen sich und der Tochter (vgl. zu *head tilts* Debras und Cienki 2012).

Auch in der Fortsetzung der schwach zustimmenden Äußerung (Z. 197: es ist nich HOLLand,) finden sich weitere Ressourcen der abschwächenden Modalisierung wie ein Schulterzucken sowie ein erneutes Abknicken des Kopfes, die zugleich als sogenannte *Shrug*-Komponenten die Funktion übernehmen, eine gewisse Distanzierung gegenüber der Tochter auszudrücken (vgl. Streeck 2009b; Debras und Cienki 2012). Es schließt dann der Einwand (Y) der Konzession, eingeleitet durch den adversativen Junktor *aber* (Z. 198), an.

Zusammenfassend markieren die körperlich-sichtbaren Ressourcen mit den Affirmationspartikeln ein schwaches Zustimmen als Teil einer Konzession. Feste Bestandteile der multimodalen Gestalt des schwachen Zustimmens sind die sprachlich abgeschwächte Affirmation (z. B. mit den Partikeln *ja gut*), die kurzzeitige Extension einzelner Finger oder das Aufdrehen der Hand mit abgewandtem Blick. I. d. R. wird die multimodale Gestalt von einem (langsamen) Absenken des Kopfes entgegengesetzt der Richtung des Gegenübers begleitet.

8.3.3 Schwaches Zustimmen zu evidenten Begründungen & Positionen

Interagierende stimmen einander nicht nur abgeschwächt in Konzessionen zu (vgl. Kap. 8.3.2), sondern sie zeigen abgeschwächte Formen der Zustimmung auch gegenüber redundanten Begründungen oder Positionen, beispielsweise wenn bereits getroffene Entscheidungen durch das Einbringen weiterer Begründungen erweitert oder besprochene bzw. evidente Begründungen wiederaufgenommen werden. Diese werden dann erst durch den Einsatz körperlich-sichtbarer Ressourcen, d. h. durch eine Variante der *Palm Up* Geste (vgl. 8.3.2; vgl. Cienki 2021) sowie *reduzierte* Formen von Kopfgesten (vgl. Whitehead 2011), wie einfaches, z. T. seitliches und auf den*die Rezipient*in ausgerichtetes Nicken sowie zustimmende laterale Kopfbewegungen, abgeschwächt bestätigt, wobei der Blick vom Gegenüber abgewandt ist. Dann erfolgt auch auf sprachlicher Ebene eine abgeschwächte Affirmation (vgl. Tab. 20). Dabei laufen die körperlich-sichtbaren Ressourcen üblicherweise den sprachlichen voraus (vgl. Beispiel 1 & 2) und bilden mit diesen gemeinsam die multimodale Gestalt des schwachen Zustimmens zu evidenten Begründungen und Positionen. An diese können dann z. B. beipflichtende Äußerungen (vgl. Beispiel 1 & 2) oder aber ein Themenwechsel (vgl. Beispiel 3) anschließen, sodass die multimodale Gestalt des schwachen Zustimmens das Abschließen eines Diskussionspunktes anbahnt.

Schwaches Zustimmen zu evidenten Begründungen & Positionen	
Beschreibung der multimodalen Gestalt	• Aufdrehen der Hand bzw. Extension einzelner Finger + Blick abgewandt + seitliche Kopfgeste + schwache Affirmation • läuft sequenziell verbalem Beipflichten und/oder Themenwechseln voraus

Tabelle 20: Beschreibung der multimodalen Gestalt „Schwaches Zustimmen zu evidenten Begründungen & Positionen“

⁸⁵ Debras und Cienki (2012: 935) können keinen Zusammenhang zwischen der Richtung des abgesenkten Kopfes (nach links oder rechts), *head tilt*, und der jeweiligen selbst- oder auf den anderen ausgerichteten Positionierung erkennen. Unbeleuchtet bleibt in der Studie jedoch, neben qualitativen Einsichten in die Daten, ob der Kopf vom Gegenüber abgewandt oder auf ihn hinführend abgesenkt worden ist.

In dem ersten Beispiel erfolgt die multimodale Gestalt des *schwachen Zustimmens zu einer evidenten Begründung*, nachdem beide Interagierenden bereits zu einer gemeinsamen Entscheidung gekommen waren, die Tochter nun aber expandierend eine Konzession⁸⁶ und anschließend eine weitere Begründung einbringt und somit den von der Mutter angebahnten Job des Abschließens verzögert. Die Mutter stimmt hier nun der expandierten Begründung der Tochter schwach zu.

Beispiel 1 – Schwaches Zustimmen zu evidenten Begründungen & Positionen

212 MU: würd ich auch vorschlagen;
 213 das_s [für] ne GROße klassengemeinschaft bestimmt °h (1.7) am
 spannendsten sei[n könnte;]
 214 TO: [ja:;]
 215 [SPRICHT] bestimmt nicht ALle an==
 216 =aber: wahrsch vielleICHT die mEisten;
 217 AUßerdem (-) <<p>will man ja auch etwas zusAmmen machen;>=
 218 ==<<pp>und-> (1.2)



219 to-gz jetzt NICHTS mit shoppen (-) so GROSS|artig;
 ----->|blickt zu MU-->>



220 MU: ((bewegt Kopf nach links, öffnet ineinandergelegte Hände,

⁸⁶ Die Konzession (Z. 215-216) wird nicht als „Cardinal Concessive“ (Couper-Kuhlen & Thompson 2000: 384) mit dem Muster X, X', Y geäußert, sondern die Tochter bringt zuerst einen Einwand (Y) ein, auf den dann die Zustimmung (X') erfolgt. In dieser Reihenfolge und an dieser sequenziellen Position wird nun die *Zustimmung*, und nicht der Einwand, relevant für die Fortsetzung der Interaktion. Die Mutter markiert durch ihre körperlichen Ressourcen eine abgeschwächte Zustimmung, indem sie die Daumen ausstreckt, die Verschränkung der Hände dadurch löst, anhebt und den Kopf *entgegengesetzt* der Tochter absenkt. In der sequenziellen Organisation zeigt sich dann, dass die Tochter ihre Zustimmung (X'), mit der sie die Konzession beendet hat, durch eine *weitere* Begründung stützt und die Mutter anschließend dieser expandierten Begründung schwach zustimmt: Die Kopf-geste zeigt nun keine entgegengesetzte, sondern eine auf die Tochter ausgerichtete Orientierung (vgl. Analyse).



spread lips & bewegt Kopf zurück))

221 TO: un_hol[land] hat-

222 MU: [<p>mjoa;>]

223 (2.8)

224 TO: s_is oKEE-

225 MU: geNAU;

Die schwache Zustimmung der Mutter (Z. 220-221) bezieht sich in dem vorliegenden Transkriptausschnitt auf die expandierende Begründung der Tochter (Z. 217-219). Die schwache Zustimmung wird in einem eigenen *Turn* (Z. 220) durch körperlich-sichtbare Ressourcen angezeigt: Die Mutter öffnet die ineinandergelegten Hände, indem sie den linken Daumen anhebt und die Finger der rechten Hand aus der Verschränkung löst und eine bestätigende *Palm Up* Geste äußert (vgl. Kap. 8.3.2), den Kopf dann zurückführt und zeitgleich die geschlossenen Lippen geschürzt nach vorne schiebt. Diese faciale Geste wird auch als „spread lips“ (Streeck 2009b: 190) oder „lip-pout“ (Givens 1977: 13, zit. n. Debras 2017: 2) beschrieben. Sie kann als eine *Shrug*-Komponente auftreten (vgl. u. a. Morris 1994; Streeck 2009b; Debras 2017) und eine Distanzierung der Mutter gegenüber der Tochter anzeigen; sie übernimmt somit zugleich modalisierend abschwächende Funktion.

Insgesamt ähneln die körperlich-sichtbaren Ressourcen der multimodalen Gestalt des schwachen Zustimmens in Konzessionen (vgl. Kap. 8.3.2), jedoch fungiert die Kopfgeste hier als eine langsame und reduziert ausgeführte Form des Kopfschüttelns. Mit dieser stimmt die Mutter der Negation (Z. 219) zu, mit der die Tochter hervorhebt, dass Freizeitaktivitäten wie etwa das Shoppen keine angemessene Aktivität für eine Abschlussfahrt sind (vgl. u. a. Kendon 2002). Die geschürzten Lippen, als *Shrug*-Komponente der Markierung von Distanzierung (vgl. Streeck 2009b), schwächen die Zustimmung, die durch das Öffnen der Hände zu einer *Palm Up* Geste markiert und durch das affirmierende Kopfschütteln verstärkt wird, ab. Die körperlich-sichtbaren Ressourcen führen an dieser sequenziellen Position, an der die Tochter ihre Begründung abgeschlossen hat, zu einer Verzögerung der verbalsprachlichen Rückmeldung – die Ressourcen dienen hier also auch einer Dispräferenzmarkierung (vgl. auch Pomerantz 1984). Die Mutter schließt dann in verminderter Lautstärke eine abgeschwächte Affirmationspartikel (Z. 222: <p>mjoa;>) an, auf die eine Pause und dann eine bewertende Äußerung der Tochter folgt (Z. 224), der die Mutter nun mit der bestätigenden Partikel *genau* (Z. 225) beipflichtet (vgl. Oloff 2017). Die Expansion einer Begründung, die an eine bereits gemeinsam getroffene Entscheidung angeschlossen wird, bedarf seitens der Mutter also nur noch einer abgeschwächten Zustimmung, bevor nun endgültig der Kernjob des Begründens von den Interagierenden global abgeschlossen wird.

Doch nicht nur das Verhindern des Abschlüssens der Kernjobs kann durch eine (verzögerte) schwache Zustimmung quittiert werden, auch das Benennen evidenter, bekannter Begründungsinhalte kann als nicht-präferiert markiert werden. Der nachfolgende Transkriptausschnitt zeigt, wie die körperlich-sichtbaren Ressourcen des Vaters (Hochziehen der Augenbraue, *Palm Up* Geste, langsames Absenken des Kopfes) eine schwache Zustimmung anzeigen, die anschließend auch sprachlich realisiert wird.

Beispiel 2 – Schwaches Zustimmen zu evidenten Begründungen & Positionen

093 VA: [°hh]
 094 aber ihr lernt zuSAMmenzuarbeiten;
 095 zuSAMmen zu sEgeln;
 096 die aufgaben alle zuSAMmen zu erledigen;
 ((...))
 148 SO: und der segeltörn das is natürlich: (---) joa;
 149 VA: da muss man sich selbst verPFLegen,
 va-gz >>blickt zu SO->>
 so-gz >>blickt auf Tischplatte (>Z. 153)
 150 muss man vielleicht auch mal selbst KÖchen; (.)
 151 selbst SAÜber machen.
 152 SO: das wär dann halt für_n zuSAMmenhalt gut. |
 so-gz -----> |
 153 | (0.4)
 so-gz |blickt zu VA->>



154 VA: ((zieht Augenbraue hoch, senkt Kopf Ri SO ab, dreht bHde Palm Up auf))
 155 ((schnalzt))
 156 VA: ZUM beispiel.
 157 (1.0)
 158 VA: (und) DAS is nur abhängen.

Der abgebildeten Sequenz geht eine Ablehnung des Sohnes gegenüber der Option des Segeltörns voraus, die sich auch in Zeile 148 zeigt, in der ein erwartbares evaluatives Element (Z. 148: und der segeltörn das is natürlich:) ausbleibt und stattdessen durch die Partikel *joa* ersetzt wird. Eine Entscheidung für diese Option ist möglich, aber nicht unbedingt präferiert. Der Vater hingegen bringt eine Liste von Argumenten für den Segeltörn ein (Z. 149-151), die Ähnlichkeiten zu einer zuvor eingebrachten Aufzählung (Z. 094-096) aufweist: Beide Begründungen heben auf das selbstverantwortliche, gemeinsame Handeln der Schüler*innen ab. Als nun der Sohn an die zweite Liste anschließend resümiert, „das wär dann halt für_n zuSAMmenhalt gut.“ (Z. 152) und damit den Kern der Begründungen des Vaters für den Segeltörn zusammenfasst, bleibt zuerst eine Rückmeldung des Vaters aus, obwohl eine Turnübernahme sowohl durch die abgeschlossene Turnkonstruktionseinheit der Vorgängeräußerung sowie dadurch, dass der Sohn ihm den Blick zuwendet (Z. 152), wahrscheinlich ist. Anschließend stimmt der Vater dieser Konklusion nur schwach zu und markiert sie als evident bzw. redundant: Er zieht die Augenbraue(n) hoch, senkt den Kopf in Form eines langsamen Nickens (vgl. Whitehead 2011) in Richtung des Sohnes ab und dreht beide Hände bestätigend zu einer *Palm Up* Geste mit lateraler Ausrichtung auf. Nicht nur durch den sequenziellen Kontext, sondern auch durch die eingesetzten Ressourcen wie beispielsweise die *Palm Up* Geste mit lateraler Ausrichtung sowie das Hochziehen der Augenbrauen wird der Inhalt der Äußerung des Sohnes als offensichtlich markiert (vgl. u. a. Debras 2017). Dabei verzögern sowohl die vorangegangene Pause als auch die körperlich-sichtbaren Ressourcen selbst die sprachliche Rückmeldung. Sie dienen gemeinsam als Dispräferenzmarkierung (vgl. Pomerantz 1984). Es folgt anschließend, nach einem Schnalzen (Z. 155), durch die Phrase „ZUM beispiel“ (Z. 156) auch sprachlich

nur eine abgeschwächte Form der Zustimmung, die nicht weiter ausgebaut wird. Eine weitere Bearbeitung der Konklusion des Sohnes bleibt aus; vielmehr wendet sich der Vater nun einer weiteren Option (dem Bungalowpark) zu (Z. 158) und leitet damit einen thematischen Wechsel ein.

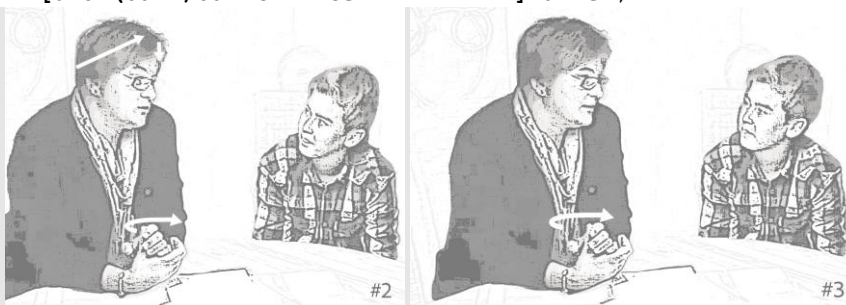
Dass die körperlich-sichtbaren Ressourcen den sprachlichen nicht ausschließlich vorauslaufen (vgl. Beispiel 1 & 2), sondern auch mit diesen in ihrer zeitlichen Koordination zusammenfallen können, illustriert das nachfolgende und letzte Beispiel.

Beispiel 3 – Schwaches zustimmen zu evidenten Begründungen & Positionen

157 MU: aber dann macht_a nich als KLASse was zusammen;
 158 SO: ja KÖNN_n wir ja;
 159 MU: WAS denn;
 160 SO: wir KÖNNten ja zum beispiel alle gemein[sam-]
 161 MU: [IHR geht] alle shoppen;
 162 SO: SCHWIMmen gehen oder so;
 163 MU: gut--man hätte (ne gemeinschafts);
 164 °h ja [und BOWling,]
 165 SO: [aber man] !KÖNN!te ja au:ch,
 166 °h (--)
 167 mit dem (.) !SE!gelschiff (.) zum zu irgendeiner freizeitmöglichkeit
 fahren-
 168 und (.) dann (.) da: einen LANDaufenthalt machen,



169 und vom landaufenthalt dann wieder auf (.) SCHIFF gehen;
 170 [und (dann/damit WEiter]fahren;



171 MU: [|KÖNNte |man| auch machen;]
 mu-fa |zieht AB hoch--|
 mu-gz |schließt Augen-|blickt zu SO (>Z. 176)
 mu-he |senkt Kopf Ri SO---|
 mu-gs |öffnet rH-----|
 172 SO: also könnte man das ja eigentlich kombiNIeren;
 173 [(--) so;]
 174 MU: [|((richtet Hände parallel übereinander aus: rH Palm Up, lH Palm Down))]
 175 ((lässt Hände ineinander fallen))
 176 MU: DAS |könnte man auch machen;=
 mu-gz --->|blickt auf Blatt->>
 177 MU: =das HEISST also (.) CENTer parcs find ich jetzt nich so gut-

Der Beginn der Transkripts zeigt noch die Ausläufer einer Aushandlung zwischen den Interagierenden, in der der Sohn seine Position für den Bungalowpark vertritt und die Mutter gegen diese einwendet und zu einzelnen Behauptungen Ergänzungsfragen stellt (z. B. Z. 157-159). Obwohl die Mutter in Zeile 163 einzulenken beginnt, initiiert der Sohn nun einen neuen Vorschlag (Z. 165-170) zu einer möglichen Kombination zweier Optionen – eine davon ist der von ihm favorisierte Aufenthalt in einem Bungalowpark. Überlappend mit dem Ende dieser Expansion stimmt die Mutter dieser schwach zu (Z. 171): Zeitgleich mit einer partiellen Wiederholung des *Turns* des Sohnes (Z. 165: aber man !KÖNN!te ja auch, / Z. 171: KÖNNte man auch machen;), mit der die Mutter den Vorschlag des Sohnes bestätigt, aber keine weiteren lexikalischen Affirmationen einführt, um ein ausgebautes Einverständnis mit diesem Vorschlag zu markieren, löst sie die rechte Hand kurzzeitig aus der Verschränkung, öffnet somit die Hand leicht zu einer *Palm Up* Geste (vgl. Cienki 2021) und senkt den Kopf in Richtung des Sohnes ab (vgl. Whitehead 2011). Den Konjunktiv hervorhebend zieht die Sprecherin aligniert mit dem Fokusakzent „KÖNNte“ (Z. 171) beide Augenbrauen intensivierend nach oben (vgl. Ekman 1979). Dass es sich um eine schwache Form der Zustimmung handelt, wird nun in der Fortsetzung der Interaktion sichtbar: Der Sohn expandiert seinen *Turn* erneut (Z. 172f.) und es erfolgt eine wiederholte schwache Zustimmung der Mutter (Z. 176), bevor die Sprecherin dann zu einer eigenen Position überleitet (Z. 177). Stevanovic (2012: 790) bezeichnet diese Form der Affirmation auch als „agreement“ (Zustimmen), das allerdings noch kein „commitment“ (keine Zusage) zu einer gemeinsamen Entscheidung beinhalten muss. Der Anschluss einer eigenen, aber gegenläufigen Position (Z. 177) wird dadurch möglich.

Zusammenfassend konnte gezeigt werden, dass das schwache Zustimmung in sequenziellen Kontexten hervorgebracht wird, in denen ein (lokales oder globales) Abschließen von den Interagierenden bereits angebahnt, aber durch eine expandierende Handlung des Gegenübers unterbrochen worden ist, oder aber in sequenziellen Kontexten realisiert wird, in denen evidente Begründungen eingebracht worden sind. Durch die multimodale Gestalt des schwachen Zustimmens werden die Handlungen dann als redundant erkennbar angezeigt – und eine anschließende Fortsetzung im Anschluss an die schwache Zustimmung durch lokales Abschließen oder Themenwechsel umgangen. In der temporalen Entfaltung der multimodalen Gestalt können sprachliche und körperlich-sichtbare Ressourcen nun gemeinsam realisiert werden (vgl. Beispiel 3) oder die körperlich-sichtbaren Ressourcen den sprachlichen vorauslaufen und dabei als Dispräferenzmarkierungen eine sprachliche Rückmeldung verzögern (vgl. Beispiel 1 und 2; vgl. Pomerantz 1984). Die körperlich-sichtbaren Ressourcen enthalten dabei bestätigende Komponenten, die aber aufgrund ihrer Realisierung, d. h. durch das *kleinräumige* und *kurzzeitige* Aufdrehen und/oder Öffnen der Hand bzw. Hände und Extension einzelner Finger zu einer *Palm Up* Geste sowie durch die *reduzierte* Geste des Nickens bzw. Kopfschüttelns, *abgeschwächte* Affirmation markieren. Die Untersuchung legt darüber hinaus dar, dass der Ressourcenverbund der Hand- und Kopfgeste sowie der abgeschwächten sprachlichen Affirmation durch Ressourcen erweitert werden kann, die zusätzliche Funktionen übernehmen, wie die geschürzten Lippen (vgl. Beispiel 1) als weiteres *Display* der Distanzierung, die laterale Bewegung der *Palm Up* Geste zur Markierung von Evidenz (vgl. Beispiel 2) oder auch das Hochziehen der Augenbrauen (vgl. Beispiel 3), das in Alignierung mit dem Fokusakzent ein wiederholtes Lexem im Konjunktiv hervorhebt.

8.3.4 Zusammenfassung

In mündlichen Argumentationen können zustimmende Handlungen nachdrücklich oder abgeschwächt hervorgebracht werden. Die Untersuchung zeigt, dass die zeitliche und intrapersonelle Koordinierung der multimodalen Ressourcen sowie ihre sequenzielle Einbettung zu ihrer jeweiligen pragmatischen Funktion beitragen.

- So nutzen Interagierende in persuasiv bzw. antagonistisch gerahmten Anteilen der Interaktion das *nachdrückliche Zustimmung*, um in ausgebauten, z. T. satzübergreifenden Äußerungen den Bezugspunkt ihrer Zustimmung durch deiktische Gesten sichtbar zu machen. In dem Wechsel von der opponierenden zur beipflichtenden Rolle des*der Sprecher*in werden die iterativen deiktischen Gesten eingesetzt, um den sprecher*innenseitigen epistemischen *Access* zu der nun eingenommenen zustimmenden Position demonstrativ zu verstärken und somit die Position ‚glaubhaft‘ zu vertreten (vgl. Kap. 8.3.1).
- Dass Interagierende ihre Zustimmung auch abmildern und somit eine (vorerst) nicht uneingeschränkte Zustimmung markieren, konnte sowohl innerhalb des Einlenkens in Konzessionen (vgl. Kap. 8.3.2) als auch im Kontext redundanter Begründungen und Positionen (vgl. Kap. 8.3.3) gezeigt werden. Trotz der äußerlichen Ähnlichkeit der Koordinierung der körperlich-sichtbaren Ressourcen, wie dem Aufdrehen oder der Extension der Finger sowie der Kopfgeste, zeigen sich deutliche Unterschiede in ihrer temporalen sowie sequenziellen Entfaltung in der Interaktion: In der multimodalen Gestalt des schwachen Zustimmens in Konzessionen werden die sprachlichen Affirmationen und die körperlich-sichtbaren Ressourcen zeitgleich geäußert, während der Kopf in die vom Gegenüber entgegengesetzte Richtung als Markierung eines Perspektivwechsels sowie der Distanzierung abgesenkt wird. In der multimodalen Gestalt des schwachen Zustimmens zu evidenten Begründungen und Positionen können die körperlich-sichtbaren Ressourcen hingegen auch den sprachlich abgeschwächten Affirmationen vorauslaufen und somit zu deren zeitlicher Verzögerung beitragen. Hier erfolgt die Kopfgeste, die als reduzierte Form des Nickens oder Kopfschüttelns geäußert wird, in Richtung des Gegenübers und trägt somit gemeinsam mit der Handgeste ebenfalls eine abgeschwächt bestätigende Funktion. Gemein ist beiden multimodalen Gestalten des schwachen Zustimmens, dass sie eine *Palm Up* Geste beinhalten, deren Funktion sich im gesamten Ressourcenverbund und unter Berücksichtigung des sequenziellen Kontextes als schwach bestätigend erwies. Die Untersuchung zeigte ferner eine Varianz in der zeitlichen Entfaltung der Ressourcen der multimodalen Gestalt des schwachen Zustimmens zu evidenten Begründungen und Positionen: Die Hand- und Kopfgeste, die immer zeitgleich miteinander koordiniert geäußert werden, laufen dabei i. d. R. der verbalen Ressource voraus. Dabei bleibt die multimodale Gestalt stets in sich beständig, kann aber in ihrer zeitlichen Entfaltung der Ressourcen variieren (vgl. Beispiel 1 & 2 gegenüber Beispiel 3 in Kap. 8.3.3). Insgesamt zeigt sich in den letztgenannten multimodalen Gestalten des schwachen Zustimmens eine Präferenz zur *reduzierten* Produktion der multimodalen Ressourcen, wohingegen beim nachdrücklichen Zustimmung die körperlich-sichtbaren Ressourcen durch iterative Bewegungen oder die lange Vorbereitungsphase demonstrativ geäußert werden.

8.4 Einwenden

In mündlichen Argumentationen zeigen sich Interagierende durch Einwände gegenseitig an, dass eine vorangegangene Äußerung problematisiert wird (vgl. Kap. 2.2.2). Dabei können die Problemstellen, oder auch „trouble source[s]“ wie Goodwin und Goodwin (1987: 207) sie nennen, in den Einwänden implizit oder explizit sichtbar gemacht werden und so z. B. Plausibilisierungen oder Elaborationen vom Gegenüber evozieren. Dabei können die Einwände bzw. einzelne Segmente durch *Palm Up* Gesten für die Beteiligten präsentiert als auch ihre jeweilige Bezugnahme auf beteiligte Personen oder das Material angezeigt werden (vgl. Kap. 8.4.1). Die Charakteristik eines Einwandes legt nahe, dass durch die Opposition Teile oder auch gesamte Vorgängeräußerungen negiert werden. Die Studie zeigt, dass gestische

Ressourcen zu einer sichtbaren Explizierung dieser Negationen beitragen können (vgl. Kap. 8.4.2). Einwände können auch in größere Strukturen wie Konzessionen (vgl. Kap. 2.2) eingebettet sein. Schließt die Konzession mit einer schwachen Zustimmung ab, kann das Anhalten des Einwandes durch *Holding Away* Gesten angezeigt werden (vgl. Kap. 8.4.3). Gesten, wie beispielsweise der erhobene Zeigefinger, können in Einwänden auch zu einer besonderen Aufmerksamkeitsfokussierung beitragen, indem sie nachfolgende Inhalte als relevant ankündigen und im Kontext des Einwendens den Inhalt einer Vorgängeräußerung anmahnen (vgl. Kap. 8.4.4). Diese multimodalen Gestalten werden im Folgenden vorgestellt.

8.4.1 Präsentieren während des Einwendens

Innerhalb des Einwendens können Interagierende entweder gesamte Einwände oder größere Segmente dieser im Verbund mit der gestischen Ressource der *Palm Up* Geste als abstrakte diskursive Objekte präsentieren (vgl. Müller 2004; Schönfelder und Heller 2019) und dabei, je nach Orientierung der Geste, andere Interagierende oder weitere Objekte in den Einwand einbeziehen. Durch die zeitliche Alignierung mit der Turnkonstruktionseinheit markiert die Geste ferner den Einwand bzw. ein Segment des Einwandes als zusammengehörig und übernimmt neben der performativen präsentierenden auch eine diskursstrukturierende Funktion, wie der tabellarischen Übersicht 21 zu entnehmen ist.

Präsentieren während des Einwendens	
Beschreibung der multimodalen Gestalt	• Einwand + <i>Palm Up</i> Geste (in unterschiedlicher Ausrichtung der Geste) • erfolgt sequenziell responsiv

Tabelle 21: Beschreibung der multimodalen Gestalt „Präsentieren während des Einwendens“

Die multimodale Gestalt des Präsentierens während des Einwendens ist geprägt durch eine mit dem Einwand zeitlich alignierte *Palm Up* Geste (vgl. Beispiel 1-3). Durch die Orientierung der Geste können Interagierende unterschiedliche, aber für den jeweiligen Einwand relevante Bezugnahmen herstellen: So können interpersonell ausgerichtete *Palm Up* Gesten (vgl. *Palm Addressed*, Kendon 2004) das Gegenüber als Auslöser*in bzw. Empfänger*in des Einwandes adressieren (vgl. Beispiel 2) oder aber *Palm Up* Gesten ein Zeigeziel explizit für den Einwand relevant machen (vgl. Beispiel 3).

In dem ersten Transkriptauszug führt die Mutter einen Einwand ein und präsentiert dabei das erste Segment des Einwandes auf der geöffneten Handfläche.

Beispiel 1 – Präsentieren während des Einwendens

050 SO: [MU]sical?
 051 also WIRKlich;
 052 das is das ist doch so Opernmusik und so(was/weiter);
 053 das ist doch SCHEI[Sse.]



054 MU: |[ab]er |VIER tage hAm |burg;|
 mu-gs |.....|öffnet lH Palm Up|,,,,,|
 mu-gz |blickt zu SO----->>

055 mit [(.)] STADT[rundfahrt-]
 056 SO: [ja oKAY-]
 057 [JA aber-]
 058 MU: HAFen-
 059 SO: ja Ach KOMM-

Oberflächensprachlich wird der Einwand durch den Adversativjunktorkonkordanz *aber* eingeführt (Z. 054), den die Sprecherin überlappend mit der abschließenden Evaluation des Sohnes (Z. 053: das ist doch SCHEISse.) äußert. Aufgrund der sequenziellen Position des *Turns* sowie der lexikalischen Markierung der Opposition wird ersichtlich, dass die Sprecherin auf den vorangegangenen *Turn* Bezug nimmt. Dabei entfaltet sich gemeinsam mit der Turnkonstruktionseinheit eine *Palm Up* Geste (Z. 054: aber VIER tage hAm-burg; vgl. #1): Temporal aligniert mit dem Fokusakzent der Intonationsphrase wird die Hand zu einer *Palm Up* Geste geöffnet und der Blick dem Sohn zugewandt. Kurz vor dem Beenden der Turnkonstruktionseinheit, das prosodisch durch die fallende Tonhöhenbewegung markiert wird, wird nun auch die Geste in eine Ruheposition zurückgeführt. Insgesamt wird in diesem Beispiel also das erste, syntaktisch unvollständige bzw. elliptische Segment des Einwandes präsentiert, in dem die Sprecherin die Option opponierend zur Ablehnung des Sohnes benennt und durch die Fokusakzentuierung der Anzahl der Tage deren mögliche Relevanz markiert. Der Geste kommt innerhalb des Ressourcenverbundes damit nicht nur eine performative Funktion zu, indem sie den elliptischen Einwand einleitend präsentiert. In ihrer Entfaltung, die innerhalb einer Gesteneinheit mit Beginn der Turnkonstruktionseinheit die Ruheposition verlässt und in diese am Ende zurückkehrt, zeigt sie ferner die Turnkonstruktionseinheit als zusammengehörige Einheit (vgl. Kendon 2004: 112) an und übernimmt somit auch eine äußerungsstrukturierende Funktion.

In dem zweiten Beispiel erbringt der Sohn einen Einwand gegen eine Belegerzählung (vgl. Heller 2012: 86) des Vaters, mit der dieser eine persönliche Erfahrung als Begründung für die Auswahl des Segeltörns anführt. Der Einwand des Sohnes wird dabei auf der geöffneten Handfläche präsentiert, die zugleich durch ihre Ausrichtung den Vater zum Adressaten des Einwandes sichtbar markiert (vgl. Kendon 2004: 271ff.).

Beispiel 2 – Präsentieren während des Einwendens

261 VA: =ICH hab so_n: (.) so_n segeltörn ja mal mit meinem lEistungskurs
 gemacht;
 262 <<dim>auf_m iJSSselmeer;
 263 °hh und das fand ich UNglaublich gut;>
 264 [man] man [hängt GANZ ENG bei_nander,]
 265 SO: | [°hh]
 so-gz |blickt i. d. Raum->



266 SO: [JA: ;=|LEIS|tungskurs;]
 so-gs |zeigt mit lH Palm Up auf VA-->
 so-gz ----->|blickt zu VA (>Z. 272)
 267 SO: lEIS|tungskurs| [kurs;]
 so-gs ----->|
 268 VA: [JA-=]

269		=die sind [<<nickend>die sind (ne stufe/natürlich)] Älter> gewesen;=
270	SO:	[das_s Oberstufe;
271	VA:	=KLAR;

Das vorliegende Beispiel zeigt, dass die Rekonstruktion einer argumentativen Handlung nicht allein aufgrund der verbalsprachlichen Ressourcen erfolgen kann: Nachdem der Vater in seiner Belegerzählung seine persönliche Segelerfahrung mit einem Leistungskurs (Z. 261f.) in einer *Extreme Case Formulation* (vgl. Pomerantz 1986) als „UNglaublich gut“ (Z. 263) bewertet hat und das Äußerungspaket mit einer Begründung (Z. 264) fortzusetzen beginnt, setzt der Sohn überlappend mit einer betonten sowie gedehnten Affirmationspartikel ein (Z. 266: JA:;) und wiederholt das vom Vater verwendete Lexem „LEISTungskurs“ (Z. 266, 267) zweimal mit einer prosodischen Markierung durch einen Fokusakzent sowie mit fallender Tonhöhenbewegung. Die Lexemwiederaufnahme bzw. -wiederholung schafft, i. S. des *Format Tyings* (vgl. Goodwin und Goodwin 1987; Goodwin 1990), somit eine Verbindung zwischen beiden Äußerungen. Sequenziell schließt der Einwand direkt an die Äußerung des Vaters an – eine mögliche Turnübernahme wird bereits durch das wahrnehmbare Einatmen (Z. 265) angedeutet. Doch erst durch die kontrastierende Fokusakzentuierung wird das Lexem „LEISTungskurs“ (Z. 266) problematisiert. Denn in der Belegerzählung nimmt der Vater eben nicht Bezug auf eine Fahrt, die er mit einem *Unter-* bzw. *Mittelstufenkurs*, sondern mit einem *Leistungskurs* und somit einem Kurs der *Oberstufe* (vgl. Z. 270) unternommen hat – in der vorliegenden Entscheidungsfindung geht es aber um eine Abschlussfahrt am Ende der Mittelstufe, die der Sohn besucht.

Zeitlich aligniert mit dem problematisierten und als „trouble source“ (Goodwin 1983: 659) identifizierbaren Lexem „LEISTungskurs“ (Z. 266) zeigt der Sohn mit der geöffneten Handfläche auf den Vater (vgl. #1) und adressiert diesen mit einer *Palm Up* Geste als Empfänger des Einwandes, „as if to indicate that one expects or hopes to receive something specific from the other to whom it is directed, such as some information, an explanation or a justification for something“ (Kendon 2004: 272). Die Ausrichtung der Geste macht also nicht nur sichtbar, wer der Adressat der Äußerung ist, sondern durch die gestische Ressource wird das Lexem *Leistungskurs*, das Teil der Begründung des Vaters war (Z. 261), einwendend an das Elternteil *zurückgegeben*. Dabei besteht der Einwand oberflächensprachlich nur aus einem einzigen Lexem. Der Sohn reduziert hier also in seiner Opposition seinen Redebeitrag so weit, dass er in der Fortsetzung der Interaktion nur bedingt selbst opponierend durch das Elternteil aufgegriffen werden kann – die multimodale Gestalt ermöglicht hier also eine Reduktion der verbalsprachlichen ‚Angriffsfläche‘. Es schließt dann eine Zustimmung des Vaters gegenüber dem Einwand an, den er, in Überlappung mit dem Sohn (Z. 268-269), elaboriert.

Erst die umfassende Beschreibung der sequenziellen Einbettung der Äußerung des Sohnes sowie der prosodischen und gestischen Ressourcen lässt erkennen, dass vermittels gestischer Ressourcen nicht nur ein Einwand als präsentiert kontextualisiert wird, sondern auch an das Gegenüber zurückgegeben werden kann. Die Geste erfüllt hier also im Rahmen der gesamten multimodalen Gestalt performative Funktionen.

Innerhalb des Präsentierens von Einwänden können nicht nur Bezugnahmen auf das Gegenüber, sondern auch auf das Material angezeigt werden. Im folgenden Beispiel bezieht der Sohn in seiner Begründung gegen den Segeltörn eine Erfahrung mit seinen Mitschüler*innen ein, von denen ein Großteil auf der Fährfahrt nach England seekrank geworden sei. Nach einer ersten Zustimmung kontrastiert der Vater nun die Gewässer (*Ärmelkanal* vs. *Ijsselmeer*) und präsentiert den ersten Teil des Einwandes auf der geöffneten Handfläche.

Beispiel 3 – Präsentieren während des Einwendens

434 SO: ja: der NACHteil is halt wirklich-
 435 dass ähm (.) das für seekranke GAR keinen spaß macht,
 436 VA: hm_hm,
 437 klar.
 438 SO: weil ähm bei mir MINdestens nen DRITttel der klasse,
 439 von der fahrt nach caLAIS von caLAIS nach dOver,
 440 VA: ((räuspert sich))
 441 SO: wieder [zurück] jeweils (.) ähm (.) schOn- (-)
 442 VA: [hm_hm,]
 443 SO: für ne hAl für ne (.) für ANderthalb stunden waren_se schon (.)
 seekrank,
 444 und wenn sie dann FÜNf tage auf_dem meer sind,
 445 mit [SCHLafen und allem möglichen,]
 446 VA: [dann IS das: dann is das unter] umständen ziemlich ätzend;
 447 das SEH ich ein.
 448 SO: tja:;



449 VA: Aber |das is |natürlich auch nicht |ÄR |melkanal;
 va-gs |.....|führt lH Palm Up auf Blatt-|Beat-|verschränkt Arme
 (>Z. 460)
 450 |(-) das is ja IJsselmeer.
 va-gz |blickt zu SO ----->
 451 und ijssel|meer is ja mehr so: (.) [BINnenmeer] binnen-
 va-gz ----->|
 452 SO: [MEER is meer;]
 453 VA: |nee nee nee nee nee:..|
 va-fa |zieht AB hoch-----|

Der Vater bringt zu Beginn des Einwandes (Z. 449-451) eine opponierende Proposition ein (Z. 449: Aber das is natürlich auch nicht ÄRmelkanal;), deren argumentative Funktion erst ab Zeile 451 erkennbar wird. Nach dem Adversativjunktore *aber* setzt die Vorbereitung der Gestenausführung ein, bei der er mit der nach oben geöffneten Handfläche auf das Material zeigt und mit dem Fokusakzent die Hand in einem *Beat* auf das Blatt absenkt (vgl. #1). In der einwendenden und durch das Satzadverb *natürlich* (Z. 449) epistemisch verstärkt modalisierten Proposition werden zwei Gewässer kontrastierend einander gegenübergestellt: Durch das unakzentuierte, turninitiale Demonstrativpronomen *das* (Z. 449)⁸⁷ nimmt der Sprecher einerseits, wie retrospektiv deutlich wird, Bezug auf das *Ijsselmeer* (Z. 450), während er sich andererseits am Ende der Turnkonstruktionseinheit auf das explizit-sprachlich benannte Gewässer des Ärmelkanals bezieht. Die *Palm Up* Geste übernimmt für den Einwand nun zwei Funktionen: Zum einen kontextualisiert sie die einwendende Proposition als präsentiert. Zum anderen schafft die Geste eine für den Einwand relevante Kontrastierung, indem sie während der Benennung des „Ärmelkanal[s]“ (Z. 449) auf das Material und somit auf das Gewässer des Ijsselmeers verweist.

⁸⁷ Unakzentuierte Demonstrativpronomen, an die Zeigehandlungen anschließen, werden nicht in der Funktion der Lokalisierung oder Identifizierung verwendet, wie Stukenbrock (2015: Kap. 5.1.2) für Indexfingergesten mit vertikal orientierter Handfläche feststellt.

Zusammenfassend zeigt sich, dass sich die *Palm Up* Geste zeitlich koordiniert gemeinsam mit der argumentativen Handlung des Einwendens entfaltet. Sie setzt um den Beginn der Turnkonstruktionseinheit ein und wird mit deren Ende zurückgezogen, sodass sie den Einwand bzw. Segmente des Einwandes als präsentiert kontextualisiert. Durch die zeitliche Alignierung von Turnkonstruktionseinheit und Geste wird darüber hinaus das Segment als zusammengehörig markiert (vgl. Kendon 2004: 112), sodass der Geste neben einer performativen ebenfalls eine äusserungsstrukturierende Funktion zukommt (vgl. u. a. Beispiel 1). Die Ausrichtung der Geste kann dabei interpersonelle Bezugnahmen (vgl. Beispiel 2) oder Referenzierungen auf Objekte (vgl. Beispiel 3) anzeigen. Dabei deutet sich an, dass in satzförmigen Oppositionen Gesten in ihren Ressourcenverbünden relevante Hinweise liefern, wie z. B. die Rückgabe eines Begründungselementes (Beispiel 2), ohne diese explizieren zu müssen. Möglicherweise kann so die *verbalsprachliche* ‚Angriffsfläche‘, auf die das Gegenüber in der Fortsetzung der Interaktion Bezug nehmen kann, reduziert werden (vgl. dazu auch Goodwin und Goodwin 1987).

8.4.2 Sichtbares Explizieren von Negationen in Einwänden

In einwendenden Äußerungen stellen u. a. Negationen ein wichtiges Mittel dar, um eine Opposition zur Vorgängeräußerung herzustellen (vgl. u. a. Spranz-Fogasy 2005b). Negationen können z. B. durch Negationsmorpheme wie *nicht*, *kein/e/r* oder durch körperlich-sichtbare Ressourcen wie das Kopfschütteln sowie auch Gesten ausgedrückt werden (vgl. u. a. Kendon 2002; Li 2021). Verbleiben die Negationen auf sprachlicher Ebene implizit, können die körperlich-sichtbaren Ressourcen zu ihrer Explizierung beitragen oder, wenn die Negation explizit-sprachlich markiert wird, diese modalisierend verstärken (vgl. u. a. Kendon 2004; Streeck 2009b; Harrison 2010; Bressemer und Müller 2014a; Harrison und Larrivé 2016; Harrison 2018; vgl. Kap. 2.2.2 & Kap. 3.2.2.1).

In den vorliegenden Daten werden gestische Ressourcen sowohl zur Verstärkung einer explizit-sprachlichen Negation eingesetzt (vgl. Beispiel 1) sowie auch, um implizite Negationen, die zur Opposition beitragen, sichtbar zu machen (vgl. Beispiel 2). Dazu nutzen Sprecher*innen zum einen materialgebundene Gesten, bei denen z. B. ein Stift horizontal durch die Luft geführt wird (vgl. Beispiel 1), sowie zum anderen *Horizontal Palm* (vgl. Kendon 2004: 255ff.) bzw. *Palm Down Horizontal Across* Gesten (vgl. Harrison und Larrivé 2016), bei denen die Hand lateral vom Körper des*der Sprecher*in weg bewegt wird (vgl. Beispiel 2). Durch die Orientierung der Handfläche, die nach unten ausgerichtet ist, sowie die Bewegungsrichtung der Geste lateral nach außen können Interagierende anzeigen, „that something in progress is interrupted, that some process is cut off“ (Kendon 2004: 263). Es ist zu beobachten, dass die multimodale Gestalt entweder am Ende eines *Turns* eingesetzt wird oder aber der*die Sprecher*in innerhalb eines größeren Äußerungspaketes ein Argument des*der vorangegangenen Sprecher*in aus dem weiteren Diskurs ausschließt und dann einen Themenwechsel einleitet (vgl. Tab. 22).

Sichtbares Explizieren von Negationen in Einwänden	
Beschreibung der multimodalen Gestalt	Geste mit horizontaler Bewegungsrichtung (z. B. <i>Palm Horizontal/Palm Down Horizontal Across</i> oder horizontale Bewegung eines Stiftes) (+ optional sprachliche Negation)

Tabelle 22: Beschreibung der multimodalen Gestalt „Sichtbares Explizieren von Negationen in Einwänden“

In dem nachfolgenden Beispiel erfolgt der Einwand des Vaters im Anschluss auf den Vorschlag des Sohnes, getrennte Bungalows für Jungen und Mädchen in dem Bungalowpark zu organisieren. In einer eingeschobenen Einschätzung opponiert der Vater und markiert darin sowohl sprachlich als auch gestisch Negationen.

Beispiel 1 – Sichtbares Explizieren von Negationen in Einwänden

539 SO: [O:der,]
 540 dass die JUNgen und die MÄDchen-
 541 e die JUNgen, (-)
 542 ham zwei BUNgalows; (-)
 543 auf der EInen seite des parks;
 544 mit dem LEHrer,
 545 (0.9)
 546 und die mädchen haben zwei bungalows auf der ANdern seite des parks,
 547 [mit der lehreRIN,]
 548 VA: [es geht nich DARum-=]
 549 =dass die dass die mÄdchen und die JUNgen irgendwie mitei[nander was
 anfangen;]
 550 SO: [aber wenn
 die MÄD]



551 VA: [|DA hab ich keine sorgen | bei euch;]
 va-gs |bew. Stift mehrm. horizontal hin & her|
 552 SO: [chen wenn die mädchen al |LEine sind;
 553 dann ähm-|
 554 VA: MÄchen die nix.
 555 SO: n dann TRINKen die aber nich;

Mit der phrasenhaften Formel „es geht nich DARum-=/=dass“ (Z. 548f.) beginnt der Einwand des Vaters gegen den Vorschlag des Sohnes. Er weist darin die inhaltliche Ausrichtung der Begründung zurück (Z. 549: dass die dass die mÄdchen und die JUNgen irgendwie miteinander was anfangen;) und markiert bereits in der Formel durch die Negationspartikel *nicht* seine Opposition. Die anschließend eingeschobene Einschätzung (Z. 551) stärkt die vorangegangene Opposition, indem durch die negierte Nominphrase „keine sorgen“ erneut eine explizit-sprachliche Negation erfolgt, mit der sich koordiniert eine Geste entfaltet. Es handelt sich um eine materialgebundene Geste, bei der der Sprecher einen Stift, ähnlich der *Horizontal Palm* Geste (vgl. Kendon 2004: 255ff.), horizontal durch die Luft führt – da der Stift durch die Finger wie in einem Gelenk gehalten wird, kann der Sprecher ihn in zügigen Bewegungen wiederholt von links nach rechts führen. Die horizontale, wiederholte Bewegung des Stiftes markiert hier gemeinsam mit dem Negationsartikel *keine* (Z. 551) verstärkt die Opposition des Vaters – und somit auch die Zurückweisung des Vorschlags des Sohnes. Nach Beendigung des *Turns*, die durch eine fallende Tonhöhenbewegung markiert wird, findet nun in der weiteren Aushandlung eine thematische Verschiebung statt und das ursprünglich implizierte und vom Vater elaborierte Argument möglicher Anbahnungen zwischen den Mitschüler*innen wird fallengelassen.

In dem zweiten Beispiel begründet der Sohn seine Präferenz für die Auswahl der Option ‚Hamburg‘ und schließt dabei vergleichend ‚Rom‘ als Ziel der Abschlussfahrt aus mehreren Gründen aus. In der anschließenden Äußerung der Mutter wendet diese gegen seine Begründung ein, indem sie auf die Besonderheit der Fahrt als *Abschlussfahrt* verweist. Während eine sprachliche Negation implizit bleibt, entfaltet sich innerhalb des Einwandes der Mutter eine *Horizontal Palm* Geste bzw. *Palm Down Horizontal Across* Geste, bei der die nach unten geöffnete Handfläche von einer Seite zur anderen geführt wird (vgl. Kendon 2004; Harrison und Larrivée 2016).

Beispiel 2 – Sichtbares Explizieren von Negationen in Einwänden

044 SO: und zwar einfach WEIL:- (.)
 045 hamburg erstens NICH so weit weg is wie ROM, (0.8)
 046 ((schnalzt)) und deswegen auch nicht TEUER wär-=
 047 =ich WÜRde n- (-)
 048 mal ABgesehen davon äh (.) würde n musicalbeSU:CH (.) schüler aus
 049 meiner klasse wohl Eher ansprechen;=
 050 =als zum beispiel ein kultURprogramm,
 051 ROM.
 052 das wär auch wesentlich !BIL!liger.
 053 (1.7)
 053 SO: [((zeigt mit geöffneter HF Ri Blatt))]
 054 MU: [ja aber] GUCK mal;
 055 ihr habt ja jetzt mal das GLÜCK,
 056 dass der förder|verein: (.) euch das FI|nanziert. |
 so-gz |blickt zu MU-----|
 mu-gz |blickt zu SO|
 057 SO.
 058 und das is ja |eure |ALLer |letzte (-) |fährt. |
 mu-gs |.....|führt rH Palm Down n. außen-|hält-----|
 mu-gz |blickt
 zu SO (>Z. 062)
 059 DAnach löst sich ja eure gemeinschaft aus.
 060 das_s das LETzte was ihr nochmal zusammen macht. (.)
 061 °hh und ich DENke- (.)
 062 solltest du mal über|LEgen-
 mu-gz ----->|
 063 was (.) jungs und mädels (.) so gerne MACHen;
 064 MEINSte nicht;=
 065 =dass so_n segeltörn (--) mal ne GANZ tolle erfahrung is auf so_m
 schiff,
 066 (um) nochmal so um die gemeinschaft (1.0) zuSAMmenzubringen und SPASS
 zu haben,=



Die Begründung des Sohnes (Z. 044-053) ist gefolgt von einer schwachen Zustimmung (vgl. Pomerantz 1984) sowie einem Adversativjunktore (Z. 054: ja aber), mit dem die Mutter in ihren Einwand einleitet: In diesem entkräftet sie zuerst das Argument des Sohnes, dass eine Abschlussfahrt nach Hamburg günstiger sei als nach Rom (vgl. Z. 046, 051), indem sie auf die Modalitäten der Abschlussfahrt verweist (Z. 055f.: ihr habt ja jetzt mal das GLÜCK, / dass der förderverein: (.) euch das FINANZIERT.). Syntaktisch mit der ersten Proposition durch die Konjunktion *und* zurückgebunden nimmt sie anschließend auf die Besonderheit der Klassenfahrt als „ALLerletzte (-) fährt“ (Z. 058) Bezug und verbindet somit das Argument der Finanzierung mit der Bedeutung einer Abschlussfahrt. Dabei wird das Adjektiv „ALLerletzte“ (Z. 058) durch die Fokusakzentuierung, das Steigerungspräfix sowie die anschließende Pause, die das Lexem vom Kopf der Nominalphrase absetzt, besonders hervorgehoben. Zeitlich koordiniert die Sprecherin eine *Horizontal Palm* Geste mit dem Lexem „ALLerletzte“ (Z. 058) und führt die nach unten orientierte Handfläche in einer lateralen Bewegung vom eigenen Körper weg (vgl. Kendon 2004). Während die Äußerung oberflächensprachlich keine explizite Negation enthält, sondern nur *ex negativo* erschlossen werden kann, dass auf die „ALLerletzte (-) fährt“ (Z. 058) *keine* weitere folgen wird, deckt die

Gestenäußerung die Negation sichtbar für alle Beteiligten auf: Die *Horizontal Palm* Geste „makes kinestically explicit the [...] negation that is implied by what is in the word“ (Kendon 2004: 255, Hervorh. i. O.). Während zu Beginn des Äußerungspakets noch die erste Proposition des Einwandes mit einem Adversativjunktore eingeführt und deutlich als Opposition markiert wird, expliziert nun in der angeschlossenen zweiten Proposition die Geste eine sprachlich ausschließlich implizit markierte Negation – durch sie wird die Negation in der Opposition der elternseitigen Äußerung für alle Beteiligten sichtbar und überdies durch den *Post-stroke Hold* (vgl. Z. 058) bis zum Ende der Turnkonstruktionseinheit gehalten, bevor die Sprecherin dann eine andere Option zur Aushandlung relevant macht (Z.061-066). Das Halten der Geste „could help speakers to indicate the scope of their negations more clearly“ (Harrison 2010: 46) – in der vorliegenden Sequenz ermöglicht das Halten der Geste das Bezugsobjekt der Negation, das erst am Ende der Turnkonstruktionseinheit benannt wird (Z. 058: fährt), zu identifizieren.

Zusammenfassend kann festgehalten werden, dass sprachlich explizierte Negationen in Einwänden gestisch verstärkt (vgl. Beispiel 1) oder, bei sprachlich impliziter Negation, durch *Horizontal Palm* Gesten erst sichtbar gemacht werden (vgl. Beispiel 2). Dabei kann die Geste, wie das zweite Beispiel nahelegt, dazu beitragen, dass Oppositionen gesichtsschonend markiert werden.

8.4.3 Anhalten einer vorangegangenen Opposition in Konzessionen

In ko-konstruierten Begründungen, in denen Interagierende kollaborativ für oder gegen die Wahl einer der vorgegebenen Optionen argumentieren, können die Beteiligten ihre Zustimmung auch im Rahmen von umgekehrten Konzessionen (vgl. Kap. 2.2.2) ausdrücken. Dabei endet die Konzession dann nicht mit einer opponierenden Handlung, wie von Couper-Kuhlen und Thompson (2000) für die *Cardinal Concessive* (X, X', Y) beschrieben worden ist, sondern mit einer schwachen Zustimmung (X, Y, X').⁸⁸ Aufgrund der Umkehr der konzessiven Struktur steht nun also die abgeschwächte Zustimmung X' zu Position X am Ende des *Turns* und kann so potenziell in die Fortsetzung der Interaktion aufgenommen werden (vgl. Couper-Kuhlen und Thompson 2000).

In den vorliegenden Daten zeigt sich innerhalb der beschriebenen Struktur der umgekehrten Konzession nun, dass die opponierende Handlung (Y) durch körperlich-sichtbare Ressourcen als ‚angehalten‘ markiert wird: Zeitlich koordiniert mit dem Einsetzen der schwachen Zustimmung, die i. d. R. mit dem Junktore *aber* eingeleitet wird, wird eine kleinräumige *Holding Away* Geste (vgl. Bressem und Müller 2014a: 1582) geäußert, bei der die Sprecher*innen aus einer Ruheposition heraus die Handfläche oder Fingerspitzen ausstrecken (vgl. Tab. 23). Die Handfläche ist leicht vom Körper des*der Sprecher*in weg orientiert und markiert ein „stopping from continuation“ (Bressem und Müller 2014a: 1582) – sie hält somit den vorangegangenen Einwand an, wie Beispiel 1 zeigt.

Anhalten einer vorangegangenen Opposition in Konzessionen	
Beschreibung der multimodalen Gestalt	schwache Zustimmung in einer Konzession (X, Y, X') + <i>Holding Away</i> Geste + Absenken des Kopfes

Tabelle 23: Beschreibung der multimodalen Gestalt „Anhalten einer vorangegangenen Opposition in Konzessionen“

⁸⁸ Nach Couper-Kuhlen und Thompson (2000: 392ff.) handelt es sich um eine Variante der *Cardinal Concessive*.

Das folgende Beispiel illustriert, wie die Tochter eine abschließende Evaluation (X) der Mutter erst zurückweist (Y), sie korrigiert und ihr dann schwach zustimmt (X'). Dabei erklärt sie ihre Opposition retrospektiv durch eine *Holding Away* Geste für ungültig.

Beispiel 1 –Anhalten einer vorangegangenen Opposition in Konzessionen

210 MU: ((schnalzt)) Okay. |
 mu-gz >>blickt auf Blatt->>
 to-gz ----->|blickt auf Blatt (>Z. 219)
 211 (2.0)
 212 MU: würd ich auch vorschlagen;
 213 |das_s [für]|ne GROSse klassengemeinschaft bestimmt |°h (1.7) am
 spannendsten sei[n könnte;]
 214 TO: [ja:;]
 mu-he |nickt-----|
 to-he |nickt-----|
 215 [SPRICHT] bestimmt nicht ALle an==

 216 =a|ber: |wahrsch |vielleICHT|
 to-gs |öffnet lH Palm Down|,,,,,,|
 to-he |senkt Kopf n. li ab|
 mu-he |.....|
 |die mEis |ten;
 mu-he |senkt Kopf nach re ab-----|
 mu-gs |öffnet lDaumen & rDaumen, hebt ineinandergelegt Hde an|
 217 AUSserdem (-) <<p>will man ja auch etwas zusAmmen machen;>=
 218 =<<pp>und-> (1.2)
 219 jetzt NICHTS mit shoppen (-) so GROSS|artig;
 to-gz ----->|blickt zu MU->>

In dem vorliegenden Transkriptauszug stimmt die Mutter einem letzten Argument der Tochter für die Auswahl des Segeltörns zu (Z. 210, 212) und schließt die Begründung mit einer durch das Satzadverb modalisierten und gesteigerten Evaluation (Z. 213: *bestimmt °h (1.7) am spannendsten*) ab. Im direkten Anschluss an die Bewertung (X) beginnt die Konzession der Tochter, in der sie noch überlappend mit dem *Turn* der Mutter eine Gegenbehauptung formuliert (Y, Z. 215) und anschließend, epistemisch abschwächt (Z. 216: *vielleICHT*), eine Zustimmung (X') äußert. Die Gegenbehauptung ist auf sprachlicher Ebene durch die Negation *nicht* markiert und ihr Bezug zur Äußerung der Mutter wird durch das Recyceln (vgl. Goodwin 1990; Arendt und Zadunaisky Ehrlich 2020) des Satzadverbs *bestimmt* (Z. 213, 215) erkennbar. Der Einwand wendet sich inhaltlich gegen den Aspekt, dass der Segeltörn für die „GROSse klassengemeinschaft“ (Z. 213) und somit für die Klasse insgesamt „am spannendsten sein könnte“ (Z. 213). Die Tochter schränkt mit der Konzession die Behauptung der Mutter ein, indem sie nicht nur die Bezugsgröße (*Klassengemeinschaft* vs. *nicht alle/die meisten*), sondern auch die Evaluation durch die Verwendung eines semantisch qualitativ schwächeren Verbs (*könnte am spannendsten sein* vs. *ansprechen*) herabstufend korrigiert (vgl. Kotthoff 1993; Couper-Kuhlen und Thompson 2000). Mit dem in die Zustimmung einleitenden Adversativjunktorkor *aber* hebt sie die Finger der linken Hand gestreckt und vertikal öffnend als *Holding Away* Geste an (vgl. #1) und markiert (retrospektiv) ein Anhal-

ten und Nichtfortsetzen der vorangegangenen Opposition, während sie gleichzeitig eine schwache Zustimmung zur Bewertung der Mutter äußert. Mit dem Anheben der Hand senkt sie zugleich den Kopf ab. Durch das Absenken des Kopfes zu einer Seite können Sprecher*innen einerseits auf äusserungsstrukturierender Ebene einen Übergang zu einem neuen Abschnitt innerhalb ihres Beitrags markieren, wie es beispielsweise in der Erwerbsstudie von Reig Alamillo et al. (2012) für das Erzählen nachgezeichnet werden konnte (vgl. auch Kendon 1972b), sowie auch inhaltliche Perspektivwechsel sichtbar einleiten (vgl. Maury-Ruan 2011). Hier markiert die Sprecher*in durch das Absenken des Kopfes auf dem turninitialen Adversativjunktorkonjunktoren *aber* den Übergang zu einer Zustimmung, während sie zugleich durch die *Holding Away* Geste die Opposition anhält. Interessant erscheint, dass die gegenläufige Bewegung von Kopf- und Handgeste (nach unten bzw. oben) hier metaphorisch die inhaltliche Gegenbewegung von Gegenbehauptung (Y) und Zustimmung (X') körperlich-sichtbar zum Ausdruck bringt. Insgesamt übernimmt die Sprecherin mit Einbringen der Konzession nicht nur die Verantwortung für das Abschließen der Begründung, sondern auch für die Bewertung der verhandelten Option, der die Mutter durch körperliche Ressourcen schwach zustimmt (Z. 216, vgl. dazu ausführlich Kap. 8.3.3).

Zusammenfassend können also Interagierende in Konzessionen mit der Struktur X, Y, X' eine selbst eingebrachte Opposition retrospektiv in der dann anschließenden, abgeschwächten Zustimmung als ‚angehalten‘ markieren, indem an der Scharnierstelle der Konjunktion *aber*, die die Opposition abschließt und die Zustimmung einleitet, eine *Palm Away* Geste geäußert wird. Dieser Übergang von dem vorherigen Einwand zur schwachen Zustimmung wird ferner durch das Absenken des Kopfes begleitet, sodass sich in einem metaphorischen Sinne die Gegenbewegung der Konzessionselemente in der Gegenläufigkeit der Hand- und Kopfgeste wiederfindet. Insgesamt konnte gezeigt werden, dass in dieser gesamten multimodalen Gestalt die eingebrachte Opposition in der Konzession retrospektiv durch die Geste als ‚angehalten‘ kontextualisiert wird. Der Geste kommt dabei innerhalb des Ressourcenverbundes eine performative Funktion zu.

8.4.4 Mahnendes Einwenden

Wenn Sprecher*innen etwas gegen eine Vorgängeräußerung einwenden, können sie den Einwand durch das Erheben des Zeigefingers besonders in den Fokus der Beteiligten rücken. Die Geste selbst wird dabei entweder als emblematische „Achtung-Geste“ (Fricke 2007: 153) beschrieben, der i. S. einer Mahnung performative Funktionen zukommen. Sie wird aber auch als *Recurrent Gesture* beschrieben, deren pragmatische Funktion darin besteht, „[to draw] the attention of other participants to new particular important topics of talk as well as to signal[] thematic shifts, such as when dismissing the statement of others“ (Bresse und Müller 2014a: 1583). Der Geste kommt somit auch eine kataphorische, äusserungsstrukturierende Funktion zu, denn sie verweist (projizierend) auf Neues oder Wichtiges innerhalb des eigenen *Turns* oder zeigt einen Themenwechsel an (vgl. Bresse und Müller 2014a: 1583; Inbar 2022; vgl. auch Kendon 2004: 142). Inbar (2022) beobachtet das Vorkommen der Geste an sequenziellen Positionen, in denen die Äußerungen des*der Sprecher*in die Erwartung des Gegenübers verletzt. Besonders häufig kommt die Geste dabei in Oppositionen vor (vgl. Inbar 2022, 2025). In den vorliegenden Daten wird die Geste innerhalb von Einwänden geäußert (vgl. Tab. 24).

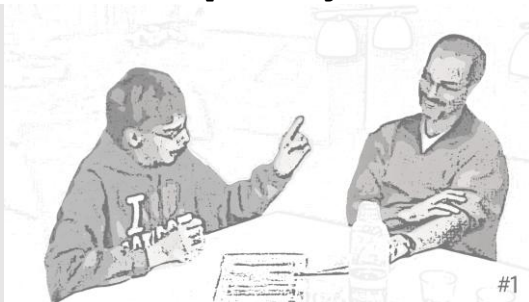
Mahnendes Einwenden	
Beschreibung der multimodalen Gestalt	Einwand + turninitial erhobener Zeigefinger

Tabelle 24: Beschreibung der multimodalen Gestalt „Mahnendes Einwenden“

Das folgende Beispiel illustriert, wie der Einwand des Sohnes durch das Erheben des Zeigefingers angekündigt und durch das Wiederholen der Geste mit der anderen Hand über den Einwand hinweg aufrechterhalten wird.

Beispiel 1 – Mahnendes Einwenden

328 VA: al|so bei FÜNF tage SEgeltörn auf !EI!nem großen segel|schiff,
so-gz |blickt auf Blatt-----|zu VA (>Z. 330)
329 VA: sind !HUN!dertprozentig (1.3) !BEI!de lehrer mit an bord.



330 SO: |(-) |A | |ber; ||
so-gs |weist mit lIF Richtung VA|,,,,,,|
so-gz ---->|blickt auf Blatt----->>

331 VA: |!GANZ!| !ENG! bei [euch.]|
va-gs ----->|



332 SO: | |es gibt |
so-gs |zeigt mit rIF Ri VA|



HUNdertproz[entig auch mindestens] |(!) !EI! |nen (.) seekranken;
so-gs |rIF Ri VA|

333 VA: [und bei |v::]

334 |(!) |
va-fa |zieht AB zusammen|

335 SO: und [DEM macht das] dann keinen spaß.

In der vorliegenden Sequenz bringt der Vater eine Begründung gegenüber der Option des ‚Segeltörns‘ ein (Z. 328f.), in der er durch die Verwendung des Indikativs, des epistemisch verstärkenden Adjektivs *hundertprozentig* sowie die gehäuft auftretenden Fokusakzentuierungen eine starke epistemische Haltung markiert (vgl. Heritage 2012b). Wie in dem Fortgang der Interaktion deutlich wird, interpretiert der Sohn die Begründung des Vaters als Fürsprache für den Segeltörn, gegen die er nun einen Einwand

vorbringt (Z. 330, 332, 335): Während einer kurzen Pause (Z. 330) und vor der Turnübernahme weist der Sohn bereits mit dem linken Indexfinger in Richtung des Vaters (vgl. #1), dringt dabei relativ weit im gemeinsamen Gestenraum vor (vgl. Sweetser und Sizemore 2008; vgl. Kap. 6.2) und rückt die Geste somit in den Wahrnehmungsraum des Vaters. Mit der Geste verweist der Sprecher hier kataphorisch auf die Relevanz der nun folgenden Äußerung (vgl. Bressemer und Müller 2014a). Die unmittelbare Einleitung des Turns durch den fokusakzentuierten Adversativjunktorkonkret *aber* markiert ihn dann sprachlich als Opposition. Im Verbund mit der sprachlich markierten Opposition übernimmt die Geste zu der kataphorischen auch eine mahnende Funktion. Die Rückbindung des Einwandes des Sohnes an die Äußerung des Vaters wird ferner durch das Recyclen (vgl. Goodwin und Goodwin 1987) einzelner Lexeme samt ihrer Fokusakzentuierung (Z. 329: !HUN!dertprozentig / Z. 332: HUNdertprozentig) sowie kontrastierender Gegenüberstellungen (Z. 329: !BEI!de / Z. 332: !EI!nen) markiert. Auch der Sohn nimmt dabei eine starke epistemische Haltung ein und weist die Begründung des Vaters durch sein Gegenargument deutlich zurück.

Sowohl durch den verbalsprachlich geäußerten Einwand als auch durch die vertikal ausgerichtete Geste, mit der Interagierende eine Unterbrechung oder das Anhalten des Gesprächs anzeigen können (vgl. *Vertical Palm* Geste Kendon 2004: 251ff.), wird insgesamt die Progression des vom Vater eingebrachten Arguments vorläufig unterbrochen und *zugleich* in strukturierender Funktion kataphorisch auf die Relevanz der nun folgenden Äußerung hingewiesen. Die Geste schafft außerdem einen interpretativen Rahmen der nun folgenden Sprechhandlung: Sie zeigt gemeinsam mit dem verbalsprachlich geäußerten Einwand eine mahnende Zurückweisung der Begründung des Vaters an – die dann, indem sie durch die andere Hand des Sprechers während des Einwandes wiederholt geäußert wird (vgl. #2, #3), diesen Rahmen aufrechterhält. Insgesamt kontextualisiert die Geste die argumentative Handlung also als mahndes Einwenden. Sie trägt sowohl performative als auch äuerungsstrukturierende Funktionen.

8.4.5 Zusammenfassung

Im Rahmen des vorliegenden Kapitels konnten die folgenden wiederkehrenden und funktional für das Einwenden eingesetzten multimodalen Gestalten rekonstruiert werden:

- Einwände oder Segmente von Einwänden innerhalb größerer Äußerungspakete können auf der geöffneten Handfläche präsentiert werden (vgl. Kap. 8.4.1). Dabei können auch interpersonelle Bezugnahmen sowie Bezugnahmen auf Objekte der Umgebung hergestellt und für den Einwand relevant gemacht werden. Interpersonelle Bezugnahmen können ferner anzeigen, dass problematisierte Anteile einer Äußerung an das Gegenüber zurückgegeben werden und dadurch eine Rückmeldung, z. B. in Form einer weiteren Begründung, elizitieren – ohne dass eine sprachliche Explizierung dabei notwendig wird. Möglicherweise kann somit die *sprachliche* Angriffsfläche reduziert werden.
- In Einwänden können Negationen durch Gesten mit horizontaler Bewegungsrichtung die sprachlich explizierte Negation epistemisch verstärkend oder aber sprachlich implizite Negationen sichtbar machen (vgl. Kap. 8.4.2) – und so auch dazu beitragen, eine Opposition gesichtschonend(er) einzubringen.
- In Konzessionen mit der Struktur X, Y, X' können Interagierende ihren selbst eingebrachten Einwand (Y) innerhalb der schwachen Zustimmung (X') durch eine *Holding Away* Geste anhalten und zugleich durch das Absenken des Kopfes einen Perspektivwechsel markieren (vgl. Kap. 8.4.3). Durch diese Variante der *Cardinal Concessive* (vgl. Couper-Kuhlen und Thompson 2000) steht nun die schwache Zustimmung am Ende des *Turns*, die potenziell in der weiteren

Interaktion aufgenommen werden kann. In der vorliegenden Studie kommen diese multimodalen Gestalten in Kontexten vor, in denen Interagierende konsensual sich gegenseitig ein (lokales) Abschließen z. B. einer Aushandlungssequenz angezeigt haben und durch das Einbringen der Konzession nun dieses Abschließen in den Verantwortungsbereich der konzедierenden Person übergeht.

- Interagierende können Einwände in den besonderen Fokus der Aufmerksamkeit stellen und durch Zeigefingergesten kataphorisch auf Relevantes verweisen, das nun in die Interaktion eingebracht wird. In der argumentativen Handlung des Einwendens kann dann die Zeigefingergeste den Inhalt einer Vorgängeräußerung anmahnen und sowohl durch das Einwenden selbst als auch durch die Zeigefingergeste auf eine Unterbrechung des Fortgangs der Progression hindeuten (vgl. Kap. 8.4.4).

8.5 Einfordern von Rückmeldungen

Im Rahmen argumentativer Entscheidungsfindungen kann das Einfordern von Rückmeldungen erfolgen, um beispielsweise lokale Präferenzabfragen wie *Wofür bist du?* an das Gegenüber zu richten, oder aber durch global operierende Zugzwänge wie *Warum?* eine Plausibilisierung für bezogene Positionen oder eingebrachte Behauptungen einzufordern (vgl. Kap. 2.2.2). Um eine *gemeinsame* Grundlage der Entscheidungsfindung sicherzustellen, können Interagierende außerdem eine gemeinsame Wissensbasis etablieren, indem sie beispielsweise einander auffordern, Argumente für bzw. gegen eine Option zusammenzutragen, und dieses Wissen dann für die gemeinsame Entscheidung nutzen. Auch hierzu nutzen sie die argumentative Handlung des Einforderns von Rückmeldungen.

Wenn Interagierende nun durch Fragen eine Antwortverpflichtung etablieren, können unterschiedliche Ressourcen eingesetzt werden, um beispielsweise die konditionelle Relevanz für eine Rückmeldung zu verstärken. So wird in der Untersuchung von Stivers und Rossano (2012) ersichtlich, dass in Abhängigkeit vom Turndesign sowie den eingesetzten Ressourcen, d. h. abhängig von lexikalischen, morphosyntaktischen und prosodischen Merkmalen sowie körperlich-sichtbaren Ressourcen wie dem Blick, die Antwortrelevanz (*response relevance*, vgl. Stivers und Rossano 2012: 78) bzw. die Antwortverpflichtung verstärkt werden kann: „We suggest that speakers mobilise response through the combination of many resources employed simultaneously” (Stivers und Rossano 2012: 59). Dass Gesten ebenfalls zu Ressourcen gehören, die Rückmeldungen des Gegenübers einfordern können, zeigt beispielsweise Streck (2007: 168), der die Funktion der turnfinal gehaltenen *Palm Up* Geste untersucht, die sichtbar „markiert, dass die von der Äußerung, zu der sie gehört, gesetzte konditionelle Relevanz noch unerfüllt ist“. Und auch Holler et al. (2018) arbeiten heraus, dass körperlich-sichtbare Ressourcen auf die Rückmeldung des Gegenübers Einfluss nehmen können, denn der Einsatz gestischer Ressourcen während des Einbringens einer Frage kann die Zeit der Turnübernahme durch das Gegenüber sichtlich verkürzen.

Die Daten der vorliegenden Studie zeigen nun außerdem, dass gestische Ressourcen in fragenden, eine Antwort einfordernden Sequenzen auch eingesetzt werden können, um das Gegenüber in ein spezifisches Antwortformat zu steuern: Dazu verwenden die Fragenden bereits im Aufbau ihrer Frage Zählgesten, die die Erwartung anzeigen, dass die Antwort des Gegenübers mehr als ein Argument beinhaltet und in einer Aufzählung vorgebracht wird (vgl. Kap. 8.5.1). Ferner können Sprecher*innen den Gesprächsgegenstand, auf den sie sich in ihrer Frage beziehen, durch deiktische Gesten auf das Material auch in der Antwortverpflichtung des Gegenübers präsent halten und diesen dadurch zugleich relevant setzen (vgl. Kap. 8.5.2) – sie tragen somit zur Gesprächsthemensteuerung bei. Im Folgenden werden beide multimodalen Gestalten vorgestellt.

8.5.1 Einfordern von Rückmeldungen im Listenformat

Interagierende können in argumentativen Entscheidungsfindungen auf gemeinsames Wissen rekurren und dieses sowohl für das Abschließen der Begründung als auch für weitere Aushandlungen relevant setzen. Das Wissen kann dabei entweder von den Sprecher*innen selbst eingebracht oder aber fragend vom Gegenüber elizitiert werden. Es zeigt sich, dass mit Einbringen der Frage nicht nur das elizitierte Wissen relevant gesetzt wird, sondern dass durch gestische Ressourcen auch die Erwartungen über das *Antwortformat* erkennbar gemacht werden können: Durch das Produzieren und turnfinale Halten von Zählgesten zeigen Sprecher*innen an, dass für die Bearbeitung der von ihnen eingeforderten Rückmeldung mehrere Inhaltselemente erwartet werden. Die Fragenden steuern somit die Rückmeldung des Gegenübers in ein spezifisches Antwortformat – nämlich in die Produktion einer Liste (vgl. Selting 2004). Die nachfolgende Übersicht 25 enthält eine zusammenfassende Beschreibung der multimodalen Gestalt sowie ihrer Funktion.

Einfordern von Rückmeldungen im Listenformat	
Beschreibung der multimodalen Gestalt	• Zugzwang + Zählgeste + Halten der Geste über Sprecher*innenwechsel hinweg

Tabelle 25: Beschreibung der multimodalen Gestalt „Einfordern von Rückmeldungen im Listenformat“

Die nachfolgenden Beispiele illustrieren, wie Interagierende eine Rückmeldung ihres Gegenübers in dem Antwortformat einer Liste bzw. Aufzählung elizitieren. Das Einlösen des erwarteten Antwortformats kann dann von der Person, die die Rückmeldung eingefordert hat, durch das Abzählen an der eigenen Hand ratifiziert werden (vgl. Beispiel 1), während das Ausbleiben einer solchen Zählgeste während der Rückmeldung des Gegenübers anzeigen kann, dass die Antwort nicht den Erwartungen der Fragen initiiierenden Person entspricht (vgl. Beispiel 2).

In dem nachfolgenden Transkriptauszug, Beispiel 1, kündigt die Mutter eine resümierende Zusammenfassung der bisherigen Argumentation an. Mit einem dann einsetzenden Zugzwang äußert sie die erste Zählgeste, die sie in dem anschließenden Äußerungspaket der Tochter fortsetzt.

Beispiel 1 – Einfordern von Rückmeldungen im Listenformat

165	MU:	also-
166		(.)
	mu-gs	...
	mu-gz	>>auf Blatt->>
167		wir halten FEST,
		
168		argument EINS (.) wäre?
	mu-gs	Zählgeste mit rIF an lDaumen (>Z. 171)
169		(1.0)
170	TO:	((räuspert sich)) (.) (1.0)
	to-fa	zieht AB hoch
	to-gz	blickt zu MU (>Z. 173)
171	MU:	wenn man von
	mu-gs	

		[TÄgen ge	
	mu-gs	zeigt mit rIF Ri Option 4 & 1, hält lDaumen ausgestreckt-	
		ht;]	
	mu-gs	führt Hde in Zählgeste zurück->	
172	TO:	[fü (.) achso	
		] jetzt (.) ich dachte n (nun) ALLgemein,	
173		O de:r,	
	to-gz	-> blickt zu MU (>Z. 175)	
174	MU:	ist der	LÄNGS te ausflug,
	mu-gs	Zählgeste mit rIF an lDaumen-	Beat- hält Zählgeste (>Z. 179)
175	TO:	al so ALLgemein ja;	
	to-gz	-> blickt auf Blatt (>Z. 178)	
176		(--)	
177		er is <<p>das:> (.) LÄNG ste,	
178		man macht was zuSAMmen,	
	to-gz	 blickt zu MU- blickt auf Blatt (>Z. 182)	
			
179	MU:	((zählt weiter mit der rH an lH ab))	
180		(1.3)	
			
181	TO:	man fährt (.) naja (.) so MITtelmäßig weg;	
	to-gs	rotiert bHde Palm Away vor & zurück	
	to-he	bewegt Kopf von re n. li	
	mu-gs	zählt an MF ab, hält lH mit Zählgeste (>Z. 208)	
		((...))	
207		(1.0)	
208	TO:	ja;	
	mu-gs	-->	

Die Mutter fordert in der vorliegenden Sequenz zu einer resümierenden Konklusion auf (Z. 165f.). Dabei zeigt sie durch den Verbund sprachlicher und gestischer Ressourcen die Erwartung an, dass die eingeforderte Konklusion aus mehr als einem Argument besteht (Z. 168: argument EINS wäre?; vgl. #1). Es wird somit erwartbar, dass eingeforderte Rückmeldung aus einer Liste von Argumenten besteht. Die Mutter hält die Zählgeste turnfinal und setzt sie somit für die Rückmeldung der Tochter relevant. Nach anfänglichen Verzögerungen (Z. 169-170) sowie Reparaturinitiationen (Z. 172-173) bringt die Mutter nun das erste Argument selbst ein und wiederholt die anfängliche Zählgeste (vgl. Z. 168, 174). Nach einer Wiederholung des ersten Arguments der Mutter (Z. 177) setzt nun die Tochter die Liste fort (Z. 178, 181). Mit jedem neuen Listenelement, d. h. mit jedem neuen Argument, das sie einbringt, zeigen sich Veränderungen der körperlich-sichtbaren Ressourcen der Mutter, denn diese erweitert sukzessive

ihre Zählgeste (vgl. #2, #3). Dadurch ratifiziert die Mutter die jeweils eingebrachten Argumente der Tochter und zeigt zugleich an, dass das erwartete Antwortformat eingelöst wird.

Auch in der nachfolgenden Sequenz, Beispiel 2, fordert die Mutter zu einer Rückmeldung in Form einer Liste auf. Allerdings entspricht hier die Rückmeldung des Sohnes nicht dem von der Mutter projizierten bzw. angesteuerten Listenformat, sodass sie diese anschließend selbst erbringt.

Beispiel 2 – Einfordern von Rückmeldungen im Listenformat

146 MU: dann hast DU: jetzt noch (mir/mehr) gesagt-

147 (1.5)

148 hier-

149 | (---) | CEN | ter | parc,

mu-gs | | tippt mit rIF auf Blatt |

mu-gz

| blickt zu SO (>Z. 164)



150 wir | WÄren ja schonmal in so_m | center parc-

mu-gs | | Zählgeste: tippt mit rIF gegen lDaumen |

151 | was kann man da alles MA | chen,

mu-gs | tippt mehrmals mit rIF an Daumen- | hält (>Z. 155)

152 SO: ja;=was man WILL eigentlich;

153 MU: naja:.

154 was man will-



155 | (--) |

mu-gs | zählt mit rH bis lMF ab |



156 | fällt dir was EIN, | (.)

mu-gs | zählt weiter am lIF & klFi ab- | tippt an lDaumen; hält (>Z. 158)

157 SO: ja;

158 <<lachend>ach> | man kann ACHterbahn fa?-

mu-gs -----> |

159 also | man | kann ALles ma|chen;

so-bo | Shoulder Shrug |

mu-gs

| hebt lDaumen erneut an (>Z. 161)

160 eigentlich so;

161 MU: |da is_n SCHWIMMbad |dabei,
 mu-gs |tippt mehrmals mit rIF abzählend an lDaumen, hält (>Z. 164)
 so-gz |blickt Ri Geste der MU (>Z. 164)
 162 SO: ja;
 163 MU: (da könnt_a/man könnt da) SCHWIMMen gehen-
 164 wenn die |LEHrer-| | (-)
 mu-gz ----->|blickt auf Blatt (>Z. 168)
 mu-gs ----->|zeigt mit rIF Ri Blatt; lDaumen hält Zählgeste|
 so-gz ----->|blickt auf Blatt----->>

Die Mutter etabliert zu Beginn des Ausschnitts eine weiter zu verhandelnde Option (,Center Parc‘) und zeigt durch den Verweis auf die gemeinsame Erfahrung (Z. 150: wie WAreN ja schonmal in so_m center parc-) an, dass die Interagierenden in der Verhandlung dieser Option auf ein gemeinsames Wissen (vgl. Clark und Marshall 1981) zurückgreifen können. So kann das Erfahrungswissen beispielsweise eingesetzt werden, um hypothetische Überlegungen anzustellen, welchen Aktivitäten die Klasse in einem Center Parc möglicherweise nachgehen kann (vgl. Z. 161f.: da is_n SCHWIMMbad dabei, / (da könnt_a/man könnt da) SCHWIMMen gehen-).

Dabei fordert nicht allein das Frageformat, d. h. das Fragepronomen in Kombination mit dem Zahlwort *alles* (Z. 151: was kann man da alles MACHen,), die Benennung mehrerer Aktivitäten ein, sondern es wird bereits in Zeile 150 eine Zählgeste etabliert, in der der Zeigefinger der einen mehrmals gegen den Daumen der anderen Hand tippt (vgl. #1): Das Turndesign sowie die Zählgeste zeigen gemeinsam die Erwartung der Mutter hinsichtlich des Antwortformats an – sie fordert eine Liste möglicher Aktivitäten ein. In der Rückmeldung kommt der Sohn dieser Aufforderung allerdings nur unzureichend nach, denn seine Antwort fällt inhaltlich unspezifisch aus (Z. 154: was man WILL eigentlich;). Diese Bearbeitung der Frage weist die Mutter anschließend durch das Vorlaufelement *naja* zurück und erneuert nun den Zugzwang (Z. 153ff.: naja:./ was man will-/ (--)/ fällt dir was EIN,). In der entstehenden Pause äußert sie eine weitere abzählende Geste (vgl. #2-#5), die abschließend in die ursprüngliche Gestenform (Zeigefinger an Daumen zur Markierung eines ersten Listenelements, vgl. #6) wieder zurückkehrt und gehalten wird (Z. 156ff.). Der Sohn bringt nun, lachend, eine Beispielaktivität (Z. 158: <<lachend>ach> man kann ACHterbahn fah-) ein, die er abbricht. Anschließend formuliert er die erste Frage der Mutter (Z. 151: was kann man da alles MACHen,) als allgemeine, unspezifische Proposition (Z. 159: also man kann ALles machen;). Auffällig ist, dass die Mutter dieses Transkriptbeispiels (anders als in Beispiel 1) bei der Benennung der unspezifischen Aktivität keine Zählgeste äußert, mit der sie das erste Listenelement ratifiziert, sondern den Daumen in der Ruheposition hält und erst am Turnenende des Sohnes die Geste wieder ‚reaktiviert‘. Nun übernimmt die Mutter selbst die Aufzählung (Z. 161).

Zusammenfassend kann festgehalten werden, dass in eingeforderten Rückmeldungen die multimodale Gestalt relevante Hinweise für die Beteiligten über das erwartete Antwortformat – in Form einer Liste bzw. Aufzählung – enthält. So können bereits im Zugzwang Nummerierungen auf sprachlicher Ebene eingeführt werden (z. B. *Argument eins wäre?*) und/oder vermittlels Zählgesten markiert werden. Einzelne Listenelemente können dann rezipient*innenseitig durch das Fortführen der Zählgeste ratifiziert werden (vgl. Beispiel 1) und unzureichende Bearbeitungen, z. B. in Form einer unspezifischen satzförmigen Antwort, die keine (weiteren) Listenelemente enthält, zurückgewiesen bzw. als nicht erwartungskonform markiert werden (vgl. Beispiel 2). Es handelt sich um eine die Folgehandlungen stark steuernde multimodale Gestalt des Einforderns von Rückmeldungen, der eine interaktionsorganisatorische Funktion zukommt. Darüber wirkt sie ebenfalls durch das Einfordern der Rückmeldung in Form einer Liste äusserungsstrukturierend auf den Beitrag des Gegenübers.

8.5.2 Präsenthalten des Gesprächsgegenstandes in eingeforderten Rückmeldungen

Doch nicht nur das Format der Antwort kann visuell sichtbar gemacht bzw. äusserungsstrukturierend beeinflusst werden (vgl. Kap. 8.5.1); durch das Zusammenspiel unterschiedlicher Ressourcen zeigen die Sprecher*innen an, welche Optionen sowie inhaltlichen Aspekte der Aufgabe sie für die Rückmeldung des Gegenübers präsent halten, die es rezipient*innenseitig zu bearbeiten gilt (vgl. Tab. 26). Dabei kann sich das Präsenthalten des Gesprächsgegenstandes über einzelne satzförmige (vgl. Beispiel 1) sowie längere, satzübergreifende Antwortsequenzen (vgl. Beispiel 2) erstrecken. Durch Veränderungen der Position der deiktischen Geste auf dem Material kann die Antwort des Gegenübers interaktiv inhaltlich ‚nachgesteuert‘ werden, indem weitere bzw. andere Inhaltselemente der Aufgabe für die Antwort des Gegenübers relevant gemacht werden (vgl. Beispiel 2).

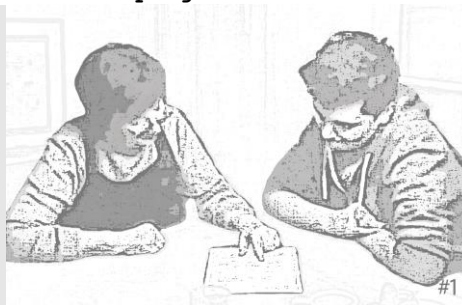
Präsenthalten des Gesprächsgegenstandes in eingeforderten Rückmeldungen	
Beschreibung der multimodalen Gestalt	Benennen eines Gesprächsgegenstandes + deiktische Geste auf Material + Halten der Geste über Sprecher*innenwechsel hinweg

Tabelle 26: Beschreibung der multimodalen Gestalt „Präsenthalten des Gesprächsgegenstandes in eingeforderten Rückmeldungen“

Das erste Beispiel legt dar, wie Sprecher*innen die Option, auf die sie sich beziehen, durch deiktische Gesten auf das Material einführen und diese innerhalb einer etablierten Antwortverpflichtung als Gesprächsgegenstand für das Gegenüber präsent halten können. Im vorliegenden Fall referenziert die Mutter in einer Rück- sowie einer Ergänzungsfrage auf eine spezifische Option, die sie auch innerhalb der Rückmeldungen des Sohnes weiter präsent hält.

Beispiel 1 – Präsenthalten eines Gesprächsgegenstandes in eingeforderten Rückmeldungen

132 SO: =ja gut das STIMMT. (-)
 133 aber im BUNgalowpark könnte man ja dann:- (--)
 134 ja:: gut das kann man in hamburg auch versCHIEdene gruppen bilden jeden tag. (---)
 135 m::: (--)
 136 ha: ich glaube aber dass das der bunga bungalowpark da BESSer geeignet ist.=
 137 =weil: (.) wenn man da so: SCHWIMmen kann;=
 138 =ANDere so (.) so freizeitaktivitäten machen kann.=
 139 =wäR das eigentlich glaub ich BESSer dann auch mit (-) shoppen gehen und so.
 140 ja in HAMburg könnte man dann vielleicht EINen tag shoppen gehen.=
 141 =aber (.) vielleicht könnt (.) also ich glaub im bungalowpark könnte man dann so jeden TAG machen was man will,
 142 und in hamburg hat man glaub ich FESTes programm.



143 MU: also |meinst du nummer |vier wäre am BES |ten;
 mu-gz >>blickt auf Blatt (>Z. 144)
 mu-gs |.....|tippt mit rIF auf Option-|hält rIF auf Option (>Z. 148)
 144 SO: joa ich DENke [schon.]

145 MU: [und] vom PREIS |her?
 mu-gz ----->|blickt zu SO (>Z. 147)
 146 (---) auch ne?
 147 SO: joa ich| denke;
 mu-gz ----->|
 148 MU: |ja gut |ob du nun VIER ta|ge hamburg oder |VIER |
 mu-bo |Shoulder Shrug-|kleinere Shrugs--|
 mu-gs ----->|tippt mit rIF|
 |tage bungalow |park| ()
 mu-gs |hält IF auf Blatt|,,,|

Innerhalb der Begründung des Sohnes werden zwei Optionen („Bungalowpark“ und „Hamburg“) verhandelt, in der er den Bungalowpark durch die Verwendung eines Komparativs als „BESSer“ (Z. 136, 139) bewertet als die Abschlussfahrt nach Hamburg. Allerdings schließt das Äußerungspaket mit einem Bezug zu Hamburg (Z. 142: und in hamburg hat man glaub ich FEStes programm.) ab. An dieser Stelle scheint eine Unschärfe in seiner Begründung zu entstehen, welches Ziel der Abschlussfahrt der Sohn nun bevorzugt, denn es schließt eine Rückversicherungsfrage seitens der Mutter an (Z. 143: also meinst du nummer vier wäre am BESTen;). Die Rückfrage wird in Form einer Ja-/Nein-Frage gestellt, in der die Sprecherin bereits die Annahme i. S. eines *Candidate Understanding* (vgl. Heritage 1984) formuliert, dass aufseiten des Sohnes eine Präferenz für den Bungalowpark vorliegt – der Sohn stimmt der Annahme zu (Z. 144). Zeitlich wird die Ja-/Nein-Frage mit einer deiktischen Geste auf das Material bzw. auf die Option koordiniert (vgl. #1) und auch nach der Affirmation des Sohnes weiter auf der Option gehalten, sodass eine weitere Bezugnahme der Mutter auf diese Option in der Fortsetzung der Interaktion projiziert wird. In der Folge schließt die Sprecherin nun eine Ergänzungsfrage mit Bezug zu der Option des Bungalowparks an (Z. 145: und vom PREIS her?) und hält die Option so lange gestisch präsent, bis sie anschließend selbst Stellung bezüglich des Vergleichs beider Optionen bezieht (Z. 148). Die Geste wird nun in eine Ruheposition überführt. Insgesamt kommt der deiktischen Geste in der abgebildeten Sequenz die Funktion zu, den Gesprächsgegenstand, auf den sich die Sprecherin bezieht, über mehrere Sprecher*innenwechsel hinweg als Bezugspunkt der aktuellen Interaktion präsent zu halten.

Innerhalb des zweiten Beispiels wird nicht nur deutlich, wie Sprecher*innen spezifische Gesprächsgegenstände präsent halten, sondern wie durch das Verändern der Geste gemeinsam mit Ergänzungsfragen weitere Gesprächsgegenstände für die Antwort des Gegenübers relevant gemacht und somit die Antwort des Gegenübers inhaltlich nachgesteuert werden kann.

Beispiel 2 – Präsenthalten eines Gesprächsgegenstandes in eingeforderten Rückmeldungen

159 °h oKAY;
 160 | (0.8)
 mu-gs |zeigt mit rIF auf Blatt (>Z. 163)
 161 MU: drei tage ROM,
 mu-gz >>blickt auf Blatt (>Z. 166)
 162 (1.5)
 163 °h ich WÜRD |das-
 mu-gs ----->|schiebt rIF an andere Stelle der Option->
 164 WEISS ich nich; |(-)
 mu-gs ----->|...->
 165 |siehst du das NUR,
 mu-gs ->|tippt mehrfach--->
 166 weil (.) das (.) zu viel |kultUR is,=
 mu-gz ----->|blickt zu TO (>Z. 171)
 mu-gs |hält rIF auf Blatt--->



```

167      =oder was |würd dich da NOCH |dran stören;
      mu-gs ----->|tippt silbisch----|hält rIF auf Blatt (>Z. 178)
168      (2.5)
169      TO:   mh::;
170      s_sind ja nur drei TAge;
171      | (0.4) | (0.5)
      mu-he |nickt----->
      mu-fa |mouth shrug->
      mu-gz ----->|blickt auf Blatt (>Z. 190)
172      TO:   das is: | (.) von allen ja das WEnigs|te;-
      mu-he ----->|
      mu-fa ----->|
173      (0.3)
174      TO:   das:-| (-)
      mu-he ---->|
175      <<kopfschüttelnd> also ZWEI gar nich;>
176      weil das sind nur drei TAge,
177      un[d-      ]

```



```

178      MU:      [ge|NAU;]                                |
      mu-gs      -->|schiebt rIF an andere Stelle der Option-|hält->
179      das WEnigste;
180      und,
181      (2.1)

```



182	TO:	kultUR	schock.	
	to-gs	öffnet rH Palm Up		
	mu-gs		fährt mit rIF an den Anfang der Option	
183	TO:	ROM,		
	mu-gs	,,,		
184		<<pp>>ja was IS denn mit rom;>		
185	MU:	<<nickend>>ist auch am wei testen entFERNT;=ne,>		
	mu-gz	-----> blickt zu TO----->>		

Dem Transkriptausschnitt vorangehend hat die Tochter einen Einwand gegen den Segeltörn eingebracht. Die Mutter ratifiziert ihre Begründung (Z. 159) und bringt nun ihrerseits die Option ‚Rom‘ (Z. 161) ein.

Dabei richtet sie eine Rückfrage (Z. 165ff.) an die Tochter, die zuvor Rom mit Verweis auf das Kulturprogramm ausgeschlossen hat. Sprachlich markiert die Sprecherin durch das Adverb *nur* (Z. 165) sowie die akzentuierte Partikel „NOCH“ (Z. 167), dass neben dem Kulturprogramm noch weitere Begründungen gegen Rom sprechen – und die Sprecherin diese bereits antizipiert. Die bereits frühzeitig eingebrachte deiktische Geste auf das Material (Z. 160) hält die Mutter turnfinal (Z. 167) und wendet der Tochter als Adressatin der Frage den Blick zu. Der Gesprächsgegenstand bleibt somit auch innerhalb der Begründungssequenz der Tochter präsent. Diese bezieht sich sodann auf die präsent gehaltene Option (Z. 169-177). Die Begründung wird seitens der Mutter ratifiziert (Z. 178: *geNAU*;) und zeitgleich mit der Ratifizierung verändert die Mutter nun die Position des Indexfingers: Sie wiederholt einen Teil der Begründung der Tochter (Z. 179: *das WEnigste*;) und stellt eine Ergänzungsfrage (Z. 180: *und*,). Es kommt zu einer langen Pause (Z. 181) und die Tochter bringt nun „*kulTURschock*“ (Z. 182) als potenzielle Begründung ein. Die Mutter fährt mit dem Indexfinger überlappend zum Turnende der Tochter über die Option auf dem Material, und zwar an die Stelle, an der die Tage der jeweiligen Ziele notiert sind. Durch die Geste sowie durch die ausbleibende Ratifizierung und/oder Turnübernahme macht sie kenntlich, dass der Zugzwang noch nicht ausreichend bearbeitet worden ist, und ein Begründungselement, auf das sie gestisch verweist, noch fehlt. Die Tochter, die ebenfalls auf das Material hin orientiert ist, markiert nun deutlich ein Verstehensdefizit, das sich erst in der Rückfrage (Z. 183: *ROM*,) und dann in einer *denn*-Frage (Z. 184: <<pp> ja was IS denn mit rom;>) äußert (vgl. Deppermann 2009; Heller 2012). Dadurch wird die Mutter verpflichtet, das Verstehensdefizit zu bearbeiten. Mit Anzeige der ersten Rückfrage löst die Mutter die Geste auf und zeigt sich nun verantwortlich für das Aufklären des Verstehensdefizits. Sie erbringt eine Begründung unter Berücksichtigung der Entfernung des Ziels. Mit der Verschiebung der Geste hat die Sprecherin einen Bezug zur Dauer der Abschlussfahrt herstellen wollen, der sie nun die Entfernung gegenüberstellt (Z. 185: *ist auch weitesten entFERNT*;=ne,).

Dieses Beispiel zeigt, dass die Mutter hier durch das anfängliche Format der Ergänzungsfrage über weitere Begründungen gegen ‚Rom‘ verfügt, die allerdings von der Tochter eingebracht werden sollen. Fehlen in der Rückmeldung des Gegenübers dann z. B. spezifische inhaltliche Aspekte, können diese durch das Verschieben der deiktischen Geste auf andere Referenten versucht werden nachzusteuern – wie das vorliegende Beispiel zeigt, gelingt dies nicht notwendigerweise.

8.5.3 Zusammenfassung

In diesem Kapitel zeigt sich, dass Sprecher*innen durch den Einsatz gestischer Ressourcen (Zählgesten und deiktischer Gesten) sowohl auf das Antwortformat, i. S. einer Aufzählung, als auch den inhaltlichen Gesprächsgegenstand Einfluss nehmen können. Dabei wird sichtbar, dass die Gesten über den sprecher*innenseitigen *Turn* hinaus eine interaktive, steuernde Funktion übernehmen können. Diese Art der Steuerung lässt sich in den vorliegenden Daten i. d. R. ausschließlich seitens der erwachsenen Sprecher*innen beobachten. Dass in Erwachsenen-Kind-Argumentationen Eltern eine steuernde Funktion übernehmen, zeigen beispielsweise auch Quasthoff und Kluger (2021: 148) für die Bearbeitung der ein- und ausleitenden diskursiven Aufgaben, die besonders häufig von Eltern übernommen werden. An dieser Stelle werden also spezifische Kontextbedingungen sichtbar, die vermutlich auf die sozialen bzw. familialen Rollen der Interagierenden zurückzuführen sind (vgl. auch Kap. 2.1).

Folgende Befunde für die argumentativen Handlungen des Einforderns von Rückmeldungen können zusammengefasst werden:

- Die multimodale Gestalt des Frageformats kann Aufschluss über das erwartete *Antwortformat* liefern (vgl. Kap. 8.5.1): So können die Fragenden durch das Produzieren von Zählgesten deutlich für ihr Gegenüber erkennbar machen, dass in der Bearbeitung der eingeforderten Rückmeldung mehrere Inhaltselemente, i. S. einer Liste, erwartet werden. Das anschließende Halten der Geste über den Sprecher*innenwechsel hinweg macht die eingeführte Zählgeste nun auch für die Rückmeldung des Gegenübers relevant. In der Rückmeldung des Gegenübers kann nun die Zählgeste fortgesetzt werden und dadurch sowohl anzeigen, dass das Listenformat erfüllt wird, als auch die eingebrachten Listenelemente ratifizieren. Wird die Zählgeste in der Rückmeldung des Gegenübers nicht fortgesetzt, kann dies darauf hindeuten, dass die Antwort des Gegenübers nicht den Erwartungen der Rückmeldung initiiierenden Person entspricht.
- Innerhalb von eingeforderten Rückmeldungen können deiktische Gesten auf das Material einen Gesprächsgegenstand turnfinal und somit auch während der Rückmeldung des Gegenübers präsent halten (vgl. 8.5.2). Verändern Gestenproduzent*innen innerhalb der Antwort des Gegenübers die deiktische Geste und bewegen sie auf einen anderen Referenten, kann dadurch versucht werden, eine Antwort inhaltlich nachzusteuern. Die Geste übernimmt innerhalb der multimodalen Gestalt eine gesprächsthemensteuernde Funktion.

8.6 Markieren epistemischer Modalisierung

Die vorangegangenen Kapitel haben gezeigt, wie multimodale Ressourcen zur Realisierung jeweils spezifischer argumentativer Handlungen beitragen. In dem vorliegenden Kapitel wird nun die Ebene der handlungsbezogenen multimodalen Gestalten verlassen und es werden *handlungsübergreifende* multimodale Gestalten vorgestellt. Im Fokus stehen Formen der epistemischen Modalisierung. Diese übernehmen für das Argumentieren insofern eine relevante Funktion, da sie z. B. Aufschluss über die Verhandelbarkeit einer Position oder die Gewissheit bzw. Ungewissheit des eingebrachten Wissens eines*einer Sprecher*in geben können. Da die Ebene der epistemischen Modalisierung für das mündliche Argumentieren – und dessen Erwerb (vgl. Kap. 10) – also von Bedeutung ist, wird sie hier separat von den handlungsbezogenen multimodalen Gestalten vorgestellt. Der Fokus der Untersuchung bleibt, wie in der vorangegangenen Untersuchung, auf die Gesten ausgerichtet.

Sprecher*innen können ihre Äußerung epistemisch modalisieren und somit in mündlichen Argumentationen die eigene epistemische Haltung, i. S. des *epistemic stance* (vgl. Heritage 2012a, 2012b; Heritage 2013), anzeigen. Die epistemische Haltung der Sprecher*innen kann dabei mit Bezug auf das (angenommene) Wissen des Gegenübers als unsicher (K–) oder sicher (K+) markiert werden (vgl. Heritage 2012b). Je nach epistemischer Modalisierung können dann Äußerungen als Vorschlag und somit als verhandelbar oder aber als informierender Hinweis und somit als feststehend bzw. gesetzt in die Interaktion eingebracht werden (vgl. Sidnell 2011; Stevanovic 2012, 2013; Stevanovic und Peräkylä 2012; Heller 2018a, 2021b; Morek 2020). Dabei wird mittels des *epistemic stance* epistemische Verantwortung (vgl. Stivers et al. 2011) beansprucht, mit der sich die Interagierenden beispielsweise in argumentativen Aushandlungen gegenseitig anzeigen, ob sie die Entscheidungsfindung als Prozess einer *gemeinsamen Aushandlung* forcieren oder einzelne Beteiligte sich als *allein verantwortlich* für die Entscheidungsfindung verstehen (vgl. Stivers et al. 2011; Heller 2018a; Weiser-Zurmühlen 2021). Epistemische Modalisierungen können dabei durch unterschiedliche multimodale Ressourcen vorgenommen werden und dazu beitragen, ein Argument bzw. den Inhalt eines Arguments als un-/sicher bzw. un-/verhandelbar

zu markieren: So zeigen linguistische Untersuchungen, dass epistemisch abschwächende Modalisierungen, wie Satzadverbien (*vielleicht, eventuell*), Konjunktive sowie *verba dicendi* + *würde* (*ich würde sagen*), dazu beitragen, ein Argument als vorläufig und vorsichtig an eine Entscheidung herantastend einzubringen (vgl. Heller 2018a; Morek 2020) – und somit auch eine gemeinsame epistemische Verantwortung zu etablieren. Gemeinsam mit gestischen Ressourcen können die Argumente außerdem auf der geöffneten Handfläche zur weiteren Aushandlung dem Gegenüber angeboten werden (vgl. Heller 2018a; vgl. auch Kap. 8.2.1.1).

Interagierende können aber auch die Inhalte ihrer Äußerungen als evident oder offensichtlich markieren. Satzadverbien wie *bestimmt* oder *sicher* sowie die Modalpartikel *ja* tragen auf verbalsprachlicher Ebene zur Evidenzmarkierung bei. Durch die Modalpartikel *ja* kann z. B. das eingebrachte Wissen als *geteiltes* Wissen markiert und „strittige Behauptungen [...] [ge]stützt[t]“ werden – „unabhängig von tatsächlicher Bekanntheit“ (Reineke 2016: 117) des dargelegten Wissens. Es handelt sich um argumentativ eingesetzte Evidenzmarkierungen. Weitere Formen der Evidenzmarkierungen zeigen sich beispielsweise auch in körperlich-sichtbaren Ressourcen, wie *Palm Up* Gesten mit lateraler Bewegung (vgl. u. a. Kendon 2004; Jehoul et al. 2017; Cooperrider et al. 2018; Borràs-Comes et al. 2019; vgl. zur Markierung von Unsicherheit auch Debras 2017, 2021; Borràs-Comes et al. 2019).

Das vorliegende Kapitel legt nun dar, wie Interagierende vornehmlich in der argumentativen Handlung des Begründens mit multimodalen Ressourcen eine Äußerung oder einen spezifischen Gesprächsgegenstand als fraglos evident markieren (vgl. Kap. 8.6.1), die alleinige epistemische Autorität übernehmen, um einen Gesprächsgegenstand in ablehnender Haltung aus der weiteren Interaktion auszuschließen (vgl. Kap. 8.6.2), oder sich gegenseitig Nicht-Wissen bzw. epistemische Unsicherheit (vgl. Kap. 8.6.3) sowie inhaltliche Unbestimmtheit (vgl. Kap. 8.6.4) anzeigen können.

8.6.1 Markieren von Evidenz

Das Markieren von Evidenz, d. h. das Darstellen eines Sachverhalts als offensichtlich oder fraglos, kann in mündlichen Argumentationen den Geltungsanspruch eines Arguments erhöhen. So kann dem Gegenüber ein Wissen zugeschrieben oder unterstellt werden – unabhängig von dem tatsächlichen Wissensbestand –, sodass neues Wissen zugleich als bekannt und somit als nicht weiter aushandlungsbedürftig markiert werden kann (vgl. Reineke 2016). Evidenzmarkierungen können dabei sowohl lexikalisch, z. B. durch den Einsatz von Modal- bzw. Abtönungspartikeln (vgl. u. a. Reineke 2016, 2018; Schoonjans 2018), gestisch, z. B. in Form von *Palm Up* Gesten z. T. mit lateraler Bewegung innerhalb der eigenen Äußerung (vgl. u. a. Kendon 2004; Cooperrider et al. 2018), sowie in Kombination mit *Shrug*-Komponenten wie dem Schulterzucken oder Abknicken des Kopfes (vgl. Debras 2017) erfolgen. Das vorliegende Kapitel legt dar, wie Interagierende im Ressourcenverbund durch das Fallenlassen der Hand bzw. Hände (vgl. Kap. 8.6.1.1), durch *Palm Up* Gesten mit lateraler Bewegung (vgl. Kap. 8.6.1.2) sowie durch das kurzzeitige Aufdrehen der Hand als *Hand Shrug* (vgl. 8.6.1.3) Evidenzmarkierungen in multimodalen Gestalten realisieren.

8.6.1.1 Markieren von Evidenz durch Fallenlassen der Hand

Interagierende können eigene Äußerungen, und im Kontext von mündlichen Argumentationen ihre eigene Begründung oder einen Einwand, als offensichtlich oder evident markieren. Dazu nutzen sie auch gestische Ressourcen. Wie die Untersuchung zeigt, wird dazu das hörbare Fallenlassen der Hand, das bei der Rückführung der Geste in eine Ruheposition entsteht, als gestischer sowie akustischer Kommentar der aktuellen Äußerung eingesetzt: Gemeinsam mit epistemisch nicht abgeschwächten Äußerungen im Indikativ sowie z. T. sprachlichen Evidenzmarkierungen markiert das akustische Signal, das aus der

Rückführung der Geste in die Ruheposition durch das Fallenlassen der Hand resultiert, den Inhalt der aktuellen Äußerung nachdrücklich als abgesichert bzw. evident (vgl. Tab. 27).

Markieren von Evidenz durch Fallenlassen der Hand	
Beschreibung der multimodalen Gestalt	epistemisch nicht abgeschwächte Äußerung + akustisches Signal einer zurückgeführten, fallengelassenen Geste (+ optional sprachliche Evidenzmarkierungen)

Tabelle 27: Beschreibung der multimodalen Gestalt „Markieren von Evidenz durch Fallenlassen der Hand“

Anhand des nachfolgenden Beispiels 1 wird die oben beschriebene Form der Evidenzmarkierung illustriert. In diesem Beispiel schließt die Mutter ihre Begründung mit einer Konklusion ab, die sie sowohl sprachlich sowie durch das geräuschvolle Fallenlassen als evident markiert.

Beispiel 1 – Markieren von Evidenz durch Fallenlassen der Hand

071	MU:	vier tage BUNgalowpark,
072		(5.0)
073	TO:	(der) klingt AU_nich schLEcht.
074		(2.0)
075	MU:	s_sind VIEle möglichkeiten (.) sicherlich dA.
076		und dann n haben viele l äh:m es sind ja viele LEUTE,=
077		=die haben alle unterschiedliche interESSen,
078		(1.3)




079	mu-gs	und sicherlich kann man da ganz	viele
		 tippt mit rMF auf Option ,,,,,	




	mu-gs	mit	°h (.) zuFRIEdenstellen;=ne,
080		lässt rH geräuschvoll fallen	
		(1.8)	

In der begründenden Elaborierung (Z. 075-079) der Bewertung der Tochter (Z. 073) formuliert die Mutter abschließend eine Konklusion, die sie durch die Verwendung des Indikativs, des zweimalig verwendeten Satzadverbs *sicherlich* (Z. 075, 079) und der Modalpartikel *ja* (Z. 076) als evident markiert. Das geräuschvolle Fallenlassen der Hand, das in der Rückführung einer deiktischen Geste in die Ruheposition entsteht und in einem akustischen Signal resultiert, verleiht der Satzaussage weiteren Nachdruck: Im gemeinsamen Verbund der multimodalen Ressourcen wird der Inhalt der Konklusion als evident

markiert und dem Gegenüber somit unterstellt, dass ihm der Inhalt ebenfalls bekannt sein müsste. Die Gestenfunktion ist dabei der *Beat* Geste (vgl. Kap. 3.2) nicht unähnlich, mit der Sprecher*innen „Argumenten Nachdruck [...] verleihen oder die Illokution [...] verstärken“ (Schoonjans 2018: 160) können. In der vorliegenden Sequenz ist es nun das aus der Rückführung der Geste resultierende Geräusch, das die Äußerung hervorhebt und letztlich das ausführt, wonach die *Beat* Gesten benannt sind: Sie realisieren ein Taktgeräusch oder auch Klopfen, „not for its own semantic content, but for its discourse-pragmatic content“, wie McNeill (1992: 15) für *Beat* Gesten resümiert. Das turninterne Fallenlassen der Hand wird mit sprachlichen Evidenzmarkierungen, aber unabhängig von fokusakzentuierten Lexemen (vgl. hierzu auch Ambrazaitis und House 2017) koordiniert: Die Geste übernimmt die Funktion, den Inhalt der Satzaussage als bekannt zu unterstellen. Die Funktion multimodalen Gestalt erstreckt sich hier über die gesamte Äußerung.

8.6.1.2 Markieren von Evidenz durch *Palm Up* Gesten mit lateraler Bewegung

Sprecher*innen können durch unterschiedliche Ressourcen, wie beispielsweise den Einsatz sprachlicher Modalisierungen, wie Partikeln (z. B. *ja*, *doch*, *wohl*) oder Satzadverbien (z. B. *bestimmt*; vgl. u. a. Reineke 2016, 2018; Schoonjans 2018; Morek 2020), körperlich-sichtbare Ressourcen, wie intensivierendes Kopfschütteln (vgl. McClave 2000; Kendon 2002), *Palm Up* sowie *Palm Up* Gesten mit seitlicher Bewegung der Hand bzw. Unterarme (vgl. Kendon 2004; Cooperrider et al. 2018; Marrese et al. 2021),⁸⁹ oder *Shrug*-Komponenten, z. B. Schulterzucken, Anheben der Augenbrauen etc. (vgl. Streeck 2009b; Debras 2017), den Inhalt einer Äußerung als offensichtlich oder evident markieren.

In den vorliegenden argumentativen Entscheidungsfindungen nutzen die Sprecher*innen *Palm Up* Gesten mit lateraler Bewegung sowie Modalpartikeln oder Satzadverbien, um sich gegenseitig die Offensichtlichkeit bzw. fraglose Evidenz ihres Arguments anzuzeigen. Durch die multimodale Gestalt wird das zugrundeliegende Wissen des eingebrachten Arguments als fraglos evident gerahmt und als *gemeinsames* Wissen (i. S. eines *Common Ground*, vgl. Clark 1996) für die Interaktion relevant gemacht (vgl. Reineke 2016). Dadurch wird auch dem Gegenüber die epistemische Haltung K+ hinsichtlich des eingebrachten Wissens zugewiesen und der Geltungsanspruch gegenüber dem präsentierten Sachverhalt verstärkt (vgl. Reineke 2016: 119f.). Dabei geschieht zweierlei: Zum einen wird dem Gegenüber auch eine wissende Haltung (K+) zugeschrieben – unabhängig davon, ob er*sie tatsächlich Kenntnis über das eingebrachte Wissen besitzt –, zum anderen wird dadurch eine Akzeptanz gegenüber dem als fraglos evident markierten Sachverhalt vorausgesetzt (vgl. Thurmair 1989; Reineke 2016). In argumentativen Kontexten kann ein solches Vorgehen hochgradig funktional sein, wenn potenziell strittige Argumente eingebracht und in ihrem Geltungsanspruch verstärkt werden sollen (vgl. Reineke 2016: 118ff.). Eine Beschreibung der multimodalen Gestalt sowie ihrer Funktion findet sich in der nachfolgenden Übersicht 28.

Markieren von Evidenz durch <i>Palm Up</i> Gesten mit lateraler Bewegung	
Beschreibung der multimodalen Gestalt	<i>Palm Up</i> Geste mit lateraler Bewegung + epistemisch verstärkende Modalisierung (z. B. Modalpartikeln, Satzadverbien)

Tabelle 28: Beschreibung der multimodalen Gestalt „Markieren von Evidenz durch *Palm Up* Geste mit lateraler Bewegung“

⁸⁹ *Palm Up* Gesten mit lateraler Bewegung können – abhängig vom sequenziellen Kontext sowie weiteren eingesetzten Ressourcen – sowohl epistemische Sicherheit als auch Unsicherheit markieren (vgl. hierzu zsf. Cooperrider et al. 2018). Eine Erklärung für diese scheinbare Gegensätzlichkeit könnte darin liegen, dass der Äußerung nichts mehr hinzuzufügen ist, wenn Sprecher*innen ausdrücken, dass etwas offensichtlich oder evident ist – „I don’t know what else I could say about it, it is obvious“ (Jehoul et al. 2017: 7). Nicht-Wissen und Evidenzmarkierungen liegen hier nah beieinander (vgl. Cooperrider et al. 2018).

Im nachfolgenden Beispiel stützt der Sohn eine erste, von der Mutter als strittig angezeigte und zurückgewiesene Behauptung durch eine weitere Begründung. Dabei markiert er den Begründungsinhalt sowohl sprachlich durch den Einsatz der Modalpartikel *ja* als auch gestisch durch eine *Palm Up* Geste mit lateraler Bewegung als evident.

Beispiel 1 – Markieren von Evidenz durch *Palm Up* Gesten mit lateraler Bewegung

197 SO: da kriegst du auch die ganze zeit also schnell LANGeweile;=
 198 =wenn du da nur auf dem MEER bist,
 199 MU: (-) <<p>vielleicht,>=
 200 SO: =[klar SCHWIMmen aber]-
 201 MU: [ja gu:t mit so viel] leuten eher NICHT,=
 202 =aber-
 203 SO: ja;



204 du |siehst ja |nur |MEER|
 so-gs |öffnet rH Palm Up, laterale Bewegung|,,,|
 so-gz ----->|blickt zu MU----->
 mu-gz >>zu SO----->
 wenn du (dich |[irgendwo] |umdrehst);|
 so-gz -----|
 so-gs |Beat rH---|
 mu-gz -----|blickt auf Tisch----->>
 205 MU: [ja-]
 206 (4.0)
 207 MU: segeln hört sich GUT an;=
 208 =aber (-) ich FINDS auch nicht so Ungefährlich;
 209 von DAher-

Nachdem gegen die erste Behauptung des Sprechers (Z. 197: da kriegst du auch die ganze zeit also schnell LANGeweile;) seitens der Mutter ein Einwand vorgebracht worden ist (Z. 201: ja gu:t mit so viel leuten eher NICHT),⁹⁰ stützt der Sprecher seine Behauptung im Folgenden durch eine Begründung (Z. 203f.: ja; / du siehst ja nur MEER wenn du (dich irgendwo umdrehst);). Er schließt seine Äußerung dabei durch die turninitiale Rezeptionsbekundung an die Äußerung der Mutter asyndetisch an und markiert sie als responsiv zu dieser (vgl. Reineke 2016).⁹¹ Die Modalpartikel *ja* wird in der stützenden Begründung nun argumentativ eingesetzt, um den Begründungsinhalt als offensichtlich bzw. fraglos evident zu markieren. Dabei positioniert sich der Sohn selbst als Wissender (K+), sodass er, auch wenn seine Äußerung an zweiter Position, d. h. responsiv erfolgt, Anspruch auf epistemische Priorität bzw. Anspruch auf primäre epistemische Rechte erhebt (vgl. Reineke 2016: 105; vgl. auch Stivers et al. 2011). Indem der Sachverhalt der Begründung aber insgesamt als bekannt angezeigt wird, wird dieses (geteilte)

⁹⁰ Innerhalb einer konzessiven Struktur stimmt die Mutter der Äußerung des Sohnes vorerst schwach zu (vgl. Z. 199), wendet dann allerdings ein (Z. 201); eine anschließende Rückkehr zu einer schwachen Zustimmung wird durch das *aber* (Z. 202) allerdings nur angedeutet und dann abgebrochen (vgl. zu den multimodalen Ressourcen des schwachen Zustimmens in Konzessionen auch Kap. 8.3.2).

⁹¹ In asyndetisch angeschlossenen Begründungen können, so „stellt Gohl fest, [...] Modalspartikeln als ‚Hinweis auf die kohäsive Verbindung zwischen zwei (oder mehreren) Äußerungen‘ (2006, S. 110) dienen“ (Reineke 2016: 120).

Wissen und somit die wissende Haltung gegenüber dem Sachverhalt ebenfalls für die Mutter geltend gemacht (vgl. Reineke 2016: 119f.) – der Geltungsanspruch der zuvor strittigen Behauptung wird dadurch innerhalb der vorliegenden Begründung ge- bzw. verstärkt. Zeitlich koordiniert wird die Äußerung mit einer *Palm Up* Geste mit lateraler Bewegung (Z. 104, vgl. #1-#4), die noch vor der Modalisierung einsetzt. Die multimodalen Ressourcen markieren hier gemeinsam den eingebrachten Sachverhalt als „obvious“, sodass die Geste „serves as a comment on the speaker’s own speech“ (Kendon 2004: 277). Der Sprecher wendet in seiner Begründung unmittelbar nach dem Einsatz der Modalpartikel und noch während er die *Palm Up* Geste nach außen bewegt seiner Mutter den Blick zu (vgl. #4). Durch die Blickadressierung lädt er sie zur Zustimmung gegenüber dem als fraglos evident markierten und die Begründung stärkenden Wissen ein. Der Blick wird nach überlappender Affirmation der Mutter (Z. 205) dann in Richtung des Tisches abgewandt (vgl. Rossano 2012). Die Zustimmung der Mutter wird anschließend in Form einer Konzession fortgeführt (X, Y, X‘; Couper-Kuhlen und Thompson 2000; vgl. Kap. 2.2.2), in der die elliptische Zustimmung turnfinal prominent gesetzt wird (vgl. Z. 208).

Insgesamt kann die multimodale Gestalt der Evidenzmarkierung durch *Palm Up* Gesten mit lateraler Bewegung in mündlichen Argumentationen dazu beitragen, das eingebrachte Wissen der Sprecher*innen in (potenziell strittigen) Begründungen bzw. Sachverhalt als fraglos evident zu markieren. Damit wird zugleich dem Gegenüber dieses Wissen zugänglich gemacht, sodass der*die Rezipient*in in Folge dieses Vorgehens auch als Wissende*r adressiert werden kann. Das Beispiel konnte aufzeigen, wie der Sprecher auf der Grundlage des ‚gemeinsamen‘ Wissens für Akzeptanz des Gegenübers in Bezug auf das eingebrachte Wissen wirbt und somit auch den Geltungsanspruch des Inhalts seiner Begründung verstärkt. Im Verbund der multimodalen Ressourcen wird die epistemische Haltung eines*einer Sprecher*in gegenüber dem Sachverhalt durch die *Palm Up* Geste mit lateraler Bewegung auch visuell wahrnehmbar angezeigt. Die Gestenausführung erfolgt dabei in der Turnkonstruktionseinheit, in der auch die Modalpartikel geäußert wird. Die gestische Ressource weist dabei, wie die Modalpartikeln selbst, einen „Skopus über die ganze Äußerungseinheit oder gar die ganze Äußerung“ (Schoonjans 2018: 248) auf, sodass sich die multimodale Evidenzmarkierung im vorliegenden Beispiel über die gesamte Turnkonstruktionseinheit erstreckt.

8.6.1.3 Markieren von Evidenz durch Hand Shrug

In mündlichen Argumentationen können Sprecher*innen multimodale Ressourcen wie *Shrug*-Komponenten, u. a. das Hochziehen der Schultern oder Augenbrauen (vgl. u. a. Streeck 2009b), einsetzen, um den Sachverhalt ihrer Äußerung als fraglos evident anzuzeigen. In der vorliegenden Studie nutzen die Sprecher*innen nicht nur *Palm Up* Gesten mit lateraler Bewegung (vgl. Kap. 8.6.1.2), sondern auch das kurzzeitige Aufdrehen der Hände im Kontext von Evidenzmarkierungen: Eine Hand oder beide Hände des*der Sprecher*in wird/werden zumeist über die Handkante nach oben gedreht. Es handelt sich um eine eher kleinräumige Gestenausführung, die nicht im Gestenhöhepunkt gehalten wird. Die Geste dient als kurzer Kommentar zur sprachlichen Äußerung des*der Sprecher*in selbst und gibt Aufschluss über seine*ihre epistemische Haltung. Wie die nachfolgende Übersicht 29 darlegt, wird die Geste im Verbund mit weiteren Ressourcen wie dem Schulterzucken oder intensivierendem Kopfschütteln verwendet, um den Äußerungsinhalt als *offensichtlich* bzw. *fraglos evident* zu markieren. Die Geste wird, aufgrund ihrer raschen Ausführung des Aufdrehens der Hand in Anlehnung an Ekman und Friesen (1972: 367) als *Hand Shrug* verstanden.⁹² Mit ihr zeigt der*die Sprecher*in nicht nur seine*ihre eigene epistemische Haltung gegenüber dem dargelegten Sachverhalt an, sondern diese eingenommene Haltung wird,

⁹² Ekman und Friesen (1972: 367) beobachten, dass *Hand Shrugs* als Embleme die Bedeutung von „helplessness or inability“ tragen. In der vorliegenden Studie wird die Bezeichnung gewählt, da der Begriff des *Shrugs* oder

indem der Gesprächsgegenstand als offensichtlich markiert wird, zugleich auch dem Gegenüber zugeschrieben (vgl. ähnlich dazu Kap. 8.6.1.2; vgl. Reineke 2016). Eine Übersicht über die beteiligten Ressourcen sowie die Funktion gibt nachfolgend Tabelle 29.

Markieren von Evidenz durch <i>Hand Shrug</i>	
Beschreibung der multimodalen Gestalt	<i>Hand Shrug</i> + weitere <i>Shrug</i>-Komponenten (z. B. Schulterzucken, Abknicken des Kopfes, Hochziehen der Augenbrauen, <i>Mouth Shrug</i>)

Tabelle 29: Beschreibung der multimodalen Gestalt „Markieren von Evidenz durch *Hand Shrug*“

Das folgende Beispiel illustriert, wie eine Sprecherin einen *Hand Shrug* in einem größeren Ressourcenverbund einsetzt, um eine Konklusion einer Begründung als offensichtlich bzw. fraglos evident zu markieren.

Beispiel 1 – Markieren von Evidenz durch *Hand Shrug*

186	MU:	die FRAGE [is-]	
187	TO:	[ja;]	
188	MU:	was HAT man von diesen drei tagen;=	
189		=UND °h ähm:-	
190		man sitzt halt VIEL im bus drin==	
191		=und m im GRUNde genommen hat die geMEINschaft (.)	
			
		die KLASse (.) nich	ts daVON; (-)
mu-gs		dreht rH vertikal auf- ,,	
mu-he		knickt Kopf nach re ab	
mu-bo		Shoulder Shrug-----	
192		°h ALso-	
mu-gs		zeigt mit rIF & MF auf Blatt-->	
193		oKAY.	
mu-gs		---->	
194		schlie ßen wa ma mal d2 (-) DIEse beiden aus; (-)	
mu-gs		-----> hebt Fi an----- zeigt mit rIF auf Option	

Die Mutter entfaltet in der vorliegenden Sequenz eine Begründung (Z. 189f.) gegen die Fahrt nach Rom und macht in einer abschließenden Konklusion (Z. 190f.) deutlich, dass bei einer so kurzen Fahrt der Mehrwert für die Klassengemeinschaft nicht gegeben sei. In dem Argument, das der Konklusion vorangestellt wird (Z. 190: man sitzt halt VIEL im bus drin-), wird durch *halt* das präsentierte Wissen als beiden Interagierenden bekannt unterstellt. Es schließt die Konklusion an, die mit einer formelhaften Phrase eingeleitet wird (Z. 191: und m im GRUNde genommen). In deren Fortsetzung zieht die Sprecherin nun die Schultern in der Minimalpause hoch, knickt den Kopf kurzzeitig ab und dreht gemeinsam mit dem Negationspronomen *nichts* die rechte Hand kurzzeitig über die Handkante auf, sodass die Handfläche auf die Sprecherin selbst orientiert ist (vgl. #1). Die Hand wird unmittelbar anschließend in eine Ruheposition zurückgeführt. Gemeinsam konstituieren die Ressourcen einen komplexeren *Shrug* (vgl.

auch ‚Zuckens‘ eine rasche(re) Ausführung der Geste impliziert. Vgl. zu der Bedeutung bzw. Funktion, die Ekman und Friesen (1972) der Geste zum Markieren von epistemischer Unsicherheit zuweisen auch Kap. 8.6.3.

Streeck 2009b; Debras 2017), mit dem das Lexem *nichts*, das auf eine Uneingeschränktheit des Äußerungsinhaltes verweist, besonders hervorgehoben wird. Zugleich wird der Äußerungsinhalt, dass die Schüler*innen von einer Klassenfahrt, bei der sie die meiste Zeit im Bus verbringen (Z. 190), nichts (Z. 191) haben, als offensichtlich markiert. Das dargelegte Argument samt Konklusion wird als – für beide Beteiligten – fraglos evident markiert. Dadurch schafft die Sprecherin einen *Common Ground* (vgl. Clark 1996), von dem aus sie nun die Interaktion fortsetzt und den Ausschluss dieser sowie einer weiteren Option beschließt (Z. 194).

Die multimodale Gestalt, in der ein *Hand Shrug* im Verbund mit weiteren *Shrug*-Komponenten, wie z. B. einem Schulterzucken, geäußert wird, dient also insgesamt der Markierung von Evidenz in einer Äußerung.

8.6.2 Markieren des ablehnenden Fallenlassens eines Gesprächsgegenstandes

Wenn Interagierende ihre Beiträge gemeinsam mit einer *Dropping of Hand* Geste (Bressem und Müller 2014a: 1584) äußern, bei der die Hand lose angehoben und dann fallengelassen wird, sodass i. d. R. ein akustisch wahrnehmbares Geräusch entsteht, dann rahmen die Beteiligten, so Bressem und Müller (2014a: 1584), Anteile einer Äußerung als nicht (weiter) relevant. Die Geste dient somit dem Ausschluss eines Themas aus der weiteren Interaktion.

In den vorliegenden Daten wird die Geste turninitial und responsiv zur vorangegangenen Äußerung produziert. Der*die aktuelle Sprecher*in verhandelt dabei den Gesprächsgegenstand der Vorgängerausäußerung weiter. Allerdings zeigt der*die Sprecher*in durch die multimodale Gestalt an, dass der Gesprächsgegenstand anschließend fallengelassen und nicht weiter von Relevanz sein wird. Das Fallenlassen der Option obliegt also der epistemischen Autorität des*der Sprecher*in. Durch die Art der Gestenausäußerung – indem die angespannte und nach oben geöffnete Hand (vgl. aber Bressem und Müller 2014a: 1584) lose fallengelassen wird – fungieren die gestischen Ressourcen auch als Teil einer Anzeige der Haltung der Sprecher*innen: Die Sprecher*innen halten nicht, wie bei den *Away* Gesten, ‚störende‘ oder ‚unliebsame‘ Gesprächsinhalte (vgl. Bressem und Müller 2014b) von sich fern. Vielmehr wird durch das unkontrollierte Fallenlassen der losen Hand eine Art Resignation bzw. Widerwillen und somit eine *ablehnende Haltung* des*der Sprecher*in gegenüber dem Gegenstand bzw. dem Inhalt der Vorgängerausäußerung für alle Beteiligten sichtbar markiert (vgl. auch Mondada et al. 2015: 67f.). In den vorliegenden Daten projizieren Sprecher*innen durch die multimodale Gestalt das Fallenlassen dispräferierter Gesprächsgegenstände. Die Handgeste der *Dropping of Hand* dient dabei sowohl der visuell als auch akustisch wahrnehmbaren Markierung von Dispräferenz (vgl. Pomerantz 1984) – sowohl gegenüber dem Gesprächsgegenstand als auch der weiteren Aushandlung.⁹³ I. d. R. ist der Blick während der Zurückweisung vom Gegenüber abgewandt (vgl. Pekarek Doehler et al. 2021; vgl. Tab. 30). Der*die Sprecher*in übernimmt hier in ablehnender Haltung auch die epistemische Autorität (vgl. Stivers et al. 2011) für das Fallenlassen des Gesprächsgegenstandes.

⁹³ Efron (1941/1972: 194) bildet diese Geste in seiner Studie ebenfalls ab und beschreibt die ihr zukommende Funktion bzw. Bedeutung (häufig auch in Kombination mit einem Schulterzucken) als „expression of doubt or rejection“. Obwohl in der Studie von Efron (1941/1972) der Kontext der Gestenausführung fehlt, erscheint interessant, dass er in seiner Beschreibung sowohl auf die Haltung des*der Sprecher*in als auch auf die performative Funktion der Geste Bezug nimmt.

Markieren des ablehnenden Fallenlassens eines Gesprächsgegenstandes	
Beschreibung der multimodalen Gestalt	responsive turninitiale <i>Dropping of Hand</i> Geste (i. d. R. + abgewandter Blick)

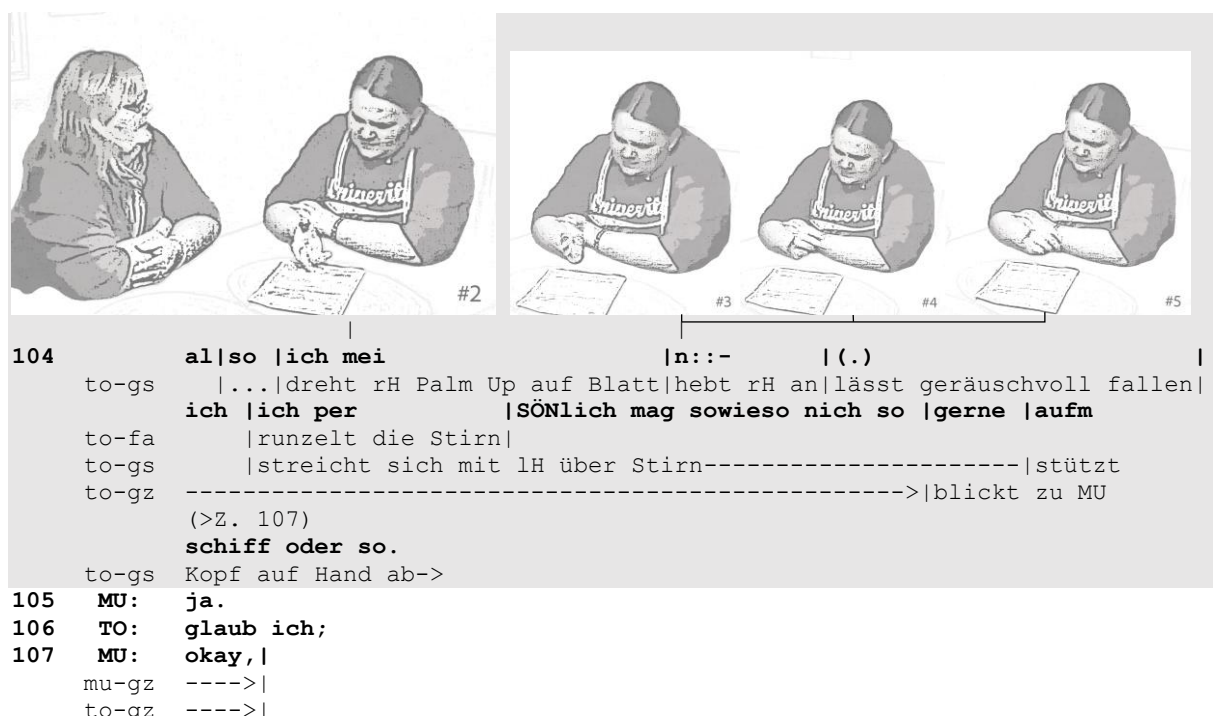
Tabelle 30: Beschreibung der multimodalen Gestalt „Markieren des ablehnenden Fallenlassens eines Gesprächsgegenstandes“

Die nachfolgenden Beispiele illustrieren, wie Sprecher*innen in Folge eines globalen Zugzwangs (vgl. Beispiel 1) oder im Rahmen wiederholten Opponierens (vgl. Beispiel 2) Gesprächsgegenstände ablehnend fallenlassen und somit aus der weiteren Interaktion ausschließen. Anhand des zweiten Beispiels wird dabei ersichtlich, dass sich der Ressourceneinsatz intensivieren kann, wenn eine erstmalige Ablehnung durch das Gegenüber nicht akzeptiert, sondern hinterfragt wird.

Beispiel 1 illustriert, wie die Tochter im Anschluss an einen Zugzwang der Mutter erst durch faziale Ressourcen, wie das Zusammenpressen der Lippen, das Kräuseln der Nase sowie das Anheben der Augenbrauen, einen *critical stance* (vgl. Heller et al. 2023) anzeigt, der eine kritische Haltung gegenüber der von der Mutter eingebrachten Option markiert. Anschließend verbalisiert die Tochter eine Ablehnung gegenüber der Option des Segeltörns. Zu Beginn der Elaboration der Ablehnung dreht die Sprecherin ihre Hand auf und lässt sie geräuschvoll fallen – sie zeigt visuell wahrnehmbar den Ausschluss des Gesprächsgegenstandes an.

Beispiel 1 – Markieren des ablehnenden Fallenlassens eines Gesprächsgegenstandes

098	MU:	und äh (.) SE geln?	
	mu-gz		blickt zu TO (>Z. 107)
099		(--)	
			
100		warum NI CH? (1.0) (0.4)	
	to-mo		presst Lippen zusammen
	to-ns		kräuselt Nase-----
	to-br		hebt AB zusammengezogen an->
	to-gz		 blickt zu MU (>Z. 102)
101		(0.4) (1.0) (0.5)	
	mu-fa	lacht-----	
	to-fa	lacht-	
	to-br	----->	
102		nee: segeln is nich so.	
	to-gz	blickt auf Blatt (>Z. 104)	
103		nein.	



Die Haltung der Tochter gegenüber dem vorgeschlagenen Segeltörn wird erst durch faziale Ressourcen (Z. 100) sowie anschließend durch mehrfache Negationen (Z. 102f.) markiert, bevor die Sprecherin dann eine Ablehnung unter Bezug auf ihr persönliches Befinden (Z. 104: als ich mein:: (.) ich ich persÖNlich mag sowieso nicht so gerne aufm schiff oder so.) begründet. Indem die Sprecherin die Ablehnung der Option auf ihrem persönlichen Befinden gründet und ihr Argument durch *sowieso* als außer Frage stehendes Faktum absichert, entzieht sie jedweder weiteren Diskussion die Grundlage (vgl. Eggs 2003): Da sie *sowieso* nicht so gern auf Schiffen ist, erscheint es wenig sinnvoll, diese Option weiter auszuhandeln. Die Ablehnung der Option wird seitens der Mutter anschließend ratifiziert (Z. 105). In der Sequenz zeigt sich nun, dass durch den zu Beginn gesetzten globalen Zugzwang der Mutter (Z. 098, 100) die Tochter dazu verpflichtet wird, eine Stellung gegenüber einer Option zu beziehen, deren Grundlage sie nicht anerkennt: Das turninitiale Aufdrehen der Hand (vgl. #2), mit dem die Tochter im Verbund mit dem Diskursmarker *ich mein* (vgl. Günthner und Imo 2003) in die Turnfortsetzung einleitet (vgl. Kap. 8.2.3), sowie das anschließende Anheben und geräuschvolle Fallenlassen (vgl. #3-#5) der Hand zeigen das Ausschließen des Gesprächsgegenstandes aus der weiteren Interaktion für die Beteiligten bereits turninitial sichtbar an. Die weitere Bearbeitung dieses Gesprächsgegenstandes wird als dispräferiert markiert, sodass die *Dropping of Hand* Geste nicht nur Hinweise über den Ausschluss eines Gesprächsgegenstandes trägt, sondern darin auch eine ablehnende Haltung der Sprecherin gegenüber einer weiteren Aushandlung dieses Gesprächsgegenstandes manifest wird. Die Sprecherin beansprucht dadurch zugleich epistemische Autorität über das Fallenlassen des Gesprächsgegenstandes. Zwar bearbeitet die Sprecherin den Zugzwang der Mutter in ihrem *Turn* satzförmig, doch wird eine weitere Aushandlung der Option nach einer Zustimmung der Mutter (Z. 105, 107) anschließend von beiden Sprecherinnen aufgegeben.

In dem nachfolgenden Beispiel 2 lehnt der Vater eine Begründung des Sohnes widersprechend ab. Als der Sohn diese Opposition erst hinterfragt und anschließend seine zuvor eingebrachte Begründung mit einer Belegerzählung (vgl. Heller 2012: 87) stützt, widerspricht der Vater ein weiteres Mal – nun mit einem gesteigerten Einsatz multimodaler Ressourcen.

Beispiel 2 – Markieren des ablehnenden Fallengelassens eines Gesprächsgegenstandes

337 SO: |und [DEM macht das] dann keinen spaß.
 va-gz |blickt auf Tisch/Blatt (>Z. 341)
 338 VA: [wiesO,]
 339 |ich hatte KEI|nen.|
 va-gs |legt Stift auf Option Segeltörn (>Z. 343)
 va-he |schüttelt Kopf----|
 340 SO: JA,
 341 VA: nö.|
 va-gz -->|
 342 SO: |n |bei UNS sind schon wel|che auf ner ANderthalb stunden fährt |nach==
 so-gs |zeigt mit ausgestrecktem rIF auf sich-----|zeigt
 Linie in Luft->
 va-gz |zu SO-----|blickt auf Blatt (>Z. 344)



343 VA: |=<<h>ICH |hatte |keinen;=>=
 va-lh ----->|
 va-rh |.....|führt rH Palm Up n. o.|lässt Hd geräuschvoll fallen|
 va-he |schüttelt Kopf-----|
 so-gs ----->|
 344 SO: =nach (.) DOver |von [caLAIS |aus] |(.) krank geworden;=
 so-gz ----->|blickt auf Blatt (>Z. 346)
 va-gz ----->|blickt zu SO|auf Blatt----->|
 345 VA: [ja=passiert;]
 346 =wobei äh ijsselmeer is natürlich (.) is is |mehr SEE als meer; (0.5)
 so-gz ----->|
 347 also (-) <<p>nich so HEftiger wellen|gang.>
 348 °hh während HIER bei vier tage bÜngalowpark,

Der Vater weist die Begründung des Sohnes zurück, dass Mitschüler*innen, die an Seekrankheit leiden, keine Freude an einer Abschlussfahrt auf einem Segelschiff haben würden. In der opponierenden Handlung bezieht er sich auf seine eigenen Segelerfahrungen (Z. 339). Indem der Sohn auf den ersten Widerspruch des Vaters (Z. 339) hin eine Belegerzählung anführt (Z. 342, 344) und sich somit gegen den Widerspruch des Vaters wendet, ist der Vater zu einer weiteren Reaktion gefordert: Es erfolgt nun eine zweite, wiederholte Zurückweisung der Begründung des Sohnes (Z. 343). Diese wird mit einem erhöhten Tonhöhenregister, einer Verschiebung des Fokusakzents auf das Personalpronomen, einem erneuten Kopfschütteln, das die Negation (Z. 343: keinen) verstärkt, sowie einer *Dropping of Hand* Geste geäußert, bei der die rechte Hand mit der Handfläche nach oben orientiert angehoben und geräuschvoll fallengelassen wird. Eine Zurückweisung der Belegerzählung des Sohnes wird also nicht nur sprachlich durch die Opposition selbst erkennbar, sondern ebenfalls durch die körperlich-sichtbaren Ressourcen, mit denen der Gesprächsgegenstand in ablehnender Haltung fallengelassen und aus der weiteren Aushandlung ausgeschlossen wird. Anschließend etabliert der Vater, nach einem letzten Einwand (Z. 346-347), nun einen neuen Gesprächsgegenstand – die Verhandlung des Segeltörns wird an dieser Stelle fallengelassen.

Insgesamt zeigt sich, dass Interagierende durch die turninitiale *Dropping of Hand* Geste einen Gesprächsgegenstand oder Inhalte der Vorgängeräußerung aus der weiteren Interaktion als auszuschließend kontextualisieren können. Das dabei entstehende hörbare Signal scheint als akustische Demarkationslinie zu fungieren, die einen anschließenden Themenwechsel bereits ankündigt. Hier übernimmt die Geste im Ressourcenverbund und in Abhängigkeit von ihrer sequenziellen Position eine performative Funktion, die das Fallenlassen eines Gesprächsgegenstandes, das sprachlich implizit verbleiben kann, für alle Beteiligten sichtbar anzeigt. Das Anheben und anschließende Fallenlassen der *losen Hand* dient außerdem einer Dispräferenzmarkierung: Die Sprecher*innen bearbeiten den Gesprächsgegenstand der Vorgängeräußerung zwar kurzzeitig in ablehnender Haltung, wenden sich aber in der interaktiven Konsequenz kurz darauf anderen Gesprächsgegenständen zu. Zu den Handgesten können auch weitere Ressourcen, wie faziale Gesten (vgl. Beispiel 1) oder Gesten des Kopfes (vgl. Beispiel 2), hinzutreten und den Ausdruck der ablehnenden Haltung verstärken. Diese erweisen sich als optional – die ablehnende Haltung der Sprecher*innen kann in Abhängigkeit vom sequenziellen Kontext auch allein durch sprachliche Ressourcen und die *Dropping of Hand* Geste angezeigt werden. Insgesamt nimmt der*die Sprecher*in dabei die epistemische Autorität (vgl. Stivers et al. 2011; Heller 2018a) für das Ausschließen des Gesprächsgegenstandes an sich; dies obliegt keinem gemeinsamen Aushandlungsprozess.

8.6.3 Anzeige von epistemischer Unsicherheit

Neben der Markierung von Evidenz (vgl. Kap. 8.6.1) und dem Beanspruchen epistemischer Autorität über das Fallenlassen eines Gesprächsgegenstandes (vgl. Kap. 8.6.2) können Interagierende sich gegenseitig aber auch eine epistemisch unsichere Haltung anzeigen. Die Interagierenden der vorliegenden Studie nutzen zum Ausdruck eines „insufficient knowledge“ (Pekarek Doehler 2016: 161) körperlich-sichtbare Ressourcen, wie das Aufdrehen der Hand/Hände und Unterarme sowie das Hochziehen der Schultern, und metakommunikative formelhafte Phrasen, wie *ich weiß (es) nicht*, *weiß ich nicht*, *was weiß ich* oder *keine Ahnung* (vgl. zu *ich weiß nicht* Helmer und Deppermann 2017; Helmer et al. 2017; vgl. für das Französische Pekarek Doehler 2016, 2019; Debras 2021; vgl. Tab. 31). Mit Helmer et al. (2017: 391) kann dabei zwischen prospektivisch sowie retrospektiv angezeigter epistemischer Unsicherheit unterschieden werden.⁹⁴

Anzeige von epistemischer Unsicherheit	
Beschreibung der multimodalen Gestalt	formelhafte Phrase (z. B. <i>ich weiß es nicht</i>) + <i>Palm Up</i> Geste (z. T. mit lateraler Bewegung) (+ optional Schulterzucken, Kopfschütteln)

Tabelle 31: Beschreibung der multimodalen Gestalt „Anzeige von epistemischer Unsicherheit“

In den vorliegenden argumentativen Entscheidungsfindungen zeigt sich, dass Interagierende sich durch multimodale Gestalten, bei denen die Sprecher*innen eine *Palm Up* Geste, z. T. mit lateraler Bewegung, Schulterzucken und Kopfschütteln produzieren, eine relative Unsicherheit in Bezug auf den präsentierten Gesprächsgegenstand anzeigen. Sprecher*innen verweisen dann in der Mitte ihrer Äußerung (turn-

⁹⁴ Die Autor*innen unterscheiden hierbei zwischen einer *epistemischen* Markierung von *ich weiß nicht*, die eine „Unsicherheit bezüglich der Wahrheit einer Folgeäußerung anzeigt“, von Nicht-Wissen, und einer *pragmatischen* Markierung, die sich „metapragmatisch auf die Hinlänglichkeit einer folgenden Äußerung“ bezieht (Helmer et al. 2017: 393). In Letzterer kann der*die Sprecher*in „durchaus eine hohe epistemische Sicherheit über den Wahrheitsgehalt der Folgeäußerung haben“, aber „unsicher über die Nützlichkeit oder Relevanz der folgenden Information“ (Helmer et al. 2017: 393) sein. In den vorliegenden Daten zeigen die Sprecher*innen sich gegenseitig allerdings durchgängig epistemische Unsicherheit an, sodass diese im Fokus des Kapitels stehen wird.

medial) prospektiv auf eine nicht abgesicherte Wissensdomäne und setzen ihre Äußerung fort (vgl. Beispiel 1). Die multimodale Gestalt zeigt sich aber auch turnfinal bzw. turnexpandierend, wenn Sprecher*innen retrospektiv ihre Begründung epistemisch herabstufen (vgl. Beispiel 2).

In dem nachfolgenden Beispiel markiert der Sprecher multimodal innerhalb eines größeren Äußerungspakets epistemische Unsicherheit in Bezug auf die Anzahl der Personen, die in einem Bungalow untergebracht werden können.

Beispiel 1 – Anzeige von epistemischer Unsicherheit

348 VA: °hh während HIER bei vier tage bUngalowpark,
349 |HEIßT das ja:-
va-gz |blickt in mittlere Distanz->
350 in dem |BUN |galow |sind- |
va-gz ----->|blickt zu SO (>Z. 355)
va-rh -----|tippt----->|legt Stift auf Option|
va-lh |öffnet Fi|,,,,,,|.....->
so-gz |blickt zu VA----->>



351 |was weiß ICH; |
va-gs |öffnet & hebt bHde Palm Up, lateral|



352 |zu SECHST oder- |
va-gs |legt bHde übereinander; lässt bHde auf Tisch fallen|

353 SO: ja::,

354 VA: |<<p>zu viert (oder) zu sechst> (.) zu |ACHT| oder so,
va-ey |verengt Augen-----|
va-br |zieht AB zusammen-----|
va-ns |kräuselt Nase-----|
va-he |Beat|

355 °hhh |(-) DA: WÜRDEN euch die lehrer natürlich NICHT so genau unter
va-gz ---->|
kontrolle haben;

356 SO: [das is doch gerade der !SINN! an so einer abschlussfahrt-]

Das Markieren von Nicht-Wissen bzw. epistemischer Unsicherheit erfolgt in der abgebildeten Sequenz (Z. 348-355) in Form zweier aufeinanderfolgender multimodaler Gestalten (Z. 351, Z. 354). Die erste Gestalt, bei der die formelhafte Phrase *was weiß ich* (Z. 351) gemeinsam mit einer beidhändigen *Palm Up* Geste mit lateraler Bewegung geäußert wird (vgl. #1), steht hier im Fokus. Die Unsicherheit des Sprechers, so wird aus der sich weiter entfaltenden Sequenz ersichtlich, bezieht sich auf die Personenanzahl, die in einem Bungalow beherbergt werden kann: Der sprachlichen sowie gestischen Markierung

von Unsicherheit vorauslaufend (Z. 349) zeigt der Sprecher durch den Blick in die mittlere Distanz (vgl. Goodwin und Goodwin 1986; Heller 2021b) einen Prozess des Nachdenkens öffentlich sichtbar an. Dieser ist für den Sohn, der seinen Blick auf ihn gerichtet hat, wahrnehmbar (Z. 350). Mit der Blickzuwendung zum Sohn (Z. 350) bereitet der Sprecher nun mit der linken Hand eine Geste vor. Er überführt dann die Hände in eine beidhändige *Palm Up* Geste, die in lateraler Bewegung nach oben geführt wird. Gemeinsam mit der formelhaften Phrase *was weiß ich* (Z. 351) zeigt der Sprecher seine Unsicherheit multimodal an. Er führt die Hände anschließend kreuzend übereinander und lässt sie dann, mit einem ersten Vorschlag zu einer möglichen Anzahl von Personen, die gemeinsam einen Bungalow beziehen können (Z. 352: zu SECHST oder-), auf den Tisch fallen (vgl. #2). In einer gemeinsamen multimodalen Gestalt markieren also die formelhafte Phrase *was weiß ich* sowie die beidhändige *Palm Up* Geste mit lateraler Bewegung prospektiv die Fortsetzung der Äußerung als epistemisch unsicher. Während nun die gestische Markierung der epistemischen Unsicherheit endet, da die Hände in einer Ruheposition auf dem Tisch verbleiben, werden die Unsicherheitsmarkierungen innerhalb einer zweiten multimodalen Gestalt fortgesetzt: Die vom Sprecher eingebrachte Vermutung über die mögliche Personenanzahl in den Bungalows (Z. 354) wird durch die verminderte Lautstärke, das Kräuseln der Nase, das Zusammenziehen der Augenbrauen und das Verengen der Augen sowie die finale Etcetera-Formel *oder so* (Z. 354) als epistemisch nicht abgesichert bzw. ungewiss markiert (vgl. Helmer und Deppermann 2017: 145; vgl. zu den fazialen Ressourcen Ruusuvuori und Peräkylä 2009; Kaukomaa et al. 2014). Die bereits prospektiv durch die formelhafte Phrase sowie die *Palm Up* Geste mit lateraler Bewegung angezeigte Unsicherheit findet sich also in der sequenziellen Entfaltung des *Turns* fortgeführt. An die Unsicherheitsmarkierungen schließt der Vater nun ein weiteres Argument an (Z. 355: °hhh (-) DA: WÜRDEN euch die lehrer natürlich NICHT so genau unter kontrolle haben;). Diesem begegnet der Sohn in interaktiver Konsequenz nun mit einer eigenen Stellungnahme, in der er das ablehnende Argument des Vaters als Fürsprache für die Auswahl des Bungalowparks anführt (Z. 356: das is doch gerade der !SINN! an so einer abschlussfahrt-). Inwiefern Interagierende epistemische Unsicherheit auch retrospektiv multimodal markieren, zeigt das nachfolgende Beispiel 2.

In dem zweiten Beispiel wird die formelhafte Phrase *ich weiß (es) nicht* (vgl. Helmer et al. 2017) expandierend zu einer Begründung eingesetzt, als eine gestisch eingeforderte Rückmeldung seitens des Vaters (Z. 253) zunächst ausbleibt (Z. 254) und anschließend ausschließlich durch die Antwortpartikel *okay* (vgl. Betz und Deppermann 2021) erbracht wird. Die turnexpandierende Markierung von Nicht-Wissen bzw. Unsicherheit wird gemeinsam mit dem Aufdrehen einer Hand sowie einem Schulterzucken als multimodale Gestalt realisiert.

Beispiel 2 – Anzeige von epistemischer Unsicherheit

253	SO:	da würd sich (ja) eine neunte klasse zu TOda langweilen.
254		(1.7) (-)
	so-gs	hält bHde auf Handkanten aufgestellt ,,, ->
255	VA:	oKAY,
	so-gs	,,,,,
	so-gz	---->



256 SO: al|so |ich |WEI|SS_s nich.
 so-gs |dreht lH auf|,,,
 so-bo |Shoulder Shrug--|

257 also-
 258 (1.3)
 259 ah- (.)
 260 VA: gut i ich meine es is EURE (.) abschlussfahrt.
 261 und dann solltet IHR auch drüber entscheiden.
 262 (2.0)
 263 SO: °hhh al[SO-]
 264 VA: [WIR s] wir sind ja nur zur finanzierung da.

Das Markieren von epistemischer Unsicherheit erfolgt in diesem Beispiel nach einer ausbleibenden Turnübernahme des Vaters in Folge einer äusserungsübergreifenden, selbstinitiierten Begründung seitens des Sohnes, in der er das Für und Wider der Abschlussfahrt nach Rom abwägt (nicht im Transkriptauszug enthalten). Der Vater ratifiziert diese, nach einer 1.7-sekündigen Pause, ausschließlich mit der Antwortpartikel *okay* (Z. 255; vgl. Betz und Deppermann 2021), die eine steigende Intonation trägt. Eine weitere Rückmeldung in Bezug auf die Begründung des Sohnes bleibt aus, obwohl diese aufgrund der abschließenden *Extreme Case Formulation* (Z. 253: zu Töda langweilen; vgl. Pomerantz 1986: 219) erwartbar wäre.⁹⁵ Es folgt nun eine multimodale Gestalt, mit der der Sohn durch die formelhafte Phrase „ich WEISS_s nich.“ (Z. 256), das Aufdrehen der linken Hand zu einer *Palm Up* sowie einem Schulterzucken epistemische Unsicherheit anzeigt (vgl. #1). Durch die Turnexpansion (vgl. zu *sequence closing thirds* Schegloff 2007: 118) markiert der Sprecher eine Unsicherheit hinsichtlich der Relevanz seiner eigenen Begründung, sodass er retrospektiv die dargelegte Begründung herabstuft (vgl. auch Helmer et al. 2017). In interaktiver Konsequenz lädt die multimodale Gestalt das Gegenüber zu einer Stellungnahme gegenüber der als unsicher markierten Begründungssequenz ein, die schließlich (Z. 260f., Z. 264) seitens des Vaters erfolgt. Der Vater tritt der Unsicherheit des Sohnes entgegen, indem er dem Sohn ein Entscheidungskriterium anbietet, das ihm in der Entscheidungsfindung helfen kann.

Insgesamt zeigt sich, dass Sprecher*innen ihre epistemische Unsicherheit durch formelhafte Phrasen wie *was weiß ich* oder *ich weiß es nicht*, die zeitlich koordiniert werden mit aufdrehenden Händen zu *Palm Up* Gesten sowie weiteren *Shrug*-Komponenten (vgl. u. a. Streeck 2009b), explizieren. Je nach sequenzieller Position können Sprecher*innen sowohl prospektiv als auch retrospektiv epistemische Unsicherheit anzeigen, die dann das Gegenüber zur Stellungnahme in Bezug auf den als unsicher markierten Inhalt einlädt.

⁹⁵ Betz und Deppermann (2021) zeigen in ihrer Untersuchung, dass der Einsatz der Antwortpartikel *okay* mit steigender Intonationskontur im Anschluss an eigentlich als abgeschlossen markierte *Turns* das Gegenüber zu einer weiteren Elaboration des *Turns* einladen kann.

8.6.4 Markieren von Unbestimmtheit

In der Ausführung von Begründungen, Einwänden etc. innerhalb mündlicher Argumentationen können Sprecher*innen auch inhaltliche Vagheit bzw. Unbestimmtheit markieren. Dabei ist die epistemische Haltung der Sprecher*innen nicht notwendigerweise eine unwissende, wie die vorliegende Studie zeigt. Denn die Sprecher*innen können durchaus wissend bzw. sicher in Bezug auf den dargelegten Gesprächsgegenstand sein und zugleich eine inhaltliche Unbestimmtheit des *Gesprächsgegenstandes selbst* durch multimodale Ressourcenverbünde ausdrücken. Die Unbestimmtheit bzw. Vagheit wird dabei sowohl sprachlich manifest als auch im Verbund mit körperlich-sichtbaren Ressourcen wie beispielsweise *Palm Down Open Hand* Gesten (vgl. Bressemer und Müller 2014a: 1580), bei denen die Handfläche nach unten orientiert geöffnet ist und die Hand rotierend vor- und zurückbewegt wird (vgl. Tab. 32).

Markieren von Unbestimmtheit	
Beschreibung der multimodalen Gestalt	sprachliche Markierung von Unbestimmtheit + vor/zurück rotierende <i>Palm Away</i> Geste bzw. iterative Vor-/Zurückbewegung einer <i>Palm Up</i> Geste

Tabelle 32: Beschreibung der multimodalen Gestalt „Markieren von Unbestimmtheit“

In den vorliegenden Beispielen markieren die Sprecher*innen sowohl inhaltliche Unbestimmtheit in Bezug auf eine Entfernungsangabe (vgl. Beispiel 1) als auch eine inhaltliche Unbestimmtheit in Bezug auf die beliebige Erweiterbarkeit eines Szenarios (vgl. Beispiel 2).

Das erste Beispiel illustriert, wie die Sprecherin hinsichtlich einer Entfernungsangabe, die die Sprecherin in Bezug auf den Segeltörn einbringt, Unbestimmtheit markiert.

Beispiel 1 – Markieren von Unbestimmtheit

178 man macht was zuSAMMEN,
179 MU: ((zählt weiter mit der rH an lH ab))
180 (1.3)



181 TO: man fährt (.) |naja (.) so MITtelmäßig |weg; |
to-gs |rotiert bHde Palm Away vor & zurück|
to-he |bewegt Kopf von re n. li|
mu-gs |zählt an MF ab, hält lH mit Zählgeste (>Z. 208)

182 naja;
183 MU: tja;
184 [sonst AUCH-]
185 TO: [(ich denk) man] SIEHT auch (.) jeden tag etwas ANderes wahrscheinlich,
186 (1.3)

Die Sprecherin führt innerhalb eines größeren Äußerungspakets eine Liste von Argumenten für den Segeltörn an (vgl. 178, 181). Der vorliegende Transkriptausschnitt zeigt dabei innerhalb eines der Argumente (Z. 181) eine Unbestimmtheitsmarkierung: Die Sprecherin rekurriert auf die Entfernung der Abschlussfahrt. Konträr zu der eigentlichen Annahme, dass bei einer Abschlussfahrt die Beteiligten

potenziell ein weiter entfernt liegendes Ziel präferieren, bringt die Sprecherin nun, im Verlaufe der Fürsprache für den Segeltörn, die „MITtelmäßig[e]“ (Z. 181) Entfernung ein. Es ist anzunehmen, dass sie hier nun mit der Pause sowie der anschließenden Partikel *naja* eine *online* Korrektur ihrer Äußerung vornimmt (von „man fährt weit weg“ zu „man fährt (.) naja (.) so MITtelmäßig weg“). Denn mit der Partikel *naja* können Sprecher*innen „einen Übergang innerhalb eines durch Diskrepanz oder Divergenz gezeichneten Wissens- oder Handlungszusammenhangs [vollziehen]“ und somit „das Nachfolgerelement als das im Kontext relevante“ markieren (Zifonun et al. 1997: 401). In dieser Korrektur markiert die Sprecherin nun eine Unbestimmtheit in Bezug auf die Entfernung der Abschlussfahrt. Dazu trägt einerseits die Unschärfepartikel *so* bei (vgl. Weinrich 2003: 687) sowie auch das Lexem „MITtelmäßig“ selbst, mit dem eine exakte Entfernungsbestimmung ausbleibt. Stattdessen zeigt die Sprecherin die inhaltliche Un- bzw. Unterbestimmtheit auch körperlich an: Sie rotiert beide Hände *Palm Away* vor und zurück und aligniert mit dieser Geste auch die Bewegung des Kopfes, der von links nach rechts geführt wird. Diese Geste beschreiben Bresse und Müller (2014a: 1580) als „PDOH [Palm Down Open Hand, N. S.] with clockwise rotation“, mit der Sprecher*innen „events, states, as well as ideas“ als „uncertain and indeterminate“ markieren können. In der vorliegenden Sequenz nutzt die Sprecherin diese Geste im Verbund mit weiteren Ressourcen, um Unbestimmtheit zu markieren: Sie schließt die Sequenz mit einer Wiederholung der Partikel *naja* – nun in sequenzschließender Funktion (vgl. Golato 2018: 420f.) – ab.

Auch in dem zweiten Beispiel markiert eine Sprecherin multimodal eine Form der Unbestimmtheit. Sie baut in der vorliegenden Sequenz ein hypothetisches Szenario (vgl. Brünner und Gülich 2002a) auf und beschreibt darin, welchen Aktivitäten die Schüler*innen auf dem Segeltörn nachgehen können, sollte es während der gesamten Abschlussfahrt regnen. Die inhaltliche Unbestimmtheit bezieht sich nun auf die Beliebigkeit der benannten Aktivitäten.

Beispiel 2 – Markieren von Unbestimmtheit

190 MU: =NEHmen wir mal an das WETter ist schlecht.
 191 ne?
 192 SO: |ja.=[dann HASse n problem.]|
 so-he |nickt-----|
 193 MU: [und dann sitzt man (.)] FÜNF tage |(.) UNter |deck, ne,|
 mu-gz |blickt zu SO (>Z. 195)|
 mu-he |nickt----|
 194 SO: [ja.]
 195 MU: |[spielt] |viel |leicht|
 mu-gz ----->
 mu-gs |.....|öffnet LH vertikal|,,,,,|



KARten |oder |irgendwie| geSELLschaftsspiele;
 mu-gz ----->|blickt auf Blatt----->
 mu-gs |bewegt LH vertikal vor & zurück--|
 mu-he |leichtes Kopfschütteln|

196 °hh ähm: (-) hat KEine möglichkeit irgendwie: (.) WEISS ich nich-|
 mu-gz ----->|
 blickt zu SO->>
 197 ipod oder <<lachend>halt ins INternet zu kommen.>=ne?

In ihrem *Multi-Unit Turn*, von dem hier nur ein Ausschnitt wiedergegeben wird, entwickelt die Sprecherin innerhalb einer Begründung ein hypothetisches Szenario (Z. 190-197; vgl. Brünner und Gülich 2002a), dessen Imagination vom Gegenüber durch eine metakommunikative Einleitung (Z. 190: =NEHMEN wir mal an das WETter ist schlecht.) eliziert wird. Innerhalb des Szenarios benennt die Sprecherin dann Aktivitäten, denen die Schüler*innen bei schlechtem Wetter auf dem Segeltörn nachgehen können (Z. 195: spielt vielleicht KARten oder irgendwie geSELLschaftsspiele;). Es handelt sich dabei um Beispielaktivitäten, die die Sprecherin mit dem Satzadverbial *vielleicht* sowie den Identikativ-Artikel (vgl. Weinrich 2003: 475) *irgendwie* als *mögliche*, aber durch den Rezipienten durchaus erweiter- bzw. beliebig austauschbare Aktivitäten einordnet. Verstärkt wird dieser Eindruck ebenfalls durch den Selektiv-Junktor *oder* (vgl. Weinrich 2003: 810), der die ersten beiden Beispielaktivitäten verbindet und auf eine Offenheit, diese zu erweitern, hindeutet. Zeitlich aligniert mit „oder irgendwie“ (Z. 195) produziert die Sprecherin mit der linken Hand, die vertikal geöffnet ist, mehrmalige Bewegungen vor und zurück sowie ein leichtes Kopfschütteln, mit dem Sprecher*innen „another possibility or missing piece of information“ (McClave 2000: 862) einräumen können, während der Blick auf das Material gerichtet ist. Die Geste fungiert hier als multimodale Etcetera-Formel (vgl. dazu *oder so* in Beispiel 1), mit der die Unbestimmtheit der konkreten Aktivitäten der Jugendlichen sowie zugleich deren beliebige Erweiterbarkeit markiert wird.

Zusammengefasst zeigt sich also, dass Interagierende die inhaltliche Unbestimmtheit oder Vagheit ihrer Äußerung multimodal markieren. Gestische Ressourcen, wie das mehrfache Rotieren der geöffneten Hand sowie das mehrfache Vor- und Zurückbewegen der geöffneten Hand, und sprachliche Ressourcen, wie die Markierung durch Unbestimmtheitspartikeln oder Identikativ-Artikel, realisieren dabei gemeinsam eine multimodale Gestalt der Unbestimmtheit. Trotz der Unterschiede der Gesten in ihrer Form teilen sie ein gemeinsames, relevantes Merkmal in ihrer Bewegung: Eine iterative Vor- und Zurückbewegung, mit der die Sprecher*innen die Unbestimmtheit (auch i. S. einer beliebigen Erweiterbarkeit) ausdrücken.⁹⁶

8.6.5 Zusammenfassung

In dem vorangegangenen Kapitel konnten multimodale Gestalten rekonstruiert werden, die zur epistemischen Modalisierung beitragen.

- Interagierende können durch unterschiedliche multimodale Gestalten den Inhalt ihrer Äußerung als fraglos evident markieren und dadurch den Geltungsanspruch ihres Beitrags erhöhen (vgl. Kap. 8.6.1). Diese modalisierende Funktion kann durch unterschiedliche multimodale Gestalten übernommen werden: durch das Fallenlassen der Hand während der Rückführung der Geste in ihre Ruheposition in einer epistemisch nicht abgeschwächten, z. T. epistemisch verstärkt modalisierten Äußerung (Kap. 8.6.1.1); durch eine *Palm Up* Geste mit lateraler Bewegungsrichtung, während die Satzaussage epistemisch verstärkt wird (Kap. 8.6.1.2); sowie durch einen *Hand Shrug*, d. h. durch das kurzzeitige Aufdrehen der Hand in Kombination mit weiteren *Shrug*-Komponenten (z. B. Schulterzucken, Abnicken des Kopfes, Hochziehen der Augenbrauen, *Mouth Shrug*; vgl. Kap. 8.6.1.3).
- Multimodale Gestalten können auch Aufschluss über die ablehnende Haltung eines*einer Sprecher*in geben: So können in der gemeinsamen argumentativen Entscheidungsfindung verhandelte Gesprächsgegenstände der Vorgängeräußerung von dem*der aktuellen Sprecher*in durch

⁹⁶ Dass sich die Bewegungsrichtung von Gesten als relevant für ihre Funktionen erweist, wurde bereits in Bezug auf die Klassifikation der Familie der *Away* Gesten von Bressen und Müller (2014b; vgl. auch Kap. 3.2.2.2) dargelegt.

das lose Fallenlassen der Hand in Form einer *Dropping of Hand* Geste als dispräferiert markiert und aus der nachfolgenden Interaktion ausgeschlossen werden. Die eingesetzte *Dropping of Hand* Geste bildet sowohl visuell als auch akustisch eine Demarkationslinie, mit der ein anschließender Themenwechsel angekündigt wird. Dabei reklamiert der*die Sprecher*in zugleich auch die epistemische Autorität, den Gesprächsgegenstand aus der weiteren Interaktion auszuschließen (vgl. Kap. 8.6.2).

- Auch epistemische Unsicherheit können durch multimodale Gestalten angezeigt werden. Dazu setzen Interagierende Ressourcenverbünde aus formelhaften Phrasen epistemischer Unsicherheit ein (z. B. *ich weiß es nicht*), während sie zugleich eine *Palm Up* Geste (z. T. mit lateraler Bewegung) äußern und auch weitere *Shrug*-Komponenten wie das Schulterzucken oder Kopfschütteln hervorbringen (vgl. Kap. 8.6.3). In interaktiver Konsequenz kann die Anzeige epistemischer Unsicherheit in Bezug auf ein vorgebrachtes Argument das Gegenüber zu einer Stellungnahme einladen.
- Das Anzeigen von Unbestimmtheit erfolgt durch die Verbindung aus sprachlichen Unbestimmtheitsmarkierungen sowie vor und zurück rotierenden *Palm Away* Gesten bzw. iterativen, sich vor- und zurückbewegenden vertikal geöffneten *Palm Up* Gesten (vgl. Kap. 8.6.4). Durch die multimodale Gestalt können Sprecher*innen eine inhaltliche Unbestimmtheit des von ihnen verhandelten Gesprächsgegenstandes anzeigen, ohne dabei notwendigerweise selbst eine epistemisch unwissende Haltung einzunehmen.

Abschließend führt die nachfolgende tabellarische Übersicht 33 die systematischen Beschreibungen der multimodalen Gestalten und ihrer Funktionen, die in den vorangegangenen Kapiteln rekonstruiert werden konnten, komprimiert zusammen. Es folgt sodann eine Einordnung der Befunde im Hinblick auf die erste Forschungsfrage (vgl. Kap. 9).

Argumentative Handlung	Funktion der multimodalen Gestalt	Beschreibung der multimodalen Gestalt
Position beziehen	Hervorheben von Optionen beim Beziehen von Positionen	Benennen der Position (z. B. durch Stellungnahme, Bewertung) + Benennen der Option (X) + deiktische Geste auf das Material
	Position beziehen durch Ausschließen einer Option	Benennen der Option + <i>Throwing</i> bzw. <i>Holding Away</i> Geste (optional: sprachliche Explikationen des Ausschlusses)
Begründen	Präsentieren von (komplexeren) Begründungen	Begründung bzw. Begründungssegment(e) + <i>Palm Up</i> Geste + <i>Post-stroke Hold</i> (+ turnfinales Halten)
	Kurzzeitiges Präsentieren und Hervorheben relevanter Bedeutungselemente der Begründung	einzelnes Lexem (i. d. R. + Fokusakzent) + <i>Palm Up</i> Geste (optional: <i>Beat</i>)
	Etablieren und Präsenthalten eines verhandelten Gesprächsgegenstandes	- Benennen des Gesprächsgegenstandes (z. B. Option X) + deiktisch-referenzierende Geste (auf Material) + Halten der Geste äußerungsintern und/oder über Sprecher*innenwechsel hinweg - Benennen des Gesprächsgegenstandes (z. B. entscheidungsrelevante Kriterien) + Zählgeste + Halten der Geste äußerungsintern und/oder über Sprecher*innenwechsel hinweg
	Etablieren und Präsenthalten eines verhandelten Gesprächsgegenstandes bei konkurrierender Turnübernahme	Benennen des Gesprächsgegenstandes (z. B. Option X) + Halten der ausgeführten Geste in konkurrierender Turnübernahme
	Referenzieren und Präsenthalten eines verhandelten Gesprächsgegenstandes als Verstehensdokumentation (rezipient*innenseitig)	deiktische Geste auf Blatt während der Begründung des Gegenübers
	Multimodales Einführen einer Turnfortsetzung	sprachliches Strukturelement (z. B. formelhafte Phrase oder Konnektor) + <i>Palm Up</i> Geste
	Multimodale Kohäsion und Kohärenz	- Gestenwiederaufnahmen, die sich in Form und/oder Bewegung ähneln + - ähnliche lexikalische/morpho-syntaktische Elemente - Wiederaufnahme inhaltlicher bzw. thematischer Elemente (obligatorisch für das Markieren von Kohärenz) an relevanter sequenzieller Position
	Verwerfen in Begründungen	<i>Throwing Away</i> bzw. <i>Brushing Away</i> Gesten + Benennung ausgeschlossener Gesprächsgegenstände
Zustimmen	Nachdrückliches Zustimmung	(ausgebaute) sprachliche Zustimmung + iterative bzw. ostentative deiktische Geste auf das Material - anschließend an vorangegangene Opposition

Schönfelder: Multimodale Ressourcen beim Argumentieren im Gespräch

	Schwaches Zustimmung in Konzessionen	schwache sprachliche Zustimmung + Aufdrehen der Hand bzw. Extension einzelner Finger + Blick abgewandt + i. d. R. Abknicken des Kopfes weg vom Gegenüber läuft sequenziell einem Einwand/Widerspruch voraus
	Schwaches Zustimmung zu evidenten Begründungen & Positionen	- Aufdrehen der Hand bzw. Extension einzelner Finger + Blick abgewandt + seitliche Kopfgeste + schwache Affirmation - läuft sequenziell verbalem Beipflichten und/oder Themenwechseln voraus
Einwenden	Präsentieren während des Einwendens	- Einwand + <i>Palm Up</i> Geste (in unterschiedlicher Ausrichtung der Geste) - erfolgt sequenziell responsiv
	Sichtbares Explizieren von Negationen in Einwänden	Geste mit horizontaler Bewegungsrichtung (z. B. <i>Palm Horizontal/Palm Down Horizontal Across</i> oder horizontale Bewegung eines Stiftes) (+ optional sprachliche Negation)
	Anhalten einer vorangegangenen Opposition in Konzessionen	schwache Zustimmung in einer Konzession (X, Y, X') + <i>Holding Away</i> Geste + Absenken des Kopfes
	Mahnendes Einwenden	Einwand + turninitial erhobener Zeigefinger
Einfordern von Rückmeldungen	Einfordern von Rückmeldungen im Listenformat	Zugzwang + Zählgeste + Halten der Geste über Sprecher*innenwechsel hinweg
	Präsenhalten des Gesprächsgegenstandes in eingeforderten Rückmeldungen	Benennen eines Gesprächsgegenstandes + deiktische Geste auf Material + Halten der Geste über Sprecher*innenwechsel hinweg
Epistemische Modalisierung	Markieren von Evidenz durch Fallenlassen der Hand	epistemisch nicht abgeschwächte Äußerung + akustisches Signal einer zurückgeführten, fallengelassenen Geste (+ optional sprachliche Evidenzmarkierungen)
	Markieren von Evidenz durch <i>Palm Up</i> Gesten mit lateraler Bewegung	<i>Palm Up</i> Geste mit lateraler Bewegung + epistemisch verstärkende Modalisierung (z. B. Modalpartikeln, Satzadverbien)
	Markieren von Evidenz durch <i>Hand Shrug</i>	<i>Hand Shrug</i> + weitere <i>Shrug</i>-Komponenten (z. B. Schulterzucken, Abknicken des Kopfes, Hochziehen der Augenbrauen, <i>Mouth Shrug</i>)
	Markieren des ablehnenden Fallenlassen eines Gesprächsgegenstandes	responsive turninitiale <i>Dropping of Hand</i> Geste (i. d. R. + abgewandter Blick)
	Anzeige epistemischer Unsicherheit	formelhafte Phrase (z. B. <i>ich weiß es nicht</i>) + <i>Palm Up</i> Geste (z. T. mit lateraler Bewegung) (+ optional Schulterzucken, Kopfschütteln)

	Markieren von Unbestimmtheit	• sprachliche Markierung von Unbestimmtheit + vor/zurück rotierende <i>Palm Away</i> Geste bzw. iterative Vor-/Zurückbewegung einer <i>Palm Up</i> Geste

Tabelle 33: Übersicht über rekonstruierte multimodale Gestalten für das mündliche Argumentieren (inkl. Funktion und Beschreibung)

9 Erträge für eine multimodale Gattungsforschung

In dem vorangegangenen Kapitel 8 wurden die dyadischen argumentativen Entscheidungsfindungen daraufhin untersucht, welche Funktionen multimodale Ressourcen für das mündliche Argumentieren übernehmen (vgl. Kap. 5, Forschungsfrage I). Ausgehend von den Befunden der Studie wird nachfolgend zuerst beleuchtet, welche Bedeutung multimodale Gestalten für mündliches Argumentieren zukommt. Die Erträge der Studie für eine *multimodale* Gattungsforschung werden anschließend diskutiert. Das Kapitel schließt mit einer ersten Einordnung der gewonnenen Erkenntnisse für die dann folgende längsschnittliche Fallstudie (vgl. Kap. 10).

In der Untersuchung des Kapitel 8 konnten *wiederkehrende* Ressourcenverbünde ermittelt werden, die an der Realisierung argumentativer Handlungen beteiligt sind. Mit Mondada (2014b) wurden diese Ressourcenverbünde als komplexe multimodale Gestalten konzeptualisiert (vgl. auch Kap. 3.1.4). In den untersuchten argumentativen Entscheidungsfindungen wurde sichtbar, dass multimodale Gestalten vielfältige pragmatische Funktionen für das Argumentieren übernehmen und beispielsweise zur *Strukturierung* argumentativer Handlungen, zur *epistemischen Modalisierung*, sowie auch zur *interpretativen Rahmung* beitragen. Darüber hinaus werden sie auch zur *Interaktionsorganisation* eingesetzt. Ausgehend von diesen Befunden ist das Verhältnis zwischen multimodalen Gestalten und argumentativen Handlungen nun wie folgt zu verstehen: Die multimodalen Gestalten, die zur Realisierung der argumentativen Handlungen eingesetzt werden, führen zu einer spezifischen *Ausprägung* der jeweiligen argumentativen Handlung. Nicht immer sind die multimodalen Gestalten dabei an die Rolle der ‚Sprecher*innen‘ zurückgebunden. Die Untersuchung zeigt, dass multimodale Gestalten einerseits auch über Sprecher*innenwechsel hinweg gehalten werden (vgl. u. a. Kap. 8.2.2.1 & Kap. 8.5.1) und somit den *Turn* der Gesten äußernden Person überdauern. Andererseits werden multimodale Gestalten auch ‚*rezipient*innenseitig*‘ als Form der Verstehensdokumentation begleitend zu der Begründung des Gegenübers eingesetzt (vgl. Kap. 8.2.2.3). Es stellt sich daher die Frage, inwiefern auch *diese* multimodalen Gestalten zu einer spezifischen Ausprägung einer argumentativen Handlung beitragen. Abhängig von der sequenziellen und temporalen Entfaltung der Geste wurden zwei unterschiedliche analytische Vorgehen gewählt: Werden die Ressourcen einer multimodalen Gestalt noch in einer *eigenen* argumentativen Handlung der Gesten äußernden Person begonnen und anschließend gehalten, kann die multimodale Gestalt an die konkrete argumentative Handlung dieses*dieser Sprecher*in zurückgebunden werden (vgl. u. a. das *Etablieren und Präsenthalten eines verhandelten Gesprächsgegenstandes*). Im Rahmen der Verstehensdokumentation (vgl. Kap. 8.2.2.3) verhält es sich jedoch anders: Die multimodale Gestalt wird geäußert, während eine *andere* beteiligte Person eine argumentative Handlung aufbaut. In der Untersuchung wurde diese multimodale Gestalt der Handlung des*der aktuellen Sprecher*in zugeordnet, denn für eben diese argumentative Handlung hat die multimodale Gestalt eine interaktive Relevanz. Insgesamt handelt es sich bei den hier vorgenommenen Einordnungen der multimodalen Gestalten um eine analytische Entscheidung, die es ermöglicht darzustellen, im Rahmen welcher argumentativen Handlung die multimodale Gestalt hervorgebracht wird und auf welche Handlung sie somit auch interaktiv Einfluss nehmen kann. In Bezug auf die Frage, welche Bedeutung multimodale Gestalten für das Argumentieren übernehmen, wird durch die mikroanalytische Untersuchung sichtbar, dass multimodale Gestalten hochgradig funktional für das mündliche Argumentieren eingesetzt werden, indem sie insgesamt zu einer spezifischen Ausprägung argumentativer Handlungen beitragen.

Verlässt man nun die Ebene der Zusammenfassung konkreter Untersuchungsbefunde, erweist sich die Studie auch als ertragreich für eine *multimodale Gattungsforschung*: Bislang wurden mündliche Argumentationen vornehmlich hinsichtlich routinierter *sprachlicher* Handlungen untersucht, die Sprecher*innen einsetzen, um ein kommunikatives Problem zu lösen (vgl. Kap. 2.2). Mit der vorliegenden Untersuchung wird allerdings sichtbar, dass Interagierende, die sich in physischer Kopräsenz befinden, argumentative Handlungen mittels *multimodalen Gestalten* realisieren. Dabei sind körperlich-sichtbare Ressourcen ebenso wie sprachliche Ressourcen gemeinsam an der Bedeutungskonstitution (vgl. Kap. 3) und dem Aufbau argumentativer Handlungen beteiligt. Anhand des analytischen Zugriffs (vgl. Kap. 7) konnten dabei zuerst die jeweiligen argumentativen Handlungen identifiziert und zweitens die an ihrer Realisierung beteiligten Ressourcen beschrieben werden. Die detaillierte Untersuchung der zeitlichen Entfaltung und intra- sowie interpersonellen Koordinierung der multimodalen Ressourcen sowie der sequenziellen Einbettung dieser multimodalen Gestalt ermöglichte es, *gattungstypische*, d. h. argumentationsspezifische Funktionen multimodaler Ressourcenverbünde zu rekonstruieren. Die Befunde der Studie bekräftigen, dass für ein tiefergehendes Verständnis kommunikativer Gattungen körperlich-sichtbare Ressourcen zu integrieren sind. Untersuchungen, die ausschließlich auf die sprachliche Entfaltung von Gattungen wie Argumentationen fokussieren, können zwar u. a. sprachliche Handlungen rekonstruieren, die zur Lösung eines kommunikativen Problems beitragen, wie z. B. *dass ein*e Sprecher*in* einen Gesprächsgegenstand in einer Begründung etabliert. Eine anhaltende gestische Referenz auf den Gesprächsgegenstand und somit eine Verankerung des Bezugsgegenstandes in der fortgesetzten Interaktion, wie sie z. B. in Kapitel 8.2.2 gezeigt worden ist, gerät durch dieses Vorgehen jedoch nicht in den Fokus von Forschenden. Interaktive Konsequenzen, die durch die *konkrete multimodale* Realisierung argumentativer Handlungen evoziert werden, würden sich dem analytischen Blick somit entziehen. Für ein umfassendes Verständnis sowohl des Verlaufs einer Interaktion als auch der beteiligten gattungstypischen Handlungen scheint also der Einbezug multimodaler Ressourcen unerlässlich. Die vorliegende Arbeit leistet durch ihren multimodalen konversationsanalytischen Zugriff sowie die Fruchtbarmachung der Konzeptualisierung multimodaler Gestalten für die Untersuchung argumentativer Handlungen einen Beitrag dazu, den lange Zeit vorherrschenden Fokus der Gattungsforschung auf oberflächensprachlich verfestigte Strukturen um eine multimodale Perspektive zu ergänzen: Rekurrente multimodale Gestalten, die an der Realisierung argumentativer Handlungen beteiligt sind, können i. S. einer multimodalen Gattungsforschung als *gattungstypische verfestigte Strukturen* verstanden werden. Die sich in der Forschung bereits abzeichnenden Hinweise darauf, dass multimodale Gestalten für das mündliche Argumentieren funktional eingesetzt werden (vgl. Kap. 3.3), konnten durch die Untersuchung in Kapitel 8 deutlich konturiert und um eine systematische Beschreibung der multimodalen Gestalten ergänzt werden. Dabei ermöglichte der gewählte forschungsmethodische Zugriff insgesamt die gattungstypischen Funktionen multimodaler Ressourcen im *interaktiven* Vollzug mündlicher Argumentationen zu rekonstruieren.

Wendet man nun den Blick von den Erträgen der Befunde für eine multimodale Gattungsforschung hin zu den Erkenntnissen, die Kapitel 8 für die erwerbsbezogene Untersuchung des nachfolgenden Kapitels 10 mit sich bringt, wird Folgendes deutlich: In den Befunden der Studie zeigt sich, dass multimodale Gestalten auch funktional zur Strukturierung von Äußerungen (z. B. durch das multimodale Markieren von Kohäsion und Kohärenz) sowie zur Einpassung eines Beitrages in den interaktiven Kontext (z. B. durch das Anzeigen des Verwerfens in Begründungen) oder auch epistemischen Modalisierung eines Beitrages eingesetzt werden. Die in Kapitel 8 rekonstruierten multimodalen Gestalten werden in der erwerbsbezogenen Untersuchung der Fallstudie als *Markierungsformen* perspektiviert, die die Jugend-

lichen für das Vertexten, Kontextualisieren und epistemische Modalisieren von bzw. in Begründungshandlungen einsetzen (vgl. auch Kap. 4.1). In die Fallstudie fließen die dargelegten Befunde aus Kapitel 8 konsequent mit ein.

Das anschließende Kapitel 10 adressiert nun die erwerbsbezogene Forschungsfrage in einer explorativen Fallstudie (vgl. Kap. 5, Forschungsfrage II).

10 Längsschnittliche Fallstudien zu erwerbsbezogenen Veränderungen der multimodalen Gestalten für das Begründen

Die vorangegangenen Analysen (vgl. Kap. 8) widmeten sich der Frage, welche Funktionen multimodale Ressourcen in wiederkehrenden multimodalen Gestalten für argumentative Handlungen übernehmen. Diese Befunde sollen im Folgenden für die längsschnittlich angelegte Untersuchung zweier Fokuskinder genutzt werden. Die Forschungsfrage, *welche Veränderungen sich in einer längsschnittlich angelegten Studie zweier Fokuskinder im Alter von 12, 13 und 15 Jahren hinsichtlich der sprachlich-diskursiven Fähigkeiten sowie des Einsatzes multimodaler Ressourcen zeigen* (vgl. Kap. 5), steht dabei im Fokus.

Da die Auswahl der Fokuskinder der Fallstudien aufgrund ihrer unterschiedlichen diskursstrukturellen Profile, die sie im Alter von 15 Jahren zeigen, erfolgte (vgl. Kap. 7), werden diese Profile in einem ersten Schritt komprimiert vorgestellt (vgl. Kap. 10.1). Auf diesen Beschreibungen aufbauend wird dann in einem zweiten Schritt anhand der längsschnittlichen Daten rekonstruiert, welche *sprachlich-diskursive* Entwicklung die Fokuskinder durchlaufen. Dazu werden qualitative Analysen durchgeführt, die darlegen, welche sprachlichen sowie parasprachlichen Formen der Markierung zur Kontextualisierung, Vertextung sowie epistemischen Modalisierung nutzen (vgl. Kap. 10.2). In einem dritten Schritt wird dann der Blick auf die Entwicklung des *multimodalen* Diskurserwerbs geweitet. Dabei stehen im Fokus der Analysen potenzielle Veränderungen der multimodalen Gestalten, die an der argumentativen Handlung des Begründens (vgl. Kap. 8) beteiligt sind. Weil die Produktion von Äußerungspaketen, wie sie für das Begründen typisch ist, eine besondere Entwicklungsaufgabe darstellt, fokussiert die Untersuchung auf die Handlung des Begründens. Hierbei wird sichtbar, wie die Fokuskinder produktiv mit potenziell größeren Äußerungspaketen umgehen (vgl. Kap. 10.3). Die analytische Trennung, die hier zwischen Kapitel 10.2 und Kapitel 10.3 eingebracht wird, widerspricht auf den ersten Blick der Konzeptualisierung multimodaler Gestalten, die eine integrierte Untersuchung von Beginn an nahelegen. Jedoch geht der Einsatz körperlich-sichtbarer Ressourcen für das Begründen zeitweise zurück, so dass Kapitel 10.2 ermöglicht, Veränderungen, die sich sprachlich-diskursiv zeigen, mit zu erfassen. Aufbauend auf der deskriptiven Rekonstruktion der Ergebnisse kann anschließend diskutiert werden, welche Zusammenhänge sich im Hinblick auf die multimodalen sowie diskursstrukturellen Veränderungen zeigen (vgl. Kap. 10.4). Zuletzt werden aus diesen Befunden der längsschnittlichen erwerbsbezogenen Analysen weiterführende Hypothesen zu einem multimodalen Diskurserwerb abgeleitet (vgl. Kap. 10.5).

10.1 Diskursstrukturelle Profile

Für die Auswahl der Fokuskinder wurden in zwei analytischen Zwischenschritten das Gesamtkorpus hinsichtlich der kindseitigen Beteiligungsstruktur gesichtet und eine Vorauswahl von sechs Fokuskindern entlang der Interaktionsmuster (vgl. Kap. 6.1 & 7) vorgenommen. Die Daten des dritten Erhebungszeitpunktes wurden dann auf interindividuelle Unterschiede hin vergleichend untersucht. Entlang der drei Teilfähigkeiten von Argumentationskompetenz (Kontextualisierung, Vertextung, Markierung; vgl. Kap. 4.1.1) konnten zwei maximal kontrastive diskursstrukturelle Profile ermittelt werden. Die diskursstrukturellen Profile zeigen „empirisch Ausmaß und Art der Varianz in den mündlichen Diskursfähigkeiten der Kinder“ (Kern 2003: 36, Fußnote 5). Anhand der Profile erfolgte anschließend die Auswahl der Fokuskinder. Es handelt sich um die Profile *Explorative Position-Beziehen und Begründen* sowie *Zustimmendes Position-Beziehen und Bewerten*. Mit den Profilbezeichnungen werden Unterschiede hinsichtlich der Vertextungskompetenz deutlich, wenngleich sich Unterschiede auf allen drei

Ebenen der Argumentationskompetenz feststellen lassen. Diese werden nachfolgend vorgestellt und in Tabelle 34 zusammengefasst.

Auf Ebene der *Kontextualisierung* unterscheiden sich die Heranwachsenden hinsichtlich der Fähigkeit, globale Zugzwänge erkennen, bearbeiten und auch hervorbringen sowie selbstständig in die Jobs, wie beispielsweise den Kernjob des Begründens, hineinsteuern zu können (vgl. Tab. 34, Kontextualisierung). Während Kinder des Profils *Exploratives Position-Beziehen und Begründen* bereits Begründungshandlungen initiieren und mit globalen Zugzwängen rezeptiv und produktiv umgehen können, zeigen Kinder des Profils *Zustimmendes Position-Beziehen und Bewerten*, dass sie zwar selbstständig initierend argumentative Handlungen zum Bearbeiten der Aufgabe einbringen, wie z. B. das Beziehen einer Position oder Nennen einer Dis-/Präferenz gegenüber einer zur Auswahl stehenden Option, allerdings insgesamt deutlich seltener Begründungshandlungen initiieren.

Auf der Ebene der *Vertexung* zeigen sich Unterschiede hinsichtlich der Fähigkeit, global strukturierte Äußerungspakete aufbauen zu können, die beispielsweise erwartbar werden, wenn (lokale) Positionen nicht nur mit Bewertungen eingebracht, sondern auch plausibilisiert werden (vgl. Tab. 34, Vertexung). Während Kinder des Profils *Exploratives Position-Beziehen und Begründen* ihre benannten Positionen begründen, indem sie z. B. Szenarios zur Veranschaulichung einbringen sowie Konklusionen aus diesen ableiten, wird innerhalb des Profils *Zustimmendes Position-Beziehen und Bewerten* eine Tendenz der Heranwachsenden sichtbar, vermehrt satzförmige Zustimmungen sowie lokal strukturierte, satzförmige Begründungen (vgl. Kap. 2.2.2) aufzubauen.

Zuletzt zeichnen sich die Unterschiede der Vertexung und Kontextualisierung auch auf der Ebene der *Markierung* ab. So ist es Sprecher*innen des Profils *Exploratives Position-Beziehen und Begründen* möglich, Hypothetisches (i. S. hypothetischer Szenarios) z. B. durch die Verwendung des Konjunktivs II zu markieren sowie Kausalitäten und Konklusionen durch entsprechende Konnektoren, z. B. *weil*, *nämlich* bzw. *deshalb*, anzuzeigen. Bei Sprecher*innen des Profils *Zustimmendes Position-Beziehen und Bewerten* bleiben diese expliziten Markierungen von Kausalitäten oder hypothetischen Szenarios hingegen aus – eher finden sich Hinweise auf eingebrachte, subjektive Positionen z. B. durch *verba sentiendi* wie *ich finde* (vgl. Tab. 34, Markierung).

Die Ermittlung der diskursstrukturellen Profile ermöglichte die Auswahl zweier Fokuskinder, die auf den Ebenen der Diskurskompetenz äußerst unterschiedliche Fähigkeitsausprägungen zeigten. Die Entwicklung ihrer sprachlich-diskursiven Fähigkeiten wird in Kapitel 10.2 anhand der längsschnittlichen Daten altersvergleichend rekonstruiert.

Diskursstrukturelles Profil	Beschreibung	Teilfähigkeiten der Argumentationskompetenz
Exploratives Position-Beziehen und Begründen	Die Sprecher*innen kontextualisieren globale Zugzwänge als solche, steuern selbstinitiiert und nicht ausschließlich reaktiv in Begründungshandlungen. Sie bauen gattungsangemessen vermehrt global strukturierte Äußerungspakete auf, in denen sie u. a. Kausalitäten und Konklusionen markieren, konzедieren, Äußerungen modalisieren sowie in Begründungen und Einwänden z. B. hypothetische Szenarios explorieren.	Kontextualisierung • Erkennen & Bearbeiten globaler Zugzwänge • Setzen globaler Zugzwänge • Selbstständiges Hineinsteuern in Begründungshandlungen
		Vertextung • gattungsangemessener Aufbau lokal und global strukturierter Äußerungen • Einbringen von – Begründungen – Einwänden – Positionen inkl. Bewertungen
		Markierung • Markieren der eigenen Position mit <i>verba cogitandi/verba dicendi</i> (z. B. <i>ich denke, ich würde sagen</i>) • Markieren von Kausalitäten (z. B. <i>weil</i>), Hypothetischem (z. B. <i>man könnte</i>), Konklusionen (z. B. <i>dann</i>) und Konzessionen (z. B. <i>nicht X, aber Y</i>) • verstärkende und abschwächende Formen der Modalisierung (z. B. <i>bestimmt, vielleicht</i>)
Zustimmendes Position-Beziehen und Bewerten	Die Sprecher*innen erkennen globale Zugzwänge, bearbeiten sie allerdings vermehrt satzförmig. Sie bearbeiten initiiierend lokale Anforderungen, indem sie z. B. Position beziehen, aber nur vereinzelt autonom in Begründungshandlungen steuern. Sie bauen selten(er) global strukturierte Äußerungspakete auf.	Kontextualisierung • Erkennen lokaler & globaler Zugzwänge • vornehmlich lokale Bearbeitung von Zugzwängen • vereinzelt selbstständiges Hineinsteuern in Begründungshandlungen
		Vertextung • Aufbau vermehrt lokaler Äußerungen • Einbringen von – Positionen inkl. Bewertungen – (lokal strukturierten) Begründungen – Zustimmungen – Widersprüchen
		Markierung • Markieren der eigenen Position mit <i>verba sentiendi</i> (z. B. <i>ich finde</i>) • Markieren von Konzessionen (z. B. <i>zwar ... aber</i>) • verstärkende und abschwächende Formen der Modalisierung (z. B. <i>auf keinen Fall, eigentlich</i>)

Tabelle 34: Übersicht über diskursstrukturelle Profile der Fokuskinder

10.2 Sprachlich-diskursive Entwicklung der Fokuskinder

Im Anschluss an die vorgestellten diskursstrukturellen Profile werden im Folgenden die sprachlich-diskursiven Entwicklungen der ausgewählten Fokuskinder qualitativ über alle drei Erhebungszeitpunkte, d. h. im Alter von ca. 12, 13 und 15 Jahren, hinweg untersucht. Die Befunde des ersten Fokuskindes mit dem Profil *Exploratives Position-Beziehen und Begründen* (vgl. Kap. 10.2.1) sowie des zweiten Kindes mit dem Profil *Zustimmendes Position-Beziehen und Bewerten* (vgl. Kap. 10.2.2) werden nachfolgend vorgestellt. Dies ermöglicht, eine dezidierte Einordnung der sprachlich-diskursiven Entwicklung der Kinder vorzunehmen, die dann anhand einzelner Belegstellen nachvollziehbar gemacht wird. Die Untersuchung der sprachlich-diskursiven Entwicklung der Fokuskinder erfolgt anhand der in Kapitel 4.1.1 vorgestellten Modellierung von Argumentationskompetenz. Dabei werden die individuellen Entwicklungstendenzen der Fokuskinder auf den Ebenen der Kontextualisierung und Vertextung unter Berücksichtigung der Ebene der Markierung rekonstruiert. Da Erwerbsstudien des Argumentierens darüber hinaus darauf hinweisen, dass mit ansteigenden argumentativen Fähigkeiten Sprecher*innen auch vermehrt epistemische Modalisierungen z. B. zur Abschwächung oder Verstärkung einer Begründung nutzen (vgl. Kreuz und Luginbühl 2020; Morek 2020) und epistemischen Modalisierungen für das Argumentieren wichtige Funktionen zukommen (vgl. Kap. 8.6), werden diese ergänzend unter Berücksichtigung der Ebene der Markierung untersucht. Durch epistemische Modalisierungen wird nicht zuletzt der Modus des Argumentierens, i. S. eines explorierenden, gemeinsamen Verhandelns oder eines konfrontativen Einbringens von Positionen, geprägt (vgl. auch Mercer 2009; Ehlich 2014). Das Kapitel schließt mit einem interindividuellen Vergleich der Ergebnisse (vgl. Kap. 10.2.3). Anschließend werden in Kapitel 10.3 die Entwicklungen des Einsatzes der multimodalen Ressourcen beider Fokuskinder untersucht. Die Befunde beider Fallstudien werden in die abschließende Diskussion der Frage eingehen, welche Zusammenhänge zwischen multimodalen und diskursstrukturellen Entwicklungslinien identifiziert werden können (vgl. Kap. 10.4).

10.2.1 Fokuskind I – Exploratives Position-Beziehen und Begründen

Das nachfolgende Kapitel wird die sprachlich-diskursiven Fähigkeiten des Fokuskindes I mit dem Profil *Exploratives Position-Beziehen und Begründen* entlang der Ebenen der Kontextualisierung (vgl. Kap. 10.2.1.1), Vertextung (vgl. Kap. 10.2.1.2) und epistemischer Modalisierung (vgl. Kap. 10.2.1.3) im Altersverlauf von 12 bis 15 Jahren nachzeichnen. Abschließend werden die Befunde in einer sprachlich-diskursiven Entwicklungslinie des Fokuskindes I zusammengefasst (vgl. Kap. 10.2.1.4).

10.2.1.1 Kontextualisierung

Das Kind mit dem Profil *Exploratives Position-Beziehen und Begründen* zeigt im Alter von 12 Jahren, dass es sowohl mit lokalen als auch globalen Kontextualisierungsanforderungen umgehen kann und die angesteuerte kommunikative Gattung erkennt. Denn es initiiert bereits eigene Positionen, kommt sprachlich explizit markierten globalen Zugzwänge der Mutter nach (wie z. B. *warum nicht?*) und produziert im weiteren Verlauf Gesprächsbeiträge, die entsprechend der kommunikativen Gattung des Argumentierens Begründungen und Einwände beinhalten. Die Tochter kontextualisiert außerdem Äußerungen der Mutter, die keine explizit sprachliche Markierung eines Zugzwanges beinhalten, als eine Aufforderung, sich gegenüber der von der Mutter benannten Option zu verhalten, wie das folgende Beispiel zeigt.

Beispiel 1 – Kontextualisierung (12 Jahre)

```

159 MU: was ham [we NOCH] im ange[bot;]
160 TO: [hm- ]
161 [ja; ]
162 MU: die SPIElekonsole,
163 (0.8)
164 TO: öh:::-
165 naJA;=
166 =also is AUCH nich wirklich schlecht;
167 aber wir KÖNNten: (1.0) dann eher das de vau de paket (.) nehmen;=
168 =wei:l: (-) wir KÖNNten uns ja die wii von opa und oma ↑leihen,
169 und dann: (---) die DE vau des oder spIEle-

```

Das Anschließen eines konzедierenden, global strukturierten Äußerungspakets (Z. 164-169) zeigt an, dass die Tochter die globalen kommunikativen Anforderungen innerhalb der übergeordneten Diskurs-einheit des Argumentierens (er)kennt und sich für die Fortsetzung (hier im Umgang mit der präsentierten Option) verantwortlich sieht; sie initiiert selbstständig eine Begründung und ist in der gesamten Interaktion v. a. an der Bearbeitung des Kernjobs beteiligt. Die Bearbeitung der Randjobs, wie beispielsweise das Steuern in das Abschließen, übernimmt zu diesem Erhebungszeitpunkt die Mutter. Insgesamt zeigt die Sprecherin eine durchgängige Orientierung an der Bearbeitung des kommunikativen Problems, d. h. der argumentativen Entscheidungsfindung, und bearbeitet dieses gemeinsam mit der Mutter.

Im Alter von 13 Jahren zeigt sich nun ein anderer Umgang mit den globalen Kontextualisierungsanforderungen (vgl. Beispiel 2). Während die Tochter ein Jahr zuvor selbstständig nach einer Äußerung der Mutter, die keinen explizit sprachlich markierten globalen Zugzwang beinhaltet, eine Position einbringt und diese auch begründet, initiiert nun die Mutter einen unspezifischen Zugzwang (Z. 024), den sie, nach Ausbleiben einer Antwort der Tochter (Z. 025), zu einer Präferenzfrage verändert (Z. 027).

Beispiel 2 – Kontextualisierung (13 Jahre)

```

024 MU: na,=was MEINste denn;
025 (2.4)
026 MU: wird ja vielleicht (n/mit) mädchen in deinem Alter sein,=
027 =wo hättest DU denn lust hinzufahren;
028 (1.7)
029 TO: <<kopfschüttelnd>muSEum nicht;>
030 MU: und waRUM nicht,
031 TO: LANGweilig;

```

Die Tochter kommt nach einer Pause von 1.7 Sekunden dem Zugzwang der Mutter nach und disqualifiziert eine zur Auswahl stehende Option (Z. 029). Im anschließenden *Turn* der Mutter (Z. 030) wird erkennbar, dass in der Frage nach der Präferenz auch eine globale Anforderung enthalten ist. Denn die benannte (Dis-)Präferenz wird begründungspflichtig gemacht. Den globalen Zugzwang (Z. 030) bearbeitet die Tochter dagegen lokal in Form eines einzigen Lexems, das als Minimalbegründung dient (Z. 031: LANGweilig;). Im weiteren Verlauf der Interaktion zeigt die Tochter allerdings durchaus die Fähigkeit, den globalen kommunikativen Anforderungen der argumentativen Entscheidungsfindung produktiv nachzukommen, wie das folgende Beispiel illustriert.

Beispiel 3 – Kontextualisierung (13 Jahre)

```

059 TO: °h also ein m muSEum wär auf jeden fall langweilig;
060 da laufen EH schon ganz viele leute rum;
061 <<p>und wenn man dann mit SO vielen kommt->
062 und kann man au_NICHTS gemeinsam machen;
063 °hh ähm (.) KLEtterpark ginge noch-=
064 =aber da(nn) muss man immer so lange WArten-

```

065 das ist AUCH doof;

Hier initiiert sie, ohne dass ihrem Beitrag ein Zugzwang der Mutter vorausgeht, eine global strukturierte Begründung, die aus mehreren Turnkonstruktionseinheiten besteht. Auffällig in der gesamten Interaktion ist, dass die *selbstinitiierten* Äußerungspakete der Tochter global strukturiert sind (vgl. Beispiel 3). Richtet allerdings die Mutter globale Anforderungen an die Tochter (vgl. Beispiel 2), kommt die Tochter diesen ausschließlich in satzförmigen Beiträgen nach. Neben der Bearbeitung des Kernjobs ist die Tochter nun auch an dem Randjob des Abschließens beteiligt (vgl. Beispiel 4).

Beispiel 4 – Kontextualisierung (13 Jahre)

129 MU: vielleicht (.) nehmen wir dann DOCH lieber das abendessen;
 130 TO: <<kopfschüttelnd>nein;=lieber> das BOWlen;
 131 MU: das ↑BOWlen,
 132 TO: ja da kann man sich unterHALten,
 133 <<nickend>(.) (und) da ist JEder> an der rEihe und so;
 134 (1.4)
 135 MU: tja;
 136 soll mir RECHT sein;
 137 TO: okay;=
 138 MU: =wenn DU das möchtest;
 139 TO: dann sind wir (BOWlen/BOWling); ja,=
 140 MU: =DANN gehen wir [bow]len;
 141 TO: [(bow)]
 142 dAnn gehen wir BOWlen;
 143 (0.9)
 144 MU: wir gehen BOWlen;
 145 TO: ja;
 146 BOWlen;

In dem vorliegenden Beispiel zeigen sich beide Sprecherinnen verantwortlich für die Bearbeitung des Randjobs. In gewisser Weise konkurrieren sie hier um die endgültige Verantwortung, den Kernjob abzuschließen, indem sie die Option, die seitens der Tochter eingebracht (Z. 130) und der elternseitig zugestimmt worden ist (Z. 135-136, 138), in gegenseitiger Zustimmung insgesamt fünfmal wiederholen, bis schließlich die Tochter ein letztes Mal die Option wiederholt (Z. 142).

Im Alter von 15 Jahren kontextualisiert die Tochter nun eine Salve an Zugzwängen der Mutter als globalen Zugzwang zum Position-Beziehen und Begründen.

Beispiel 5 – Kontextualisierung (15 Jahre)

075 TO: <<p>wollen wir das so DURCHgehen,>
 076 und dann-
 077 (was);
 078 MU: ja;=was spricht daFÜR,
 079 was spricht daGEgen;
 080 was: (-) spricht DICH als erstes an;=
 081 =wo würdest DU denn gerne hinfah[ren;]
 082 TO: [ja-]=also ich würd nach ROM fahren,
 083 MU: <<nickend>ROM;>
 084 TO: mit dem (.) klasse,
 085 weil ich DENke:-
 086 dass nich (.) so VIEle leute in rom waren,
 087 MU: oKAY,
 088 TO: noch nich so viele leute,
 089 und_s is für die WICHTig,
 090 MU: hm_hm-
 091 TO: (hab ich auch schon) gehört,
 092 ZIEMlich wichtig-
 093 dass: (.) man IRgendwo hinfährt-

094 wo man noch NICH gewesen ist,
096 MU: <<nickend>hm_hm,>

Die Tochter kontextualisiert die Zugzwänge der Mutter (Z. 078-081) als einen globalen Zugzwang und bringt ein global strukturiertes Äußerungspaket hervor (Z. 082-094). Sie schließt an dieses erste Äußerungspaket selbstinitiiert weitere (nicht im Transkriptauszug enthaltene) global strukturierte Begründungen zu weiteren Optionen an, ohne dass die Mutter diese explizit einfordert. Neben Begründungen bringt die Tochter, der Gattung des Argumentierens angemessen, auch einen Widerspruch ein, der eine Rechtfertigung seitens der Mutter auslöst.

Die Tochter trägt außerdem zum Abschließen der Aushandlung bei, wie das nachfolgende Beispiel zeigt.

Beispiel 6 – Kontextualisierung (15 Jahre)

210 MU: ((schnalzt)) Okay.
211 (2.0)
212 würd ich auch vorschlagen;
213 das_s [für] ne GROSse klassengemeinschaft bestimmt °h (1.7) am
 spAnnendsten sei[n könnte;]
214 TO: [ja:;]
215 [SPRICHT] bestimmt nicht ALle an==
216 =aber: wahrsch vielleICHT die mEisten;
217 AUSserdem (-) <<p>will man ja auch etwas zusAmmen machen;>=
218 =<<pp>und-> (1.2)
219 jetzt NICHTS mit shoppen (-) so GROSSartig;
 ((...))
224 TO: s_is oKEE-
225 MU: GENau;

Die Tochter fügt der metadiskursiven Bewertung der Mutter (Z. 210-213), mit der bereits in das Abschließen hineingesteuert wird, expandierend eine Konzession (Z. 215-216) sowie eine bekräftigende Begründung (Z. 217-219) hinzu. Die angeschlossene Bewertung der Tochter (Z. 224) wird dann elternseitig ratifiziert und der Kernjob gemeinsam abgeschlossen. Insgesamt zeigt die Sprecherin im Alter von 15 Jahren die Fähigkeit, mit globalen Kontextualisierungsansforderungen rezeptiv und produktiv umgehen zu können und sich über den Kernjob hinaus auch an dem Job des Abschließens zu beteiligen.

Zusammenfassend zeigt sich in der sprachlich-diskursiven Entwicklung des Kindes über die drei Erhebungszeitpunkte hinweg, dass es bereits mit 12 Jahren rezeptiv und produktiv mit den lokalen und globalen Anforderungen der argumentativen Entscheidungsfindung umgehen kann. Diese Fähigkeiten baut die Sprecherin bis ins Alter von 15 Jahren weiter aus. Sie beteiligt sich dann ab dem Alter von 13 Jahren nicht nur an dem Kernjob Begründen, sondern auch am Abschließen der Argumentation.

10.2.1.2 Vertexung

Mit Blick auf die Vertexung strukturiert das Kind mit 12 Jahren globale Äußerungspakete wie z. B. Begründungen intern zusammenhängend und logisch-hierarchisch, wie das nachfolgende Beispiel zeigt, in dem ein Szenario der Veranschaulichung (vgl. Brünner und Gülich 2002b; Kinalzik 2024) der vorgeschlagenen Option dient.

Beispiel 1 – Vertextung (12 Jahre)

102 TO: ähm::;
 103 (2.3)
 104 ähm;;
 105 (1.3)
 106 ja;=ein wochenende in einem freizeits(.)PA:RK; ne?
 107 (---)
 108 is ja ganz SCHÖN;=
 109 =wir könnten ja dann nämlich in_n euROpa park fahren;=
 110 =denn da sehen wir SCHNEIders; ne? (.)
 111 MU: hm_hm,
 112 TO: s_WÄR ja auch ganz praktisch;
 113 (1.5)

In diesem Äußerungspaket benennt die Sprecherin zuerst die von ihr verhandelte Option als Bezugspunkt ihrer Äußerung (Z. 106). Sie evaluiert die Option und modalisiert die Bewertung eine Evidenzmarkierung sowie eine Abtönungspartikel (Z. 108) und verbindet diese dann mit einem anschließend eingebrachten Vorschlag zur Ausgestaltung der Option des Freizeitparks (Z. 109f.). Der Vorschlag enthält eine nachgestellte Kausalitätsmarkierung (*nämlich*, vgl. Gohl 2006: 304ff.), der ihn als Begründung erkennbar markiert. Ebenfalls wird die mögliche Ausgestaltung der Option durch die Verwendung des Konjunktivs II (Z. 109: *könnten*) und des Temporaladverbials *dann* als hypothetisch und kontrafaktisch markiert (Z. 109). Es schließt eine weitere Begründung an (Z. 110), die wiederum den Vorschlag für den Europapark näher erläutert und die durch die einleitende Kausalkonjunktion *denn* beide Äußerungsteile miteinander verbindet. Das global strukturierte Äußerungspaket schließt die Sprecherin dann mit einer konkludierenden Evaluation (Z. 112) ab.

Die Sprecherin zeigt darüber hinaus in den ausgebauten Begründungen die Verwendung von Bedingungssätzen innerhalb von veranschaulichenden Verfahren, wie das nachfolgende Beispiel 2 illustriert.

Beispiel 2 – Vertextung (12 Jahre)

264 TO: aber man könnte zum beispiel auch die spiele verKAufen,
 265 wenn man die gar nicht SPIElen will,
 266 MU: hm_hm,
 267 TO: un::d äh:m: (-) DANN kann man von dem gEld (.) in den freizeitpark fahren;

Dabei gelingt es der Sprecherin, zwei Folgen (Z. 264, 267) durch die Bedingung (Z. 265) als Gelenkstelle miteinander zu verbinden.

Vereinzelt zeigen sich auch Sequenzen, in denen die Sprecherin ihre Beiträge lokal strukturiert. Dies tritt beispielsweise auf, wenn Begründungen für die Ablehnung von Optionen eingeführt werden, wie das nachfolgende Beispiel zeigt.

Beispiel 3 – Vertextung (12 Jahre)

227 TO: ich GLAUB wir würden den-
 228 aber den hat man sofort verBALLert;
 229 den gutschein ZWEI;
 230 (1.8)
 231 MU: das STIMMT;

Die Sprecherin markiert explizit sprachlich mit dem adversativen Junktor *aber*, dass nun ein Gegensatz zu ihrer ursprünglichen, abgebrochenen Position (Z. 227) erfolgt und verweist dadurch zugleich auf die Verbindung beider Äußerungselemente. Die als Begründung eingebrachte Proposition (Z. 228f.) verbleibt allerdings satzförmig und wird nicht weiter elaboriert.

Darüber hinaus reiht die Sprecherin einzelne Äußerungen auch unverbunden aneinander, ohne dass es gelingt, sie in eine globale Struktur zu überführen, wie Beispiel 4 illustriert.

Beispiel 4 – Vertextung (12 Jahre)

248 TO: [und] DANN sehen wir schnEIders;
 249 viel[leicht;]
 250 MU: [tja;]
 251 (2.7)
 252 TO: aber papa muss ja auch FREI haben;
 253 <<creaky>oder so;>
 254 aber is ja am WOchenende; (.)
 255 MU: ja.
 256 (--)
 257 TO: du musst ja auch [ZEIT] haben;
 258 MU: [in-]
 259 in ORdnung;

Die Sprecherin bringt im Anschluss an ihre Begründung (Z. 248), die aus einer Folge besteht, drei selbstständige Turnkonstruktionseinheiten ein, die unverbunden aufeinander folgen: Sie wendet erst satzförmig gegen ihre eigene Begründung ein (Z. 252), entkräftet den Einwand dann (Z. 254) und schließt zuletzt eine begründende Expansion für ihren Einwand an (Z. 257) – ebenfalls satzförmig und ohne weitere Elaboration. Mit Quasthoff et al. (2019: 215) werden diese Sequenzen als satzübergreifende Äußerungen verstanden, in denen es den Sprecher*innen noch nicht gelingt, einzelne Äußerungen zu einem global strukturierten Äußerungspaket zu verbinden.

Wenngleich die letzten beiden Beispiele (vgl. Beispiel 3 & 4) auf lokal strukturierte Beiträge der Sprecherin fokussieren, zeigt sie in der argumentativen Entscheidungsfindung insgesamt, dass sie ihre Beiträge sowohl lokal als auch global strukturieren kann (vgl. Beispiel 1 & 2).

Im Alter von 13 Jahren sind die Äußerungspakete, die die Tochter selbst initiiert und die nicht durch einen globalen Zugzwang der Mutter implizit oder explizit eingefordert werden, global strukturiert und zeigen in ihrer internen Struktur eine gewisse Musterhaftigkeit der Vertextung, wie das nachfolgende Beispiel 5 darstellt: Die Tochter benennt zuerst das Bezugsobjekt, auf das sie sich bezieht, bewertet es und plausibilisiert schließlich die Bewertung. Z. T. schließt sie den Beitrag dann mit einer erneuten, i. S. einer rahmenden Bewertung ab.

Beispiel 5 – Vertextung (13 Jahre)

059 TO: °h also ein m muSEum wär auf jeden fall langweilig;
 060 da laufen EH schon ganz viele leute rum;
 061 <<p>und wenn man dann mit SO vielen kommt->
 062 und kann man au NICHTS gemeinsam machen;
 063 °hh ähm (.) KLEtterpark ginge noch-=
 064 =aber da(nn) muss man immer so lange WArten-
 065 das ist AUCh doof;
 066 °hhh Abendessen (.) ist auch voll langweilig-=
 067 =weil man sitzt da nur neben EINem,
 068 kann sich nur mit dem NACHbarn unterhalten-
 069 und das sind wahrscheinlich die: (.) familien also aus: (.) den
 familien SELber;
 070 das wär dann auch BLÖD;

071 TO: [(und) BOWling wär dann das beste;]

Die Beiträge können sowohl durch die oben beschriebene, sich wiederholende Musterhaftigkeit als zusammengehörig identifiziert als auch durch sprachlich explizite Markierungen, wie dem Adverb *auch*, mit dem die Sprecherin anzeigt, dass die ablehnenden Bewertungen aneinander anschließen (Z. 065, 066, 070: das ist AUCH doof; / °hhh Abendessen (.) ist auch voll langweilig- / das wär dann auch BLÖD;). Das gesamte Äußerungspaket wird durch seine interne Strukturierung als zusammengehörig erkennbar. In der Struktur der einzelnen Begründungen markiert die Sprecherin außerdem Fortsetzungen durch Konjunktionen (Z. 061-062: und wenn man [...] / und kann man), Kausalitäten im Anschluss an Bewertungen mit Subjunktionen (Z. 067: weil) und sie projiziert Einwände durch konzедierende Strukturen, in denen die Verwendung des Konjunktivs II einen anschließenden Einwand erwartbar macht (Z. 063-064: ginge noch-= / =aber).

Macht jedoch die Mutter z. B. durch einen Einwand eine Reaktion der Tochter erwartbar, bearbeitet die Tochter diese Anforderungen v. a. satzförmig bzw. in aneinandergereihten Äußerungen und nicht in größeren Äußerungspaketen, wie das folgende Beispiel 6 zeigt.

Beispiel 6 – Vertextung (13 Jahre)

071 TO: [(und) BOWling wär dann das beste;]
 072 MU: [aber es GEHT ja darum doch dass man-]
 073 dass man sich da KENnenlernt;
 074 TO: ja;=GEHT ja nicht;
 075 <<p>beim Abendessen kann man ja NIX gemeinsam machen;>
 076 (0.8)

Als die Mutter gegen die Reihe an Begründungen der Tochter einwendet (Z. 072), macht sie zugleich eine Begründung oder Rechtfertigung seitens der Tochter erwartbar. Die Tochter jedoch widerspricht zuerst (Z. 074), indem sie i. S. des *Format Tyings* (vgl. Goodwin und Goodwin 1987) den Einwand der Mutter (Z. 072: aber es GEHT ja darum doch) aufnimmt und durch eine Negation (Z. 074: ja;=GEHT ja nicht;) die ursprüngliche Satzaussage ins Gegenteil verkehrt. Die dann folgende Begründung, in der die Tochter eigenständig Bezug auf die von der Mutter favorisierte Option (Abendessen) nimmt, besteht nun aus einer satzförmigen Proposition, deren Inhalt durch die Verwendung der Modalpartikel *ja* als evident bzw. offensichtlich markiert wird (vgl. Reineke 2016). Eine weitere Elaborierung der Proposition erfolgt an dieser Stelle nicht. Insgesamt zeigt sich in den lokal organisierten Begründungen, Einwänden und Widersprüchen, dass die Tochter durch Praktiken wie das *Format Tying* lokale Bezüge zwischen den *Turns* beider Sprecherinnen herstellt sowie durch Modalpartikeln Äußerungsinhalte als fraglos evident markiert. Die v. a. reaktiven, opponierenden und meist satzförmig strukturierten Beiträge der Tochter tragen letztlich zu einer konfrontativen Rahmung der gemeinsamen Entscheidungsfindung bei.

Zum dritten Erhebungszeitpunkt, an dem die Tochter 15 Jahre alt ist, baut sie globale Äußerungspakete bestehend aus mehreren *Turns* logisch-hierarchisierend auf, wie nachfolgend exemplarisch anhand von Beispiel 7 gezeigt wird.

Beispiel 7 – Vertextung (15 Jahre)

082 TO: [ja-]=also ich würd nach ROM fahren,
 083 MU: <<nickend>ROM;>
 084 TO: mit dem (.) klasse,
 085 weil ich DENke:-
 086 dass nich (.) so VIEle leute in rom waren,

087 MU: oKAY,
 088 TO: noch nich so viele leute,
 089 und_s is für die WICHtig,
 090 MU: hm_hm-
 091 TO: (hab ich auch schon) gehört,
 092 ZIEMlich wichtig-
 093 dass: (.) man IRgendwo hinfährt-
 094 wo man noch NICH gewesen ist,
 096 MU: <<nickend>hm_hm,>

Die Tochter bezieht zu Beginn des Äußerungspakets eine Position (Z. 082, 084), in der sie durch den Diskursmarker *also* den Anschluss eines größeren Äußerungspakets projiziert. Es folgt nun ein erstes Argument in Form einer satzübergreifenden Begründung, die mit der kausalen Konjunktion *weil* an die Position angeschlossen und mit einem *verbum cogitandi* (Z. 085: ich DENke) sowie einem Ergänzungssatz (Z. 086) fortgesetzt wird. Nach einer Ratifizierung der Mutter (Z. 087) baut die Tochter ihre Äußerung im Rahmen einer Reparatur weiter aus (Z. 088) und bringt dann ein zweites Argument ein: Durch die Konjunktion *und* (Z. 089) wird hier an das erste Argument eine Evaluation (Z. 089: WICHtig) angeschlossen, die im Rahmen einer Korrektur durch das nun hinzutretende fokusakzentuierte Adverb *ziemlich* verstärkt wird (Z. 092: ZIEMlich wichtig).⁹⁷ Diese Evaluation wird zugleich einer unspezifischen externen Autorität zugesprochen (Z. 091: (hab ich auch schon) geHÖRT,) und dann in einem Ergänzungssatz fortgeführt (Z. 093). Insgesamt schafft die Sprecherin in der Vertextung ihrer Äußerung eine globalstrukturelle Kohärenz, u. a. indem sie durch *verba cogitandi* (Z. 085: ich DENke:) und Evaluationen (Z. 089: und_s is für die WICHtig,) Ergänzungssätze projiziert und anschließend einlöst; sowie außerdem in Reparaturen Lexeme aus vorangegangenen Äußerungen wieder aufgreift und so ihre Zusammengehörigkeit markiert. Auch die weiteren Beiträge der Sprecherin, mit Ausnahme eines lokal operierenden Widerspruchs, in dem sie die Bezugnahme auf den vorangegangenen *Turn* durch ein *Format Tying* (vgl. Goodwin und Goodwin 1987) markiert, sind in ähnlicher Weise logisch-hierarchisch strukturiert.

In der Entwicklung der Sprecherin auf der Vertextungsebene wird ersichtlich, dass sie im Alter von 12 Jahren bereits ihre Beiträge global strukturieren kann, vereinzelt aber die Verknüpfung von satzförmigen Turnkonstruktionseinheiten zu einem Äußerungspaket noch schwerfallen. Die Fähigkeit baut sie bis zum Alter von 15 Jahren deutlich aus. Über alle Erhebungszeitpunkte hinweg gelingt es der Sprecherin allerdings, in größeren Äußerungspaketen logisch-hierarchische Beziehungen zwischen den Äußerungselementen herzustellen, u. a. durch Bedingungsgefüge (vgl. Beispiel 2 (12 Jahre)), Kausalitätsmarkierungen (vgl. Beispiel 1 (12 Jahre); Beispiel 5 (13 Jahre); Beispiel 7 (15 Jahre)), projizierte Konzessionen (vgl. Beispiel 5 (13 Jahre)) oder auch projizierte Ergänzungssätze (vgl. Beispiel 7 (15 Jahre)).

10.2.1.3 Epistemische Modalisierung

Auf der Ebene der epistemischen Modalisierung nutzt die Sprecherin im Alter von 12 Jahren hauptsächlich Modalisierungen, die die epistemische Gewissheit verstärken. So verwendet sie beispielsweise beim Beziehen einer Position, die sie im Indikativ einbringt, Intensitäts-Adverbien – während sie gleichzeitig durch den Konjunktiv anzeigt, dass Hypothetisches bzw. Zukünftiges verhandelt wird, und sie dadurch eine explorierende, tentative Annäherung an die Entscheidung markiert (vgl. Beispiel 1; vgl. auch Morek 2020).

⁹⁷ Weinrich (2003: 594) beschreibt das Intensitäts-Adverb *ziemlich* in seiner abschwächenden Funktion, das anzeigt, „daß eine Eigenschaft nicht in vollem Maße ausgeprägt ist“. In dem abgebildeten Transkriptausschnitt tritt das Adverb allerdings in einer Wiederaufnahme der Bewertung (Z. 090: und_s is für die WICHtig,) nun fokusakzentuiert verstärkend zu der Bewertung hinzu (Z. 093: ZIEMlich wichtig-).

Beispiel 1 – Epistemische Modalisierung (12 Jahre)

075 MU: [hast du_n vor][schlag?]
 076 TO: [gut]schein eins würd ich halt nun mal
 <<lachend> !GAR! nicht nehmen;>

Die Ablehnung von „gutschein eins“ (Z. 076) wird von der Sprecherin durch das besonders stark akzentuierte Intensitäts-Adverb „!GAR!“ sowie durch die Modalpartikel „halt“ (Z. 076) epistemisch verstärkt (vgl. zu *halt* auch Thurmair 1989: 125). Gerade in Begründungen zeigt sich dann eine besonders häufige Verwendung von Evidenzmarkierungen durch die Modalpartikel *ja*, mit der die Sprecherin den Geltungsanspruch eines Arguments erhöht (vgl. Reineke 2016).

Beispiel 2 – Epistemische Modalisierung (12 Jahre)

106 TO ja;=ein wochenende in einem freizeits(.)PA:RK; ne?
 107 (---)
 108 is ja ganz SCHÖN;=
 109 =wir könnten ja dann nämlich in_n euROpa park fahren;=
 110 =denn da sehen wir SCHNEIders; ne? (.)
 111 MU: hm_hm,
 112 TO: s_WÄR ja auch ganz praktisch;

In dem vorliegenden Beispiel markiert die Sprecherin sowohl ihre einleitende wie auch ihre abschließende Bewertung (Z. 108, 112) und das erste Begründungssegment (Z. 109) als fraglos evident. Nur einmal nutzt sie in der gesamten argumentativen Entscheidungsfindung das *verbum putandi* (*ich glaub*), um einen begonnenen Vorschlag für eine Entscheidung abzuschwächen. Insgesamt zeigt sich, dass die Sprecherin durch die Verwendung des Konjunktivs einerseits markiert, dass ein zukünftiges, hypothetisches Ereignis aufgebaut wird, das explorativ verhandelt werden kann. Andererseits modalisiert sie aber Positionen, Bewertungen in Begründungen etc. wiederkehrend epistemisch verstärkt und markiert sie u. a. als fraglos evident.

Im Alter von 13 Jahren nutzt die Sprecherin eher wenige explizit sprachliche epistemisch abschwächende oder verstärkende Modalisierungen. Vielmehr erweist sich der Modus durch das Ausbleiben abschwächender Modalisierungen und die frequente Verwendung des Indikativs als wenig explorierend, wie das nachfolgende Beispiel 3 exemplarisch illustriert.

Beispiel 3 – Epistemische Modalisierung (13 Jahre)

059 TO: °h also ein m muSEum wär auf jeden fall langweilig;
 060 da laufen EH schon ganz viele leute rum;
 061 <<p>und wenn man dann mit SO vielen kommt->
 062 und kann man au_NICHTS gemeinsam machen;

Die Bewertung, die die Sprecherin einbringt, wird durch die Verwendung des Konjunktivs abgeschwächt (Z. 059: *wär*) und zugleich durch die Phase wie *auf jeden Fall* (Z. 059) epistemisch verstärkt modalisiert. Anders als ein Jahr zuvor wird das dann anschließende Szenario allerdings nicht mehr als mögliches, hypothetisches Ereignis markiert, sondern durch die Verwendung des Indikativs als Faktum gesetzt. Dieser Modus durchzieht wiederkehrend die Beiträge der Sprecherin zu diesem Erhebungszeitpunkt.

Im Alter von 15 Jahren verwendet die Sprecherin *verba cogitandi* (*ich denke*), *verba putandi* (*ich glaube*) und *verba dicendi* + Konjunktiv (*ich würde sagen*) zur modalisierenden Abschwächung in Begründungen und zur Markierung eines explorierenden Modus der Begründungen (vgl. Beispiel 4).

Beispiel 4 – Epistemische Modalisierung (15 Jahre)

133 TO: den bungalowpark würde ich (hier/eh) NICHT nehmen;
134 (.) glaub das is so was ähnliches wie MOVIEpark;

Die Sprecherin schließt hier durch einen Vergleich mit einem Freizeitpark (Z. 134) eine Option begründet aus und markiert dabei durch das *verbum putandi* (*ich*) glaub eine Unsicherheit in der Angemessenheit des Vergleichs. Neben den o. g. *verba sentiendi* verwendet die Sprecherin insgesamt aber wenige andere Formen der modalisierenden Abschwächung oder Verstärkung.

Hinsichtlich der Entwicklung über alle drei Erhebungszeitpunkte hinweg wird erkennbar, dass die im Alter von 12 Jahren noch deutlich frequent auftretenden Evidenzmarkierungen bzw. epistemischen Verstärkungen sich im Alter von 13 Jahren und 15 Jahren als rückläufig erweisen. Ferner zeigen sich in den Altersstufen jeweils unterschiedliche Modi, mit denen die Äußerungsinhalte als Fakt bzw. hypothetisch gerahmt werden: Durch die häufige Verwendung des Indikativs markiert die Sprecherin im Alter von 13 Jahren die Äußerungsinhalte als wenig verhandelbar. Im Alter von 12 Jahren hingegen schwächt sie ihre Äußerungen modalisierend durch den Konjunktiv II und im Alter von 15 Jahren dann durch *verba sentiendi* ab und markiert einen explorierenden Modus.

10.2.1.4 Zusammenfassende Beschreibung der sprachlich-diskursiven Entwicklungslinie

In der längsschnittlichen Untersuchung des diskursstrukturell stärkeren Kindes mit dem Profil *Explorative Position-Beziehen und Begründen* zeigen sich im Altersvergleich von 12 bis 15 Jahren Veränderungen der sprachlich-diskursiven Entwicklungslinie auf den Ebenen der Kontextualisierung, Vertextung und epistemischen Modalisierung, die nachfolgend zusammenfassend vorgestellt werden.

Kontextualisierung

- Die Sprecherin erkennt bereits im Alter von 12 Jahren lokale und globale Zugzwänge und kann auch mit nicht explizit-sprachlich markierten globalen Anforderungen der Interaktion rezeptiv und produktiv umgehen. Dabei ist sie an dem Kernjob des Begründens beteiligt; die Randjobs werden hingegen elternseitig bearbeitet.
- Im Alter von 13 Jahren bearbeitet die Sprecherin dann die elternseitig gesetzten globalen Anforderungen ausschließlich lokal; allerdings operieren die *selbstinitiierten* Beiträge auf globaler Ebene. Die Sprecherin ist insgesamt sowohl an der Bearbeitung des Kernjobs Begründen als auch an dem Randjob des Abschließens beteiligt.
- Mit 15 Jahren dann zeigt die Sprecherin einen sicheren rezeptiven und produktiven Umgang sowohl mit lokalen als auch globalen Anforderungen der Interaktion. Sie beteiligt sich nicht nur am Kernjob des Begründens, sondern auch am Randjob des Abschließens.

Vertextung

- Im Alter von 12 Jahren gelingt es der Sprecherin bereits, global strukturierte Äußerungspakete logisch-hierarchisch aufzubauen und z. B. anhand von Szenarios zu veranschaulichen. Sie setzt dabei explizit-sprachliche Markierungen ein. Vereinzelt werden satzförmige Äußerungen aneinandergereiht und noch nicht zu einem global strukturierten Äußerungspaket verbunden.

- Mit 13 Jahren weisen die selbstinitiierten Beiträge der Sprecherin eine globale Strukturierung auf. Der Sprecherin gelingt es, nacheinander verhandelte Optionen in der immer gleichen Strukturierung zu bearbeiten und aufgrund dieser Musterhaftigkeit die Beiträge zu einem großen Äußerungspaket zu verbinden. Dabei projiziert sie in dem Äußerungspaket beispielsweise in Konzessionen anschließende Einwände. Beiträge hingegen, die durch Zugzwänge elternseitig eingefordert werden, sowie opponierende Handlungen werden i. d. R. satzförmig strukturiert.
- Im Alter von 15 Jahren baut die Sprecherin ihre Beiträge – mit Ausnahme von Widersprüchen – logisch-hierarchisch zu global strukturierten Äußerungspaketen auf, die sie kohärent durch explizite sprachliche Markierungen, wie z. B. das Projizieren und Einlösen von Ergänzungssätzen, miteinander verbindet.

Epistemische Modalisierung

- In der Aushandlung eines potenziell zukünftigen Ereignisses markiert die Sprecherin im Alter von 12 Jahren dieses Ereignis durch die Verwendung des Konjunktivs II als hypothetisch und baut veranschaulichende Szenarios auf. Damit markiert die Sprecherin einen explorativen Zugang zur gemeinsamen Aushandlung. Zugleich modalisiert sie aber auch eingebrachte Positionen, Bewertungen in Begründungen etc. wiederkehrend verstärkend als fraglos evident.
- Mit 13 Jahren verwendet die Sprecherin kaum noch verstärkende oder abschwächende Modalisierungen. Der Modus ihrer Beiträge ist durch die häufige Verwendung des Indikativs geprägt, durch den die Äußerungsinhalte als Fakt gerahmt werden – ein explorierender Modus wie im Alter von 12 Jahren wird hier nicht erkennbar.
- Im Alter von 15 Jahren verwendet die Sprecherin ähnlich wenige epistemische Modalisierungen wie mit 13 Jahren. Allerdings dienen die Formen, die sie einsetzt, einem explorierenden Modus, denn die Sprecherin rahmt durch *verba sentiendi* ihre Beiträge als subjektive – und nicht allgemeingültige – Positionen und Begründungen.

Im Anschluss an die Untersuchung der sprachlich-diskursiven Entwicklungslinie des ersten Fokuskindes, wird im Folgenden Kapitel 10.2.2 die sprachlich-diskursive Entwicklung von Fokuskind II nachvollzogen.

10.2.2 Fokuskind II – Zustimmendes Position-Beziehen und Bewerten

Das nachfolgende Kapitel wird, analog zum Aufbau des vorangegangenen Kapitels, die sprachlich-diskursiven Fähigkeiten des Fokuskindes II mit dem Profil *Zustimmendes Position-Beziehen und Bewerten* entlang der Ebenen der Kontextualisierung (vgl. Kap. 10.2.2.1), Vertextung (vgl. Kap. 10.2.2.2) und der epistemischen Modalisierung (vgl. Kap. 10.2.2.3) im Alter von 12, 13 und 15 Jahren nachzeichnen. Die Befunde der sprachlich-diskursiven Entwicklung des Fokuskindes II werden in einer abschließenden Zusammenfassung dargelegt (vgl. Kap. 10.2.2.4).

10.2.2.1 Kontextualisierung

Hinsichtlich der Kontextualisierung zeigt Fokuskind II mit dem diskursstrukturellen Profil *Zustimmendes Position-Beziehen und Bewerten*, dass es zum ersten Erhebungszeitpunkt im Alter von 12 Jahren die elternseitigen globalen Zugzwänge nicht bzw. nur lokal bearbeitet. Dabei beruft sich das Kind metakommunikativ auf die eigene Unwissenheit (z. B. durch *ich weiß nicht so genau, (ich) bin mir nicht sicher*) und bearbeitet dabei auch Erneuerungen globaler Zugzwänge ausschließlich mit dem Zucken der Schultern. Allerdings wird erkennbar, dass die Tochter grundsätzlich rezeptiv mit den Zugzwängen des Elternteils umgehen kann, ihr aber der Anschluss globaler Beiträge schwerfällt, wie in der Bearbeitung des globalen Zugzwangs im Folgenden (vgl. Beispiel 1) deutlich wird.

Beispiel 1 – Kontextualisierung (12 Jahre)

164 MU: [und jetzt] hier gutschein DREI?
 165 waRUM nich?
 166 (1.3)
 167 TO: ähm;
 168 weil wir das nich BRAUchen,
 169 (---)
 170 MU: NEIN;
 171 weil EINFach;
 172 KUMma;=
 173 =wenn WIR das nehmen;
 174 also diese SPIElekonsole;
 175 TO: außerdem BRAUchen wir_s [(auch) nicht;]
 176 MU: [dann-]
 177 erst_mal BRAUchen wir_s nich;=

Die Tochter kann rezeptiv mit dem Zugzwang umgehen, denn sie markiert ihre auf den Zugzwang folgende Äußerung, entsprechend der Anforderung des globalen Zugzwangs, durch die kausale Konjunktion *weil* als Begründung (Z. 168). Die Bearbeitung des Zugzwangs bleibt allerdings satzförmig (vgl. auch 10.2.2.2, Beispiel 2). Kurz darauf wiederholt die Sprecherin eigenaktiv initiiierend die satzförmige Begründung (Z. 175), nachdem die Mutter diese Begründung zurückgewiesen (Z. 170), platziert dabei die Wiederholung ihrer Begründung allerdings im *Turn* der Mutter und somit an einer sequenziellen Position, an der inhaltlich noch keine turnübergaberelevante Stelle erreicht ist. Darüber hinaus markiert die Heranwachsende ihre Äußerung zwar durch das Adverb *außerdem* als eine Expansion zur elternteiligen Äußerung, allerdings hat die Mutter zu diesem Zeitpunkt noch kein Argument eingebracht, an das die Begründung der Tochter inhaltlich anschließen könnte – die sequenzielle Einpassung des Beitrags gelingt der Sprecherin an dieser Stelle noch nicht. In der Interaktion zeigt sich jedoch, dass argumentative Handlungen, in denen gegenläufige Positionen zur Mutter angezeigt werden, wie z. B. beim Widersprechen, von der Tochter selbstinitiiierend und sequenziell passend eingebracht werden können. Insgesamt werden die kommunikativen Aufgaben von der Mutter übernommen, denn sie setzt globale Zugzwänge zum Begründen und steuert auch in den Job des Abschließens hinein. Die Tochter kann zwar rezeptiv mit den Kontextualisierungsanforderungen umgehen, aber die Bearbeitung der Zugzwänge bereiten ihr im produktiven Umgang noch Schwierigkeiten.

Im Alter von 13 Jahren gelingt der Tochter das Erkennen globaler Anforderungen nun partiell. Während sie im nachfolgenden Transkriptausschnitt (vgl. Beispiel 2) einen Zugzwang zur Präferenzbekundung lokal kontextualisiert, weist das darauffolgende Beispiel (vgl. Beispiel 3) diese Schwierigkeiten nicht auf – allerdings werden hier Probleme im Äußerungsaufbau sichtbar, die dann die inhaltliche Passung des Beitrags erschweren.

Beispiel 2 – Kontextualisierung (13 Jahre)

046 MU: °hhh hhh° was mEInst was was würdest du du denn auf ANhieb (-) gUt
 finden;
 047 (2.5)
 048 MU: hm?
 049 TO: ich WEISS es gerade nicht;
 050 (.) ich überLEG (-) schon; (1.4) (---)
 051 ich würd NICHT unbedingt kletterpark nehmen;
 052 MU: nee weil der dann da: (.) könnt ihr euch ja nicht unterHAL[ten oder-]
 053 TO: [ja;=des]
 WEgen ja;
 054 (2.0)
 055 also DA würde man ja nichts wirklich über (.) jemanden erfahren oder
 sowas;=ne,

Im Anschluss auf einen von der Mutter gesetzten Zugzwang (Z. 046), der nach einer Präferenz der Tochter fragt, folgt erst eine 2,5-sekündige Pause, bevor die Tochter dann den *Turn* übernimmt. Die Tochter liefert einen *Account* (vgl. Garfinkel 1967) für die zuvor ausgebliebene Turnübernahme, indem sie auf ihre epistemische Unsicherheit verweist (Z. 049). Es folgt ein metadiskursiver Kommentar (Z. 050) und erst dann benennt die Tochter eine Dispräferenz, indem sie Position bezieht (Z. 050: ich würd NICHT unbedingt kletterpark nehmen;). Dass die elternseitige Präferenzabfrage auch eine Begründung für diese erwartbar macht, zeigt sich in dem darauffolgenden *Turn* der Mutter (Z. 052), die zustimmend die kindseitige Ablehnung der Option elaboriert. Erst daran anschließend erfolgt nun auch eine Begründung der Tochter, die den *Turn* der Mutter eigenaktiv und ko-konstruierend (Z. 055) fortsetzt. Dass die Frage nach der persönlichen Präferenz im Rahmen eines argumentativen Entscheidungsdiskurses eine Plausibilisierung erfordert und somit der scheinbar lokale Zugzwang auch globale Anforderungen enthält, wird von der Tochter erst mit Unterstützung des Elternteils erkannt.

In dem nachfolgenden Beispiel 3 kontextualisiert die Tochter die globale Anforderung in der Äußerung der Mutter und kommt dieser durch den Aufbau eines satzübergreifenden Äußerungspakets nach. Allerdings führen Strukturierungsschwierigkeiten im Aufbau der Äußerung hier zu inhaltlichen Passungsproblemen des Beitrags.

Beispiel 3 – Kontextualisierung (13 Jahre)

059 MU: also ich dEnke mal (.) also wenn man sich nicht so KENNT-
 060 (1.0)
 061 um: (.) sich besser KENnenzulernen;
 062 ist einfach äh (1.3) ja (2.3);
 063 °hh BOWling kann man au:ch;=
 064 =aber (0.8) DA: (-) unterhält man sich wahrscheinlich auch nicht
 so;=oder?
 065 TO: °h ja doch bei bowling da muss man ja auch immer WARTen;
 066 also EIner °h mAcht ja sO,
 067 dann guckt man dann wieder zu==
 068 =und dann wenn man (.) WARTet dann redet man ja auch miteinander und so
 wahrscheinlich;
 069 ich WEISS es nicht so genau;
 070 (1.4)
 072 MU: tja;
 073 (2.0)

Die Äußerung der Mutter, die inhaltlich Unsicherheiten in Bezug auf die Frage aufweist, wie geeignet die Aktivität des Bowlens für ein Kennenlernen ist, schließt mit dem turnfinalen, hoch steigenden *question tag oder* (Z. 064; vgl. König 2020) und zeigt an, dass von der Tochter eine Stellungnahme gefordert ist. Die Tochter schließt einen global organisierten, satzübergreifenden Einwand an (Z. 065-068), der thematisch passend ist. Allerdings wird die Verknüpfung zwischen dem Inhaltselement des ‚Unterhaltens‘ (Z. 064) im *Turn* der Mutter sowie dem ‚Warten‘ (Z. 065) im *Turn* der Tochter erst durch die Konklusion (Z. 068) aufgelöst. Aufgrund des Aufbaus des Äußerungspakets gelingt die inhaltliche Passung hier also nur partiell. In der Interaktion folgen dann weitere, aktiv initiiierende, d. h. ohne einen Zugzwang eingebrachte satzübergreifende Äußerungspakete der Tochter, die zwar grundsätzlich einen Bezug zum lokalen Kontext herstellen und somit sequenziell eingepasst sind, aber aufgrund ihrer Struktur ähnliche Passungsprobleme aufweisen wie Beispiel 3. Neben initiiierenden globalen Äußerungspaketen fasst die Tochter außerdem die für die Entscheidungsfindung verbleibenden Optionen zusammen und stellt diese zur Disposition; sie initiiert ferner ihrer Mutter gegenüber einen Zugzwang.

Mit 15 Jahren übernimmt die Tochter zum ersten Mal einleitend in die Interaktion eigenaktiv das Benennen einer Position sowie das autonome Einbringen von nicht-elaborierten Begründungen, wie Beispiel 4 illustriert.

Beispiel 4 – Kontextualisierung (15 Jahre)

068 TO: ich würd ma sagen V:IERte möglichkeit auf keinen [fall,]
 069 MU: [also]
 das hab ich mir schon [(geDA:/geNAU:) gedacht;]
 070 TO: [(lacht)]]
 071 MU: VIERte is sch auf [<pp>(keinen)>]
 072 TO: [(räuspert sich)]
 073 MU: ZWEIte auch nich,
 074 <<f>ÄH> drITte mein ich.
 075 TO: nee (.) DRITte au_nich.=
 076 MU: =DRITte auch nich.
 077 TO: aber ich überLEge zwischen: (.) eins und zwei.
 078 ich mein musicalbesuch wäre interesSANT.
 079 MU: JA.=
 080 TO: =(na)TÜRlich ne,
 081 MU: geNAU.
 082 TO: aber ROM (.) <p>mit kulturprogramm find ich eigentlich au_noch->
 083 MU: ja find ich [AUCH.]
 084 TO: [SEHR] interessant.
 085 MU: SEHR interessant.

Die jugendliche Sprecherin zeigt in diesem Ausschnitt, dass sie initiativ eingebrachte Positionen einleiten kann, z. B. metadiskursiv durch das *verbum dicendi* im Konjunktiv II (Z. 068), und neu eingebrachte Positionen gegen andere z. B. durch den Diskontinuitätsmarker *aber* (Z. 077; vgl. Quasthoff et al. 2019: 99) abgrenzen kann. Die Tochter bringt dabei satzförmige Begründungen ein, indem sie subjektive Bewertungen als Argumente für die Wahl einer Option nutzt (Z. 078: ich mein musicalbesuch wäre interesSANT.). Globale Zugzwänge der Mutter werden im weiteren Verlauf der Interaktion ebenfalls ausschließlich satzförmig bearbeitet, sodass angenommen wird, dass sie rezeptiv zwar wahrgenommen, aber produktiv nicht umgesetzt, d. h. vertextet werden können. Neben dem Initiieren einer eigenen Position zu Beginn der Interaktion sowie einer autonom eingeleiteten, nicht-elaborierten Begründung übernimmt die Tochter dann auch den Job des Abschließens und beendet die gemeinsame Entscheidungsfindung.

In der sprachlich-diskursiven Entwicklung der Sprecherin zeigt sich auf der Ebene der Kontextualisierung, dass sie mit 12 Jahren die globalen Anforderungen zwar rezeptiv, allerdings nicht produktiv bearbeiten kann. Auch mit 13 Jahren werden globale Anforderungen z. T. noch lokal von der Heranwachsenden bearbeitet. Allerdings werden in der Interaktion auch Sequenzen sichtbar, in denen die Sprecherin durchaus produktiv mit globalen Anforderungen umgeht. Aufgrund von Schwierigkeiten im Äußerungsaufbau weisen diese Beiträge aber inhaltliche Passungsprobleme auf. Im Alter von 15 Jahren bringt die Sprecherin dann eigenständig lokal operierende Beiträge, wie das Beziehen einer Position, ein. Doch auch globale Zugzwänge werden lokal bearbeitet. Im Gegensatz zu den vorangegangenen Erhebungen zeigt die Sprecherin sich mit 15 Jahren nicht nur verantwortlich für den Kernjob Begründen, sondern leitet auch in den Job des Abschließens ein.

10.2.2.2 Vertextung

Auf der Ebene der Vertextung baut die jugendliche Sprecherin mit 12 Jahren ihre Beiträge i. d. R. satzförmig auf und produziert sowohl Zustimmungen als auch Einwände gegenüber der Mutter, wie das nachfolgende Beispiel 1 zeigt.

Beispiel 1 – Vertextung (12 Jahre)

127 MU: also ich würd SCHON sagen-
 128 j:a:,
 129 WEIL (--) wir ganz selten urlaub machen können;
 130 (1.7)
 ((...))
 136 MU: und das is EINFach mal;
 137 da können wir zuSAMmen was machen;
 138 (-)
 139 TO: oKAY,=
 140 MU: =und (weil/wenn) wir (1.2) ALle diesen gutschein gewonnen haben,
 141 finde ich;
 142 da haben wir ALle was davon;
 143 TO: ((nickt))
 144 MU: und nich nur- (---)
 145 TO: MAARTen vielleicht nich so;=
 146 =aber trotzdem;
 147 MU: [MAARTen?]
 148 TO: [(lacht))]
 149 MU: ja;;
 150 der wird doch AUCH-
 151 der wird doch auch dr SPASS dran haben;
 152 (---)
 153 TO: <<pp>der HÄLT ja nich so viel aus;
 154 leider;>
 155 (---)
 156 MU: der HÄLT nich so viel (leid/leider) aus?
 157 TO: der-
 158 ja der HÄLT nich so viel [aus;]
 159 MU: [na::;]
 160 (0.9)
 161 ja;

Die Tochter stimmt der Begründung der Mutter (Z. 137: da können wir zuSAMmen was machen; / Z. 142 da haben ALle was davon;) für die Auswahl des Freizeitparks zu (Z. 139, 143) und schließt dann einen abgeschwächt modalisierten, satzförmigen Einwand an (Z. 145), den sie anschließend in eine schwache Zustimmung transformiert, indem sie ihn mit einem schnellen Anschluss einleitend mit einem adversativen Junktor verwirft (Z. 146: aber trotzdem). Durch die umgekehrte Reihenfolge einer *Cardinal Concessive* (vgl. Couper-Kuhlen und Thompson 2000: 394; vgl. Kap. 2.2.2) wird nicht der Einwand, sondern die schwache Zustimmung in den Vordergrund gerückt. Allerdings greift die Mutter anschließend den Einwand auf, stellt eine Rückfrage mit Bewertungsdivergenz (Z. 147) und widerspricht (149-151), was aufseiten der Tochter zu einer Elaboration ihres Einwandes führt (Z. 153f.: <<pp>der HÄLT ja nich so viel aus; / leider;>). In der satzförmigen Elaboration referenziert die Tochter durch das Demonstrativpronomen *der* auf den zuvor benannten Maarten und verweist in der angeführten Proposition, i. S. einer Behauptung (vgl. Goetz und Shatz 1999; Kyratzis et al. 2010), durch die Modalpartikel *ja* darauf, dass es sich bei dem Inhalt der Äußerung um ein evidentes bzw. geteiltes Wissen der Interagierenden handelt (vgl. Reineke 2016). Die Proposition wird anschließend, nach erneuter Rückfrage der Mutter (Z. 156), ohne die Modalpartikel wiederholt (Z. 158).

Begründungen werden seitens der Tochter also satzförmig und nur im Kontext von Elaborierungshinweisen (wie in Beispiel 1) oder, wie im Folgenden (vgl. Beispiel 2), nach expliziten globalen Zugzwängen der Mutter (Z. 164-165) eingebracht.

Beispiel 2 – Vertextung (12 Jahre)

164 MU: [und jetzt] hier gutschein DREI?
 165 waRUM nich?
 166 (1.3)
 167 TO: ähm;
 168 weil wir das nich BRAUchen,
 169 (---)
 170 MU: NEIN;
 171 weil EINFach;
 172 KUMma;=
 173 =wenn WIR das nehmen;
 174 also diese SPIElekonsole;

Für ihre Begründung (Z. 168) nutzt die Tochter ein *Brauchen*-Argument (vgl. Arendt 2019: 353), durch das die Sprecherin ihre Ablehnung „lediglich evidenztopisch begründet“ (Kreuz 2021: 188). Sie leitet die nicht weiter elaborierte Proposition nach einer Zögerungspartikel (Z. 167: ähm;) mit der kausalen Subjunktion *weil* ein und markiert dadurch explizit das Einbringen einer Begründung, die in der vorliegenden Sequenz satzförmig aufgebaut ist.

Auch weitere Beiträge der Tochter sind eher lokal strukturiert, sodass beispielsweise Oppositionen durch einfache Widersprüche realisiert werden, wie es im nachfolgenden Beispiel 3 deutlich wird.

Beispiel 3 – Vertextung (12 Jahre)

192 MU: ICH finde (0.9) das das is nich für die ganze familie etwas;
 193 und gutschein VIER,
 194 (3.1)
 195 MU: ein (-) bücher de vau DE (---) paket mit inhalt eurer WAHL;
 196 (1.9)
 197 MU: also das is auch so ähnlich wie mit dem: (.) mit gutschein DREI;
 198 (2.4)
 199 de vau des (-) HAM wir;
 200 (1.0) wir ham nich so: ne GROSse auswahl die wir-
 201 TO: von de vau DE:S?
 202 MU: ja;
 203 TO: DOCH;
 204 wir ham ganz [VIEle de vau des;]
 205 MU: [ja wir HAM die;]
 206 aber eben;
 207 uns geFÄLLT <<silbisch> nicht so viel;>
 208 sach ich ma jetzt;=
 209 =dass wir jetzt unheimlich viele de vau des haben !MÜS![sen;]
 210 TO: [ja;]
 211 MU: [also für FÜNfhundert] euro;
 212 TO: [wir HAM aber schon-]
 213 mama wir HAM schon ganz viele;
 214 also;
 215 MU: ja;
 216 [aber das sind speZIELle;]

Die Tochter bindet durch die Wiederaufnahme der Lexeme (Z. 199, Z. 201: *de vau des (-) HAM wir; / von de vau DE:S?*) ihre Rückfrage mit Bewertungsdivergenz an den *Turn* der Mutter zurück. Die Bewertungsdivergenz selbst wird durch die Dehnung (Z. 201: *DE:S?*) sowie die hochsteigende Tönhöhenbewegung markiert. Es folgt eine Markierung des Widerspruchs durch das fokusakzentuierte Negationsmorphem *doch* (Z. 203) sowie lexikalische und morpho-syntaktische Wiederaufnahmen, die die Sprecherin in der Fortsetzung des Widerspruchs nutzt (Z. 204): Sie setzt dazu der abgeschwächt modalisierten Mengenangabe in der Äußerung der Mutter (Z. 200: (1.0) *wir ham nich so: ne GROSse auswahl die wir-*) eine verstärkt modalisierte entgegen (Z. 204: *wir ham ganz VIEle de vau des;*). Die Beiträge der Sprecherin bleiben aber weiterhin lokal strukturiert – auch in der anschließenden Wiederholung (Z. 212f.) wird der Widerspruch nicht weiter elaboriert. Diese Form der Vertextung erweist sich als prototypisch für die gesamte Interaktion, in der die Sprecherin ihre Beiträge v. a. satzförmig strukturiert, aber u. a. durch die Verwendung von Adverbien und Wiederaufnahmen turnübergreifende Verbindungen zu den vorangegangenen Bezugsäußerungen herstellen kann.

Ein Jahr später, im Alter von *13 Jahren*, baut die Sprecherin dann deutlich vermehrt satzübergreifende bzw. global strukturierte Äußerungspakete auf. Dabei gelingt es der Sprecherin in der Vertextung der Äußerungspakete allerdings noch nicht, diese inhaltlich kohärent aufzubauen (vgl. Beispiel 4) bzw. ihre argumentative Funktion herauszustellen (vgl. Beispiel 5). In dem folgenden Beispiel 4 produziert die Sprecherin einen Einwand in Form eines größeren Äußerungspakets. Aufgrund inhaltlicher Strukturierungsschwierigkeiten verbleibt die Kerngedanke bzw. die Hauptaussage des Szenarios allerdings vorerst unklar und wird erst durch die Konklusion verständlich. Darüber hinaus zeigt die Sprecherin ebenfalls Schwierigkeiten darin, das Äußerungspaket abzuschließen.

Beispiel 4 – Vertextung (13 Jahre)

059 MU: also ich dEnke mal (.) also wenn man sich nicht so KENNT-
 060 (1.0)
 061 um: (.) sich besser KENnenzulernen;
 062 ist einfach äh (1.3) ja (2.3);
 063 °hh BOWling kann man au:ch;=
 064 =aber (0.8) DA: (-) unterhält man sich wahrscheinlich auch nicht
 so;=oder?
 065 TO: °h ja_doch bei bowling da muss man ja auch immer WARTen;
 066 also EIner °h mAcht ja sO,
 067 dann guckt man dann wieder zu==
 068 =und dann wenn man (.) WARTet dann redet man ja auch miteinander und so
 wahrscheinlich;
 069 ich WEISS es nicht so genau;
 070 (1.4)

Die Tochter leitet, im Anschluss an einen Zugzwang der Mutter, ihren *Turn* (Z. 065ff.) mit den hier negierenden Partikeln *ja doch* (Z. 065) ein und begründet ihren Einwand anschließend. Sie stellt einen Bezug zur verhandelten Option her, indem sie diese explizit benennt (Z. 065: *bei bowling*) und eine Behauptung einbringt, deren Inhalt sie durch die Modalpartikel *ja* als evident (vgl. Reineke 2016) markiert und durch das Frequenzadverb *immer* (vgl. Weinrich 2003: 581) als ein etabliertes, wiederkehrendes Prozedere (Z. 065: *da muss man ja auch immer WARTen;*) kennzeichnet. Die Sprecherin konkretisiert diese Behauptung im Folgenden (Z. 066-068), indem sie ein veranschaulichendes Szenario (vgl. Brünner und Gülich 2002b; Kinalzik 2024) einsetzt. Sie leitet die Konkretisierung durch den Diskursmarker *also* ein, beschreibt implizit zwei Handlungsabfolgen, die mit einem Temporaladverb verknüpft sind (Z. 066f.: *EIner °h mAcht ja sO, / dann guckt man dann wieder zu=*) und die anschließend in einer Konklusion münden (Z. 068: *und dann wenn man (.) WARTet dann redet man ja auch miteinander und so wahrscheinlich;*). In der Konklusion wird nun die ursprüngliche Proposition (Z. 065) abschließend

durch ein Konditionalgefüge elaboriert. Dabei schafft die Wiederaufnahme des fokusakzentuierten Lexems *warten* (Z. 065, Z. 068) eine Kohärenz zwischen den Turnkonstruktionseinheiten – erst mit dieser Konklusion wird der Einwand der Sprecherin auch inhaltlich verständlich. Das Abschließen des Einwandes erfolgt dann allerdings durch eine metakommunikative formelhafte Phrase (Z. 069), die ferner die Begründung nachträglich abschwächt (vgl. Helmer und Deppermann 2017). Die Verwendung dieser Phrase verweist darauf, dass noch Schwierigkeiten bestehen, längere Beiträge abzuschließen.

Die Sprecherin zeigt in der weiteren Interaktion außerdem auch Schwierigkeiten darin, die argumentative Funktion global strukturierter Äußerungspakete zu explizieren, wie das folgende Beispiel 5 illustriert.

Beispiel 5 – Vertextung (13 Jahre)

078 TO: °h von wegen BOWling-
 079 (als ich jetzt) also bei saMANthas geburtstag dA war ja auch bOwling;
 080 °hh da war ich ja AUCH-=
 081 =also ich mein ich war ja als einziges KIND da;
 082 MU: ja;=[und,]
 083 TO: [aber-]
 084 TO: ich mein: (.) mit der zeit ging_es dann BESser;
 085 <<pp>deswEgen;>
 086 und dann haben sich auch alle miteinander unterHALten und so;
 087 MU: also BOWling ist auch ganz gut;=o[der?]
 088 TO: [<<pp>ja;]=ich glaub bowling ist auch
 ganz gut,>=

Die Tochter bringt eigenaktiv initiiierend einen Erlebnisbericht ein (Z. 078-086), den sie explizit-sprachlich markierend zurückbindet an die bereits verhandelte Option des Bowling-Spielens (vgl. Beispiel 4). Die Funktion des Erlebnisberichts bleibt allerdings (vorerst) unklar, wie anhand der Rückfrage der Mutter (Z. 082) deutlich wird. In der Fortsetzung des Erlebnisberichts (Z. 083ff.) rekontextualisiert die Tochter den Erlebnisbericht dann zuerst durch eine nachträgliche Kausalitätsmarkierung (Z. 085: <<pp>deswEgen;>) als Begründung;⁹⁸ anschließend wird der Erlebnisbericht durch ein neu eingebrachtes Argument (Z. 086: und dann haben sich auch alle miteinander unterHALten und so;) für eine Belegerzählung (vgl. Heller 2012: 86) genutzt. Denn hier knüpft die Sprecherin an ein Kriterium an, das von beiden Beteiligten in der vorangegangenen Interaktion für die Entscheidungsfindung als relevant erachtet worden ist: Die Schüler*innengruppen müssen die Möglichkeit haben, sich miteinander zu unterhalten. Wie in der Reaktion der Mutter erkennbar wird (Z. 087), bleibt trotz der Bezugnahme auf dieses Kriterium unklar, ob die Sprecherin das Argument nun im Sinne einer Fürsprache für die Auswahl der Option anführt. Die Mutter bietet schließlich eine positive Bewertung der Option an (Z. 087), die von der Tochter durch eine Wiederholung ratifiziert wird (Z. 088). Das Beispiel illustriert, dass die jugendliche Sprecherin in diesem Alter noch Schwierigkeiten zeigt, die argumentative Funktion global strukturierter Äußerungen zu explizieren.

Zum dritten Erhebungszeitpunkt, als die Tochter 15 Jahre alt ist, bringt sie eine Reihe von lokal organisierten Positionen sowie nicht weiter elaborierten Begründungen ein. Im Gegensatz zu den vorangegangenen Erhebungszeitpunkten initiiert sie nun eigene Positionen und Begründungen ohne einen elternseitigen Zugzwang.

⁹⁸ In diesem Beispiel manifestieren sich Schwierigkeiten der Sprecherin sowohl hinsichtlich der Vertextung als auch der Kontextualisierung: Indem es der Sprecherin (vorerst) nicht gelingt, die Funktion des eingebrachten Erlebnisberichts zu explizieren, erweist sich das Äußerungspaket auch in seiner sequenziellen Passung als problematisch.

Beispiel 6 – Vertextung (15 Jahre)

068 TO: ich würd ma sagen V:IERte möglichkeit auf keinen [fall,]
 ((...))
 073 MU: ZWEIte auch nich,
 074 <<f>ÄH> dritte mein ich.
 075 TO: nee (.) DRITte au_nich.=
 076 MU: =DRITte auch nich.
 077 TO: aber ich überLEge zwischen: (.) eins und zwei.
 078 ich mein musicalbesuch wäre interesSANT.
 079 MU: JA.=
 080 TO: =(na)TÜRlich ne,
 081 MU: geNAU.
 082 TO: aber ROM (.) <<p>mit kulturprogramm find ich eigentlich au_noch->
 083 MU: ja find ich [AUCH.]
 084 TO: [SEHR] interessant.
 085 MU: SEHR interessant.

Die Tochter schließt gleich zu Beginn mit einer initiiierenden Position eine erste Option aus (Z. 068). Diese leitet sie mit dem einem *verbum dicendi* (Z. 068: ich würd ma sagen) metadiskursiv ein und verstärkt den Ausschluss der Option durch die Phrase *auf keinen Fall* modalisierend. Nach einer übereinstimmenden Position mit der Mutter (Z. 075) setzt sie dann mit der Markierung eines thematischen Bruchs durch das *aber* in einem metadiskursiven Kommentar (Z. 077) fort und benennt anschließend ihr Interesse an zwei anderen Optionen (Z. 078, Z. 082): Die erste Option (Z. 078) leitet die Sprecherin mit dem Diskursmarker *ich mein* (vgl. Günthner und Imo 2003) ein und markiert durch die konjunktivische Form *wäre* die Vorläufigkeit ihrer Bewertung. Es handelt sich um eine nicht-elaborierte Begründung, da hier die Evaluation des Musicalbesuchs als eine allgemeingültige Äußerung (Z. 078: wäre interessant) und somit als Argument für die Auswahl der Option genutzt wird. Erneut wird ein thematischer Bruch markiert (Z. 082): Die Sprecherin bringt nun, markiert durch *find ich* als Ankündigung einer Evaluation, eine subjektive Bewertung in Bezug auf die zweite Option ein, in der das Prädikativkomplement *interessant* in gesteigerter Form (Z. 084: SEHR interessant) wiederholt wird. Obwohl die Sprecherin hier nacheinander zwei Optionen bespricht, sind sie jeweils satzförmig strukturiert und sie werden darüber hinaus nicht zu einem größeren Äußerungspaket miteinander verbunden. Anders als zum zweiten Erhebungszeitpunkt (mit 13 Jahren) verwendet die Sprecherin im Alter von 15 Jahren keine Verfahren zur Veranschaulichung einer Begründung oder eines Einwandes, sondern ihre Beiträge bestehen vornehmlich aus satzförmigen Propositionen, die z. T. die Äußerung der Mutter expandieren. Anders als mit 13 Jahren gelingt der Sprecherin allerdings nun zum dritten Erhebungszeitpunkt das Abschließen der (lokal strukturierten) Begründungen, indem sie u. a. resümierende, konklusive Evaluationen einbringt, wie das nachfolgende Beispiel 7 zeigt.

Beispiel 7 – Vertextung (15 Jahre)

087 MU: [also das a]ndere [ist nur]
 088 TO: [m::]
 089 MU: [SHOP]pen oder [sowas;]
 090 TO: [<<creaky, kopfschüttelnd>nei::;>]
 091 TO: [°h ja;=]
 092 TO: =ich mein kann man ja auch eigentlich HIER mach[en.]
 093 MU: [((nickt))]
 093 MU: <<nickend>j[a.]>
 094 TO: [al]so
 095 (.)
 096 MU: ja.=
 097 TO: =das find ich jetzt eigentlich UNnötig.
 098 MU: ja.

Die Tochter zeigt bereits ihre Ablehnung gegenüber einer Option an (Z. 088, 090), während diese gerade von der Mutter verworfen wird (Z. 087, 089). Damit bereitet die Tochter zugleich ihre Turnübernahme (Z. 091f.) vor, die sie mit der Kombination von Partikel und Diskursmarker (Z. 091f.: ja;= / =ich mein) einleitet. Die Sprecherin bringt nun eine satzförmige, nicht weiter elaborierte Begründung ein (Z. 092), die sie anschließend mit einer evaluativen Konklusion (Z. 097: das find ich jetzt eigentlich UNnötig) abschließt. Während die Tochter zum zweiten Erhebungszeitpunkt noch Schwierigkeiten zeigt, die global(er) strukturierten Äußerungspakete abzuschließen, gelingt ihr dies mit 15 Jahren, indem sie an die Proposition eine Bewertung anschließt, die sie mithilfe eines referenzierenden *das* an ihre satzförmige Begründung (Z. 092) rückbindet. Erneut erfolgt eine ratifizierende Äußerung der Mutter (Z. 098). Insgesamt operieren die Beiträge der Tochter, wie Beispiele zum Beziehen einer Position, Bewerten und auch zum Begründen gezeigt haben, zu diesem Zeitpunkt weiterhin deutlich auf der Satzebene.

Über die drei Altersstufen zeigt sich in der sprachlich-diskursiven Entwicklung des Fokuskindes, dass die lokale Strukturierung der Beiträge, die die Sprecherin im Alter von 12 Jahren zeigt, mit 13 Jahren durch Versuche aufgebrochen wird, die eigenen Beiträge zu globalen Äußerungspaketen aufzubauen, in denen z. B. veranschaulichende Szenarios eingebracht werden oder Erlebnisberichte, die als Belegerzählungen fungieren. Dabei zeigt die Sprecherin u. a. noch Schwierigkeiten, die globalen Äußerungspakete lokal abzuschließen. Im Alter von 15 Jahren werden die Beiträge dann wiederum *ausschließlich* lokal strukturiert – globale Strukturierungen von Beiträgen sowie das Einbringen z. B. von veranschaulichenden Szenarios zur Stützung einer Position, wie sie im Alter von 13 Jahren zu beobachten sind, werden im Alter von 15 Jahren nicht vorgenommen. Allerdings nutzt die Sprecherin im Alter von 15 Jahren zum Abschließen ihrer satzförmigen Begründungen konklusive Evaluationen. In der Vertextung der Beiträge ist also im Alter von 12 zu 13 Jahren ein Ausbau der Globalität der Äußerungen zu erkennen, die allerdings im Alter von 15 Jahren reduziert wird.

10.2.2.3 Epistemische Modalisierung

Auf der Ebene der epistemischen Modalisierung zeigt sich, dass die Sprecherin im Alter von 12 Jahren sowohl abschwächende als auch evidenzmarkierende epistemische Modalisierungen einsetzt. Epistemische abschwächende Modalisierungen treten vornehmlich auf, wenn elternseitige Zugzwänge an die Sprecherin gerichtet werden (vgl. Beispiel 1). Evidenzmarkierende Modalisierungen treten hingegen zu Äußerungsinhalten hinzu, die die Sprecherin durch eine frequente Verwendung des Indikativs als Fakt markiert (vgl. Beispiel 2).

Zuerst wird nun in Beispiel 1 die epistemisch abschwächende Modalisierung u. a. in der formelhaften Phrase *ich weiß es nicht so genau* illustriert, die die Sprecherin in Folge eines elternseitigen Zugzwanges äußert.

Beispiel 1 – Epistemische Modalisierung (12 Jahre)

113 MU: (---) ein wochenende für die ganze faMilie (.) in einem freizeitpark,
 114 (1.2)
 115 was mei was HÄLTST du davon;=
 116 TO: =ich WEISS nich so genau;
 117 (1.5)
 118 <<flüsternd>(ich) bin mir nich SIcher;>
 119 MU: du bist dir nich SIcher?
 120 TO: nein;
 121 MU: waRUM nich?
 122 TO: ((zuckt mit den Schultern))
 123 (3.2)
 124 MU: hm?
 125 TO: ((zuckt mit den Schultern))
 126 (2.1)

Es zeigt sich innerhalb des Transkriptausschnitts, dass die Sprecherin *alle* elternseitig gesetzten Zugzwänge bzw. deren Aufrechterhalten (Z. 115, 121, 124) mit einem Schulterzucken oder formelhafte Phrasen wie *ich weiß es nicht so genau* bearbeitet (Z. 116, 118, 122, 125). Die Sprecherin markiert hier abschwächend modalisiert ein situatives Unvermögen, den Zugzwang bearbeiten zu können.

Dass die Sprecherin hinzutretend zu epistemisch abschwächenden Modalisierungen auch Evidenzmarkierungen einsetzt, zeigt nun Beispiel 2. In diesem schwächt die Sprecherin zuerst einen Einwand ab, markiert aber die anschließende satzförmige Begründung, die nicht weiter elaboriert wird, als fraglos evident.

Beispiel 2 – Epistemische Modalisierung (12 Jahre)

140 MU: =und (weil/wenn) wir (1.2) ALle diesen gutschein gewonnen haben,
 141 finde ich;
 142 da haben wir ALle was davon;
 143 TO: ((nickt))
 145 MU: und nich nur- (---)
 146 TO: MAARten vielleicht nich so;=
 147 =aber trotzdem;
 148 MU: [MAARten?]
 149 TO: [((lacht))]
 150 MU: ja;;
 151 der wird doch AUCH-
 152 der wird doch auch dr SPASS dran haben;
 153 (--)
 154 TO: <<pp>der HÄLT ja nich so viel aus;
 155 leider;>

Der initiierte Einwand der Tochter (Z. 146) wird durch das Satzadverb *vielleicht* modalisierend abgeschwächt. Der Grad der Sicherheit in Bezug auf die Aussage, dass der Bruder Maarten ‚nichts‘ von der Wahl des Freizeitparks habe, wird somit eingeschränkt. In der satzförmigen Begründung des Einwandes (Z. 154), die auf einen Widerspruch der Mutter (Z. 151f.) folgt, verstärkt die Tochter hingegen den Äußerungsinhalt durch die Modalpartikel *ja*. Die satzförmige Begründung des Einwandes, die keine weiteren erklärenden Ausführungen der Aussage enthält, wird als fraglos evident markiert. Insgesamt zeigt sich zu diesem Erhebungszeitpunkt, dass die Sprecherin sowohl epistemische verstärkende als auch abschwächende Modalisierungen in unterschiedlichen Formen (u. a. formelhaften Phrasen, Satzadverbien, Modalpartikeln) einsetzt.

Auch ein Jahr später, mit *13 Jahren*, reagiert die Sprecherin auf elternseitig gesetzte Zugzwänge durch formelhafte Phrasen wie *ich weiß es (gerade) nicht*, wie bereits in Kapitel 10.2.2.1 (vgl. Beispiel 2) gezeigt worden ist. Die Sprecherin nutzt darüber hinaus ähnliche Phrasen der epistemischen Unsicherheit (wie *ich weiß es (auch) nicht so genau*) in turnfinaler Position von Einwänden oder Begründungen – und schließt durch diese satzübergreifende, global strukturierte Äußerungen ab (vgl. Kap. 10.2.2.2, Beispiel 4). Weitere modalisierende Abschwächungen erfolgen vereinzelt in eingebrachten Positionen (z. B. einmalig durch das *verbum putandi glaub ich* und das Satzadverb *wahrscheinlich*). In Begründungen hingegen verwendet die Sprecherin besonders häufig die Modalpartikel *ja*, wie anhand des nachfolgenden Beispiels 3 illustriert wird, und zeigt dadurch benannte Inhalte als fraglos evidente Sachverhalte an.

Beispiel 3 – Epistemische Modalisierung (13 Jahre)

119 TO: (.) ich mein wenn man (.) Abendessen (.) erstmal (.) mit
 landestypischen gerichten, (-)
 120 dann würde man ja erst mal was über (.) ESsens(.)kultur erfahren-
 121 und (.) dann REdet man ja auch während eines essens;=
 122 =ich mein man ISST ja nicht nur; ((räuspert sich))

In der Begründung (Z. 119-122), in der die Sprecherin mit dem Konjunktiv II (Z. 120: würde) ein hypothetisches Szenario eröffnet, werden die jeweiligen Satzaussagen durch die Verwendung von insgesamt drei Evidenz markierenden Modalpartikeln (Z. 120, 121, 122: ja) epistemisch modalisierend verstärkt. Die Satzaussagen selbst verbleiben aufgrund der inhaltlichen Reihung der Propositionen allerdings vage (vgl. zu Schwierigkeiten der Vertextung Kap. 10.2.2.2, Beispiel 4 & 5). Diese Verwendung erweist sich als prototypisch für die Sprecherin, die in der gesamten Interaktion eine besonders hohe Anzahl der Modalpartikel *ja* zeigt und somit die Äußerungsinhalte ihrer Beiträge als fraglos evident markiert.

Mit 15 Jahren nutzt die Sprecherin sowohl epistemisch verstärkende sowie abschwächende Formen der Modalisierung. So werden beispielsweise ablehnende Positionen durch die Partikel *sowieso* modalisiert, wie das nachfolgende Beispiel 3 zeigt.

Beispiel 4 – Epistemische Modalisierung (15 Jahre)

104 TO: also ich mein:: (.) ich ich perSÖNlich mag sowieso nich so gerne auf_m
 schiff oder so.
 105 MU: ja.
 106 TO: glaub ich;

Die Sprecherin markiert durch die Verwendung der Partikel *sowieso* (Z. 104) ihr subjektives Befinden als ein nicht anfechtbares Faktum, sodass die Partikel das Argument verstärkend unterstützt (vgl. Thurmair 1989: 136; Eggs 2003: 272). Nach einer nur minimalen Affirmation der Mutter (Z. 105) schwächt die Sprecherin nun allerdings durch das nachgestellte *verbum putandi* (Z. 106) den Inhalt ihrer Begründung retrospektiv ab. In weiteren, hier nicht aufgenommenen Sequenzen nutzt die Sprecherin auch formelhafte Phrasen wie *auf keinen Fall* und die Modalpartikel *ja*, um eine Position zu verstärken bzw. Begründungsinhalte als evident zu markieren; sowie das Satzadverb *eigentlich* zur epistemischen Abschwächung.

Insgesamt zeigt sich in der sprachlich-diskursiven Entwicklung des Kindes auf der Ebene der epistemischen Modalisierung, dass es sowohl im Alter von 12 Jahren als auch mit 13 Jahren elternseitig gesetzten Zugzwängen mit einer formelhaften Phrase epistemischer Unsicherheit begegnet, die in der Form mit 15 Jahren nicht mehr auftritt. Während dann mit 12 und 15 Jahren die Formen epistemischer Abschwächung und Verstärkung ausgeglichen eingesetzt werden, nutzt die Sprecherin im Alter von 13 Jahren eine besonders hohe Anzahl von Evidenzmarkierungen in Form der Modalpartikel *ja*, um auch vage verbleibende Äußerungsinhalte als fraglos evident zu markieren.

10.2.2.4 Zusammenfassende Beschreibung der sprachlich-diskursiven Entwicklungslinie

In der längsschnittlichen Untersuchung des diskursstrukturell schwächeren Kindes mit dem Profil *Zustimmendes Position-Beziehen und Bewerten* zeigen sich innerhalb des Alterverlaufs von 12 bis 15 Jahren Veränderungen auf den Ebenen der Kontextualisierung, Vertextung und epistemischen Modalisierung, die nachfolgend zusammenfassend vorgestellt werden.

Kontextualisierung

- Im Alter von 12 Jahren zeigt die heranwachsende Sprecherin v. a. produktive Schwierigkeiten im Umgang mit den globalen Anforderungen des Gesprächs. So erkennt die Sprecherin globale Zugzwänge, kann sie aber nicht oder nur lokal bearbeiten; und auch die sequenzielle Platzierung gelingt bei selbstinitiierten Beiträgen – außer in opponierenden Handlungen – noch nicht sicher.
- Mit 13 Jahren kann die Sprecherin nun nicht nur rezeptiv, sondern vermehrt auch produktiv mit den lokalen und globalen Anforderungen der Interaktion umgehen. Die Sprecherin bearbeitet Zugzwänge und bringt selbstständig Beiträge ein. Allerdings zeigen sich im produktiven Umgang mit globalen elternseitig gesetzten Zugzwängen noch Schwierigkeiten, denn die Bearbeitung der Zugzwänge erfolgt lokal – erst mit Unterstützung des Elternteils erkennt die Tochter auch globale Anforderungen. Die Sprecherin bringt zu diesem Erhebungszeitpunkt, anders als noch mit 12 Jahren, aber auch global operierende Äußerungspakete ein, die sie in den lokalen Kontext sequenziell einpasst; allerdings führen Strukturierungsschwierigkeiten im Aufbau der Beiträge zu inhaltlichen Passungsproblemen. Insgesamt initiiert sie aber eigene Beiträge selbstständig, fasst eigenaktiv verbleibende Optionen der argumentativen Entscheidungsfindung zusammen und stellt diese zur Disposition und etabliert auch selbst einen Zugzwang gegenüber der Mutter. Die Sprecherin zeigt in ihrem Handeln, dass sie an den Anforderungen der argumentativen Entscheidungsfindung orientiert ist.
- Zum dritten Erhebungszeitpunkt, im Alter von 15 Jahren, bearbeitet die Sprecherin vereinzelt die kommunikativen Aufgaben eigenaktiv und bringt initiiierend Beiträge ein. Dabei operieren alle Beiträge – anders als im Alter von 13 Jahren – auf lokaler Ebene. Zu diesem Erhebungszeitpunkt zeigt sich die Tochter jedoch erstmalig verantwortlich für weitere kommunikative Aufgaben wie das Abschließen der Kernjobs.

Vertextung

- Die heranwachsende Sprecherin vertextet im Alter von 12 Jahren ihre Beiträge v. a. satzförmig. Durch Kausalitätsmarkierungen bzw. Negationen werden die Beiträge als Begründungen bzw. Widersprüche erkennbar. Weiterführende Elaborationen bleiben jedoch i. d. R. aus.
- Im Alter von 13 Jahren bringt die Sprecherin dann veranschaulichende Szenarios und Beleg Erzählungen ein, die die Sprecherin explizit-sprachlich an die vorangegangene Interaktion zurückbindet, implizit Handlungsabfolgen aneinanderreicht sowie Kausalitäten markiert. Doch sowohl der inhaltlich kohärente Aufbau der global strukturierten Äußerungspakete, das lokale Abschließen sowie Explizierung der argumentativen Funktion bereiten der Sprecherin noch Schwierigkeiten.
- Mit 15 Jahren reorientiert sich die Sprecherin auf lokale Strukturierungen der Äußerungen zurück. Sie veranschaulicht ihre eingebrachten Beiträge nicht mehr wie im Alter von 13 Jahren. Allerdings gelingt ihr das lokale Abschließen nun anhand konklusiver Evaluationen.

Epistemische Modalisierung

- Zum ersten Erhebungszeitpunkt, im Alter von 12 Jahren, nutzt die Sprecherin unterschiedliche Formen der epistemischen Modalisierung, die sowohl der Abschwächung als auch der Verstärkung von Äußerungsinhalten dienen. Auffällig sind dabei die formelhaften Phrasen, mit denen sie in Reaktion auf elternseitig gesetzte Zugzwänge Unwissenheit und somit ein situatives Unvermögen anzeigt, mit dem Zugzwang umgehen zu können.
- Im Alter von 13 Jahren treten die formelhaften Phrasen der epistemischen Abschwächung weiterhin im Anschluss an elternseitig gesetzte Zugzwänge, aber auch in turnfinaler Position global

strukturierter Beiträge auf. Neben vereinzelten weiteren Abschwächungen (z. B. *glaub ich, wahrscheinlich*) sowie dem Konjunktiv II, mit dem sie vereinzelt Szenarios als hypothetisch rahmt, verwendet die Sprecherin in besonders großem Umfang die Modalpartikel *ja*, um Äußerungsinhalte als fraglos evident zu markieren.

- Mit 15 Jahren verwendet die Sprecherin sowohl epistemisch abschwächende als auch verstärkende Formen der Modalisierung. Die mit 13 Jahren noch frequent eingesetzte Modalpartikel *ja* wird in diesem Umfang nicht mehr verwendet.

10.2.3 Interindividueller Vergleich

Anhand der qualitativen Untersuchung konnten die sprachlich-diskursiven Entwicklungen der Fokuskinder hinsichtlich der Fähigkeiten der Kontextualisierung, Vertexung sowie der Markierung der epistemischen Modalisierung intraindividuell über drei Erhebungszeitpunkte rekonstruiert werden. Die interindividuellen Unterschiede werden nun altersspezifisch und entlang der untersuchten Fähigkeitsdimensionen inklusive der epistemischen Modalisierung dargelegt.

Im Alter von 12 Jahren zum ersten Erhebungszeitpunkt zeigt das *erste Fokuskind* (Profil: *Explorative Position-Beziehen und Begründen*) einen rezeptiven wie produktiven Umgang mit den lokalen und globalen kommunikativen Anforderungen der argumentativen Entscheidungsfindung: Es bearbeitet globale Zugzwänge der Mutter, initiiert autonom eigene Positionen und produziert global strukturierte Äußerungspakete wie Begründungen und Einwände. Das Kind strukturiert seine Beiträge durchaus logisch-hierarchisch, indem es Bedingungen und Folgen expliziert, und veranschaulicht seine Begründungen durch den Aufbau hypothetischer Szenarios. Letztere markiert es durch die Verwendung des Konjunktivs II als mögliche zukünftige Ereignisse. Das Kind verstärkt in diesem Alter v. a. die Inhalte der Begründungen, indem es sie durch die Verwendung von Modalpartikeln als fraglos evident anzeigt. Hinsichtlich des *zweiten Fokuskindes* (Profil: *Zustimmendes Position-Beziehen und Bewerten*) lässt sich vergleichend feststellen, dass es gerade in Bezug auf elternseitig gesetzte Zugzwänge ein (situationales) Unvermögen zeigt, diese bearbeiten zu können. Insgesamt kann die Sprecherin diese zwar richtig kontextualisieren, sie aber nicht global bearbeiten: So werden Begründungen zwar sprachlich durch Kausalitätsmarkierungen (z. B. *weil*) eingeführt, dann aber satzförmig in Form eines *Brauchen*-Arguments vertextet. Das zweite Fokuskind ist vermehrt reagierend und weniger autonom initiiierend an der Bearbeitung der kommunikativen Aufgaben beteiligt. Werden Beiträge doch eigenaktiv eingebracht, bewegen diese sich, wie die benannten Begründungen, auf Satzebene. Insgesamt schwächt Fokuskind II im Alter von 12 Jahren seine Äußerungen vornehmlich epistemisch ab.

Zum zweiten Erhebungszeitpunkt, im Alter von 13 Jahren, bearbeitet nun *Fokuskind I* die elternseitig gesetzten Zugzwänge vermehrt satzförmig, baut aber selbstinitiierte Beiträge global strukturiert gemäß einem spezifischen Muster auf: Auf eine (ablehnende) Bewertung folgt eine Begründung sowie rahmenbildend z. T. eine weitere abschließende (ablehnende) Bewertung. Das Kind markiert ferner explizit kausale, konditionale und konzедierende Strukturen. Es baut auch Szenarios auf, die allerdings keine *möglichen* Ereignisse darlegen, sondern durch ihre konsequente Vertexung im Indikativ als *Fakt* markiert werden. Insgesamt verwendet das Fokuskind I wenig abschwächende oder verstärkende Modalisierungen. Mit 13 Jahren zeigt nun das *zweite Fokuskind* Veränderungen im Hinblick auf die produktiven Kontextualisierungsanforderungen: Es werden deutliche Bemühungen erkennbar, die eigenen Beiträge global zu strukturieren, die Beiträge zu veranschaulichen sowie Belegerzählungen anzuführen. Die Sprecherin markiert dazu zeitliche Abfolgen in Szenarios explizit und setzt konkludierend Konditionalgefüge ein, um eingebrachte Propositionen abschließend zu plausibilieren. Die Inhalte der Begründungen werden dabei besonders häufig durch die Verwendung der Modalpartikel *ja* epistemisch verstärkend

als fraglos evident markiert. Dabei gelingt der Sprecherin zwar grundsätzlich ein thematisch kohärenter Anschluss der globalen Äußerungspakete an den lokalen Kontext, allerdings führen Schwierigkeiten im Aufbau der global strukturierten Äußerung zu einer erschwerten inhaltlichen Passung; und auch das lokale Abschließen begründender Äußerungspakete erweist sich noch als schwierig, sodass die Sprecherin sie wiederkehrend mit der Markierung epistemischer Unsicherheit (*ich weiß es (auch) nicht so genau*) beendet.

Im Alter von 15 Jahren geht *Fokuskind I* rezeptiv und produktiv angemessen mit den Kontextualisierungsanforderungen um, baut seine Beiträge, außer das lokal operierenden Widersprechen, global strukturiert und logisch-hierarchisierend auf und passt sie in den lokalen Kontext sequenziell sowie thematisch ein. Dabei plausibilisiert es autonom bezogene Positionen durch explizit-sprachlich markierte Begründungen, verknüpft mehrere Argumente zu satzübergreifenden Begründungen und zeigt das Fortsetzen der Äußerung durch die syntaktische Projektion von Ergänzungssätzen an. Die Sprecherin markiert durch die vermehrte Verwendung von *verba sentiendi* wie *ich glaube* insgesamt einen explorierenden Modus der gemeinsamen Entscheidungsfindung; weitere Formen der modalisierenden Abschwächung und Verstärkung kommen hingegen selten vor. *Fokuskind II* zeigt ebenfalls, dass es lokale und globale Kontextualisierungsanforderungen erkennen kann und passt nun die eigenen Beiträge sequenziell ein. Jedoch operieren die Beiträge, in denen die Sprecherin v. a. Position bezieht und Bewertungen sowie nicht-elaborierte Begründungen einbringt, vermehrt lokal und sind satzförmig strukturiert. Dabei leitet die Sprecherin z. B. durch Diskursmarker wie *ich mein* in Begründungen ein. Epistemische Modalisierungen von Unsicherheit, die sie noch im Alter von 13 Jahren beispielsweise beim lokalen Abschließen längerer Äußerungspakete aufweist, zeigen sich zu diesem Zeitpunkt nicht.

Insgesamt konnten die qualitativen Untersuchungen der ausgewählten Fokuskinder das sich abzeichnende Bild zu den diskursstrukturellen Profilen (vgl. Kap. 10.1) bestätigen und darüber hinaus die *Entwicklung* der sprachlich-diskursiven Fähigkeiten auf den Ebenen der Kontextualisierung, Vertextung und der Markierung der epistemischen Modalisierung *differenziert* darlegen. Es zeigt sich, dass die Fokuskinder deutliche intra- sowie interindividuelle Unterschiede in der sprachlich-diskursiven Entwicklung ihrer Argumentationskompetenz aufweisen. Auf Grundlage der Untersuchung kann das erste Fokuskind als diskursstrukturell stärker und das zweite Fokuskind als diskursstrukturell schwächer eingeordnet werden. Da die vorliegende Studie den Erwerb diskursiver Fähigkeiten als *multimodalen* Diskurserwerb versteht (vgl. Kap. 4.2), wird im folgenden Kapitel 10.3 untersucht, wie sich der Einsatz der multimodalen Ressourcen über die drei Erhebungszeitpunkte verändert.

10.3 Entwicklung multimodaler Gestalten für das Begründen

In dem vorliegenden Kapitel 10.3 rücken nun die *multimodalen* Entwicklungslinien der Fokuskinder in den Mittelpunkt der Untersuchung. Das Ziel des Kapitels ist es, zu rekonstruieren, welche multimodalen Gestalten, die in Kapitel 8 induktiv rekonstruiert werden konnten, die Kinder über die drei Erhebungszeitpunkte hinweg einsetzen und ob bzw. wie diese sich vom in der Altersspanne von 12 bis 15 Jahren verändern. Analog zu den sprachlich-diskursiven Entwicklungen wird im Rahmen der *multimodalen* Entwicklungslinien in einem ersten Schritt untersucht, welche multimodalen Gestalten die Sprecherinnen zur Kontextualisierung, Vertextung und epistemischen Modalisierung einsetzen, um darauf aufbauend dann individuelle Veränderungen zu eruieren. Der Fokus der Fallstudien liegt auf der argumentativen Handlung des Begründens, wodurch ermöglicht wird, Veränderungen sowohl in der sequenziellen Einpassung sowie in der (globalen) Strukturierung der Äußerungen bzw. Äußerungspakete zu erfassen

(vgl. ausführlicher Kap. 7.2). Die Untersuchung berücksichtigt u. a., wie die multimodalen Ressourcen zeitlich sowie inter- und intrapersonell koordiniert werden.

Die Ergebnisse der Fallstudien werden dabei für jedes Fokuskind separat vorgestellt (vgl. Kap. 10.3.1 & Kap. 10.3.2). Um die Nachvollziehbarkeit der Befunde zu erleichtern, werden auf jeder Analyseebene einleitend die altersvergleichenden Befunde benannt sowie anhand einer tabellarischen Übersicht illustriert, bevor die Befunde anschließend entlang der *Funktion* der eingesetzten multimodalen Gestalten mit exemplarischen Belegstellen erläutert werden. Die in diesem Kapitel rekonstruierten *intraindividuellen* Unterschiede, die sich in der Untersuchung im Alter von 12 bis 15 für das jeweilige Fokuskind zeigen, werden dann in einem *interindividuellen* Vergleich beider Fokuskinder in Kapitel 10.3.3 zusammengeführt.

10.3.1 Fokuskind I – Begründen

Im Nachfolgenden wird untersucht, welche multimodalen Gestalten das Fokuskind I, das bereits zum ersten Erhebungszeitpunkt über ausgebaute sprachlich-diskursive Fähigkeiten verfügt (vgl. Kap. 10.2.1), für die Kontextualisierung und Vertextung sowie die epistemische Modalisierung der argumentativen Handlung des Begründens nutzt. Da die Sprecherin allerdings im Alter von 13 Jahren keine multimodalen Gestalten in der Realisierung der argumentativen Handlung des Begründens einsetzt, auf diesen aber der Fokus des vorliegenden Kapitels liegt, kann der zweite Erhebungszeitpunkt in der Untersuchung der *multimodalen* Entwicklung nicht berücksichtigt werden. Über die Gründe für das Ausbleiben der multimodalen Gestalten können aus einem konversationsanalytischen Verständnis heraus im Rahmen der vorliegenden Untersuchung keine Aussagen getroffen werden.

Um die Befunde hinsichtlich der insgesamt erbrachten argumentativen Handlungen des Begründens einordnen zu können, wird eine Übersicht (vgl. Tab. 35) über die Anzahl an Begründungen des Fokuskindes zu den jeweiligen Erhebungszeitpunkten den nachfolgenden Analysen vorangestellt.

Anzahl der Begründungen	Alter		
	12 Jahre	13 Jahre	15 Jahre
GESAMT	13	4	4

Tabelle 35: Anzahl der Begründungen (Fokuskind I)

10.3.1.1 Kontextualisierung

Wie in Kap. 7.2 bereits beschrieben, können multimodale Gestalten von Sprecher*innen auf der Ebene der Kontextualisierung genutzt werden, um beispielsweise in der sequenziellen Einpassung einer Begründung in den lokalen Kontext anzuzeigen, welche Sprechhandlung, wie z. B. das Präsentieren oder Verwerfen von Begründungen bzw. Begründungsinhalten, vorliegt (vgl. Kap. 8). Dadurch weist die Begründung einen Adressat*innenzuschnitt auf, denn die eingesetzten multimodalen Gestalten machen in ihrer performativen Funktion die jeweilige Sprechhandlung erkennbar.

In der vorliegenden Fallstudie nutzt die Sprecherin im Alter von 12 Jahren und 15 Jahren unterschiedliche multimodale Gestalten, die zur Kontextualisierung der Begründung beitragen: Sie präsentiert sowohl im Alter von *12 Jahren* als auch mit *15 Jahren* eine Begründung auf der geöffneten Handfläche (vgl. Kap. 8.2.1.1). Darüber hinaus kontextualisiert sie mit *12 Jahren* einen Begründungsinhalt als verworfen (vgl. Kap. 8.2.5) und hält mit *15 Jahren* eine opponierende Handlung in einer Konzession an (vgl. Kap. 8.4.3). Ein detaillierter Einblick in die Analysen erfolgt im Anschluss an die nachfolgende Tabelle 36, die aufgeschlüsselt nach dem Alter der Sprecherin einen Überblick darüber gibt, welche

multimodalen Gestalten die Sprecherin zur *Kontextualisierung* einer Begründung verwendet und in welcher Häufigkeit. Dabei ist die Benennung der Funktion den multimodalen Gestalten vorangestellt, da sich in der Untersuchung zeigt, dass in Kapitel 8 *separat* beschriebene multimodale Gestalten im Rahmen der Kontextualisierung und Vertextung *gleiche Funktionen* übernehmen können. So kann das Präsentieren von Begründung sowohl durch eine kurzzeitige, aber auch durch gehaltene *Palm Up* Gesten angezeigt werden (vgl. Kap. 8.2.1.1 & 8.2.1.2).

Kontextualisierung			
Alter	12 Jahre	13 Jahre	15 Jahre
multimodale Gestalten und ihre Funktionen zur Kontextualisierung	Präsentieren von Begründungen Präsentieren von (komplexeren) Begründungen (2)	–	Präsentieren von Begründungen kurzzeitiges Präsentieren und Hervorheben relevanter Lexeme der Begründung (2)
	Verwerfen in Begründungen Verwerfen in Begründungen (1)		Anhalten von Oppositionen retrospektives Anhalten einer vorangegangenen Opposition in Konzessionen (1)

Tabelle 36: Funktionen und Anzahl multimodaler Gestalten zur Kontextualisierung (Fokuskind I)

Präsentieren von Begründungen

Die Sprecherin kontextualisiert sowohl im Alter von *12 Jahren* (vgl. Beispiel 1) als auch im Alter von *15 Jahren* (vgl. Beispiel 2) einzelne Begründungen bzw. Begründungssegmente als präsentiert, wie die nachfolgenden Beispiele illustrieren. Dabei werden Unterschiede hinsichtlich des Einsatzes der multimodalen Gestalt sichtbar, denn die Sprecherin nutzt die in der multimodalen Gestalt enthaltene gestische Ressource im Alter von 12 Jahren, um zuerst eine sprachlich ambige Bezugnahme durch eine deiktische Geste auf das Material zu desambiguieren. Im Alter von 15 Jahren wird die gestische Ressource dann lediglich in Richtung des Materials geäußert – eine konkrete deiktische Referenz, z. B. zur Desambiguierung einer sprachlichen Referenz, wird allerdings nicht hergestellt.

Beispiel 1 – Präsentieren von Begründungen (12 Jahre)

221 MU: also müssten wir uns zwischen SPIELEkonsole und dem FREIzeitpark
entscheiden;
222 (2.6)
223 TO: ja okay;
((...))
227 TO: ich GLAUB wir würden den-
mu-gz >>auf Blatt----->>
228 aber den hat man sofort verBALLert;
229 den gutschein ZWEI;
230 (1.8)
231 MU: das STIMMT;
232 (--)



233 TO: |((lacht)) |ob|wohl DAVon hat man |lÄnger |wa |s;|
to-gz |.....|auf Blatt (>Z. 235)
to-rh |..|rH Palm Up auf Option-|hebt rH an|,,,,,,|
to-lh |öffnet lH Palm Up|,,|

234 (---)

235 MU: das is ne gute |iDEE;
to-gz -----|zu MU->>

In dem vorliegenden Transkriptausschnitt präsentiert die 12-jährige Sprecherin eine Turnexpansion (Z. 233) zu einem anfänglichen Ausschluss der Option des ‚Freizeitparks‘ (Z. 228f.). Mit der Expansion bzw. der Fortsetzung des *Turns* kommt es nun zu einer inhaltlichen Verschiebung des Bezugsselements: Während zuvor der „gutschein ZWEI“ (Z. 229) den Bezugspunkt ihrer Argumentation bildet, verhandelt sie in der Expansion (Z. 233) den Gutschein der ‚Spielekonsole‘. Beide Argumente gleichen sich darin, dass sie den Ausschluss des ‚Freizeitparks‘ anzeigen. Durch die einleitende Konjunktion *obwohl* wird die Expansion allerdings als eine Korrektur des ersten Arguments projiziert (vgl. Günthner 1999) – der Anschluss erweist sich in seiner sprachlich kohärenten Passung an dieser Stelle als schwierig. Zugleich wird das neue Begründungssegment mit einer *Palm Up* Geste (vgl. Müller 2004; Kendon 2004) koordiniert, mit der die Sprecherin auf dem Material auf die Option der ‚Spielekonsole‘ (vgl. #1) verweist: Dadurch wird einerseits das sprachlich nicht weiter elaborierte Adverb „DAvon“ (Z. 233) desambiguiert (vgl. Stukenbrock 2015) und zugleich die Option, auf die die Sprecherin zeigt, auf der geöffneten Handfläche präsentiert (vgl. Kendon 2004: 210). Die Sprecherin hebt anschließend, mit der Benennung des konkreten Arguments (Z. 233: lÄnger), die *Palm Up* Geste an und hält sie für einen Moment präsentierend in dem gemeinsamen Interaktions- und Wahrnehmungsraum der Interagierenden (vgl. #2), bevor sie sie anschließend in eine Ruheposition, mit kurzzeitigem Aufklappen der linken Hand (vgl. #3), überführt. Der Einsatz der multimodalen Ressourcen kontextualisiert das satzförmige Begründungssegment der Sprecherin als präsentiert. Es wird anschließend von der Mutter positiv evaluiert (Z. 235).

Mit 15 Jahren erfolgt das Präsentieren eines Begründungssegments nun innerhalb eines ausgebauten *Multi-Unit Turns*, in dem die Sprecherin eine Aufzählung mit insgesamt sechs Argumenten produziert und ein einzelnes Listenelement im gemeinsamen Wahrnehmungsraum präsentiert.

Beispiel 2 – Präsentieren von Begründungen (15 Jahre)

181 TO: man fährt (.) naja (.) so MITtelmäßig weg;
182 naja;
183 MU: tja;
184 [(du hattest/sonst) AUCH-]



185 TO: [(und) ich denk |man] |SIEHT |auch |(0.8) jeden|
 to-gs |.....|rH Palm Up Ri Blatt|,,,,|
 to-gz >>blickt auf Blatt----->|
 |tag etwas ANderes wahrscheinlich,
 to-gz |blickt zu MU----->>

In diesem Transkriptauszug präsentiert die Sprecherin innerhalb einer größeren Begründungssequenz ein einzelnes Listenelement auf der geöffneten Handfläche (Z. 185): Die Sprecherin koordiniert mit dem fokusakzentuierten Lexem „SIEHT“ das Öffnen der rechten Hand zu einer *Palm Up* Geste über dem Material, mit der die Sprecherin das Listenelement im gemeinsamen Wahrnehmungsraum der Interagierenden (vgl. #1) präsentiert. Durch die Ausrichtung der gestischen Ressource wird zugleich ein Bezug zu dem Material hergestellt, ohne dabei allerdings auf eine spezifische Option zu verweisen.

Die Sprecherin nutzt also zu beiden Erhebungszeitpunkten multimodale Gestalten des Präsentierens von Begründungen zur Kontextualisierung der argumentativen Handlung. Im Alter von 12 Jahren, in dem die Sprecherin die multimodale Gestalt im Rahmen einer Turnexpansion einsetzt, besteht die Geste aus einer Gesteneinheit, die zwei Gestenphrasen umfasst (vgl. Kendon 2004: 114; vgl. Kap. 3.1.1): Zuerst wird durch eine Referenz herstellende *Palm Up* Geste ein sprachlich ambig verbleibendes Adverb desambiguiert. Anschließend wird die *Palm Up* Geste angehoben und ein einzelnes, Nebenakzent tragendes Lexem hervorhebend präsentiert. Dadurch wird die Geste, die die Sprechhandlung kontextualisiert, zugleich eingesetzt, um eine Ambiguität auf der sprachlichen Ebene auszugleichen. Im Alter von 15 Jahren präsentiert die Sprecherin nun ein fokusakzentuiertes Lexem auf der geöffneten Handfläche. Eine Bezugnahme auf die Aufgabe wird hier durch die Ausrichtung der Hand erkennbar; sie fungiert an dieser Stelle allerdings nicht als Desambiguierung, sondern markiert, dass die Sprecherin weiterhin an der Bearbeitung der grundlegenden kommunikativen Aufgabe orientiert ist.

Verwerfen in Begründungen

Die Sprecherin kontextualisiert im Alter von 12 Jahren einen Gesprächsgegenstand innerhalb einer Begründung als verworfen (vgl. Kap. 8.2.5). Durch eine *Throwing Away* Geste (vgl. Bressemer und Müller 2014b) wird die Sprechhandlung des Verwerfens für alle Beteiligten sichtbar, wie das nachfolgende Beispiel 1 illustriert. Die Sprecherin schließt hier innerhalb einer satzförmigen Proposition durch die gestische Ressource eine Option aus der weiteren Entscheidungsfindung aus.

Beispiel 1 – Verwerfen in Begründungen (12 Jahre)

114 TO: hm: -=
 115 =eine spielekonsole mit zehn neuen spielen eurer WAHL-



116 |BASTi |hat |ne wii;=
 to-gz |blickt zu MU (>Z. 119)
 to-bo |lehnt sich Ri MU->
 to-gs |Throwing Away|
 mu-gz |zu TO----->
 117 =das |REICHT;|
 to-bo ----->|
 mu-gz ---->|auf Blatt->>
 118 ((lacht))
 119 MU: hm_|hm,
 to-gz -->|auf Blatt (>Z. 124)

Die Sprecherin verhandelt eine selbst eingebrachte Option (Z. 115: eine spielekonsole mit zehn neuen spielen eurer WAHL-) und nimmt anschließend Bezug darauf, dass ein Familienmitglied bereits über eine spezifische Spielekonsole verfügt (Z. 116: BASTi hat ne wii;). Mit Beginn der Äußerung wird dann ein ganzes Bündel an Ressourcen produziert: Die Tochter adressiert die Mutter durch ihren Blick, produziert eine *Throwing Away* Geste, bei der sie den Unterarm auf den Tisch aufstellt, die Handfläche senkrecht von sich abgewandt öffnet und anschließend abknickend fallenlässt, und lehnt ihren Oberkörper in Richtung der Mutter, die ihr den Blick zuwendet (vgl. #1-#4). Die Tochter markiert hier multimodal projizierend die benannte Option (Z. 115) als verworfen. Erst in der dann folgenden Turnkonstruktionseinheit (Z. 117) wird sprachlich expliziert, dass das Vorhandensein einer bereits existierenden Spielekonsole der Sprecherin genügt. Dabei reiht die Sprecherin innerhalb des Begründungssegments zwei Propositionen (Z. 116 & Z. 117) unverbunden aneinander. Eine Explikation der Begründung, wie etwa *Wir brauchen keine Spielekonsole, da Basti eine Wii hat*, oder aber weitere Elaborationen bleiben an dieser Stelle aus. Die Sprecherin zeigt nun in der Sequenz eine starke körperliche Ausrichtung auf die Mutter: Indem sie ihren Oberkörper zu ihr hinüberlehnt markiert sie eine „heightened orientation to the recipient“ (Goodwin 2018: 160), die erst mit Abschließen des *Turns* aufgelöst wird. Die gesteigerte Involvierung der Rezipientin in das satzförmig verbleibende Argument adressiert die Mutter somit zugleich als Verbündete. Eine Ablehnung oder ein Hinterfragen des Arguments erscheint an dieser Stelle unwahrscheinlich.

Die Heranwachsende nutzt im Alter von 12 Jahren also die multimodale Gestalt im Verbund mit einer besonders starken körperlichen Orientierung auf die Rezipientin funktional, um den Ausschluss einer Option innerhalb einer nicht weiter elaborierten Begründung anzuzeigen – und durchzusetzen.

Anhalten von Oppositionen

Mit 15 Jahren markiert die Sprecherin durch den Einsatz einer *Holding Away* Geste (vgl. Bressemer und Müller 2014a, 2014b) das Anhalten einer selbst eingebrachten Opposition innerhalb einer letztlich zustimmenden Konzession, wie das nachfolgende Beispiel 1 illustriert (vgl. Kap. 8.4.3). Anders als bei der Kontextualisierung des Verwerfens wird hier nicht angezeigt, dass ein Gesprächsgegenstand aus der weiteren Interaktion *ausgeschlossen* wird, sondern dass ein Einwand *angehalten* und nicht weiter fortgesetzt wird.

Beispiel 1 – Anhalten von Oppositionen (15 Jahre)

- 210 MU: ((schnalzt)) Okay. |
 mu-gz >>blickt auf Blatt->>
 to-gz ----->|blickt auf Blatt (>Z. 219)
- 211 (2.0)
- 212 MU: würd ich auch vorschlagen;
 213 |das_s [für]|ne GROSse klassengemeinschaft bestimmt |°h (1.7) am
 spAnnendsten sei[n könnte;]
- 214 TO: [ja:;]
 mu-he |nickt-----|
 to-he |nickt-----|
- 215 [SPRICHT] bestimmt nicht ALle an==
- 
- 216 =a|ber: |wahrsch |vielleICHT die mEisten;
 to-gs |öffnet lH Palm Down|,,,,,,|
 to-he |senkt Kopf n. li ab|
- 217 AUSSerdem (-) <<p>will man ja auch etwas zusAmmen machen;>=
 218 =<<pp>und-> (1.2)
 219 jetzt NICHTS mit shoppen (-) so GROSS|artig;
 to-gz ----->|blickt zu MU->>

Nachdem die Tochter der ins Abschließen steuernden Begründung der Mutter (Z. 211-214) affirmativ nickend und überlappend mit der gedehnten Partikel „ja:“ (Z. 215) zugestimmt hat, wendet sie in Zeile 215 ein (SPRICHT bestimmt nicht ALle an-). Einleitend mit dem Adversativjunktur *aber* im Verbund mit einer *Holding Away* Geste (vgl. #1; vgl. Bressemer und Müller 2014a) hält sie nun den Einwand retrospektiv an. In zeitlicher Abstimmung koordiniert die Sprecherin mit der gestischen Ressource das Absenken des Kopfes (vgl. #1), sodass sowohl das Anhalten des Einwandes als auch ein Perspektivwechsel (vgl. Maury-Ruan 2011) hin zur Zustimmung sichtbar angezeigt werden. In der Zustimmung der Konzession korrigiert die Sprecherin nun die Begründung der Mutter u. a. hinsichtlich der benannten Bezugsgröße (*Klassengemeinschaft* vs. *die meisten*) (vgl. ausführlich Kap. 8.4.3) und setzt die Begründung anschließend expandierend fort (Z. 217-219).

Im Alter von 15 Jahren setzt die Sprecherin in Konzessionen also fein koordiniert an der Scharnierstelle zwischen Einwand und Zustimmung (auf dem Adversativjunktur *aber*) die multimodale Gestalt des Anhaltens einer vorangegangenen Opposition in der Konzession ein. In enger zeitlicher Koordinierung mit weiteren Ressourcen wie dem Absenken des Kopfes markiert sie einen nun folgenden Perspektivwechsel. Sie leitet dadurch in eine zustimmende Begründung über.

10.3.1.2 Vertextung

Multimodale Ressourcen können, wie in Kapitel 8 rekonstruiert worden ist, funktional zur Strukturierung, d. h. zur Vertextung von Äußerungen bzw. größeren Äußerungspaketen beitragen. Auch die heranwachsende Sprecherin der vorliegenden Fallstudie setzt multimodale Gestalten zur Strukturierung ihrer Beiträge ein, wie Tabelle 37 zeigt. Die Sprecherin nutzt im Alter von *12 Jahren* multimodale Ressourcen, um relevante Lexeme kurzzeitig hervorzuheben (vgl. Kap. 8.2.1.2), Turnfortsetzungen einzuführen (vgl. Kap. 8.2.3) und multimodale Kohäsion und Kohärenz zu markieren (vgl. Kap. 8.2.4). Mit *15 Jahren* verwendet die Sprecherin dann multimodale Gestalten, um relevante Lexeme hervorzuheben sowie Begründungssegmente kohäsiv und kohärent miteinander – hier in einem *Multi-Unit Turn* – zu verbinden.

Vertextung			
Alter	12 Jahre	13 Jahre	15 Jahre
multimodale Gestalten und ihre Funktionen zur Vertextung	Hervorheben begründungsrelevanter Bedeutungselemente kurzzeitiges Präsentieren und Hervorheben relevanter Bedeutungselemente der Begründung (6)	–	Hervorheben begründungsrelevanter Bedeutungselemente kurzzeitiges Präsentieren und Hervorheben relevanter Bedeutungselemente der Begründung (3)
	Einführen von Turnfortsetzungen multimodales Einführen einer Turnfortsetzung (4)		
	Markieren von Kohäsion und Kohärenz multimodale Kohäsion und Kohärenz (2)		Markieren von Kohäsion und Kohärenz multimodale Kohäsion und Kohärenz (1)

Tabelle 37: Funktionen und Anzahl multimodaler Gestalten zur Vertextung (Fokuskind I)

Hervorheben begründungsrelevanter Bedeutungselemente

Da multimodale Gestalten unterschiedliche pragmatische Funktionen zugleich erfüllen können (vgl. Kap. 8), kann die im Folgenden besprochene multimodale Gestalt des kurzzeitigen Präsentierens relevanter Bedeutungselemente der Begründung nicht nur zur Kontextualisierung (vgl. Kap. 10.3.1.1), sondern auch zur Vertextung und somit zur Strukturierung von Äußerungen beitragen: Indem die Sprecher*innen einzelne Lexeme einer Äußerung auf der Handfläche präsentieren, heben sie diese in dem jeweiligen Äußerungskontext besonders hervor und setzen sie dadurch prominent. Sowohl mit *12 Jahren* (vgl. Beispiel 1) als auch mit *15 Jahren* (vgl. Beispiel 2) hebt die Sprecherin systematisch einzelne für die Begründung relevante Bedeutungselemente durch eine *Palm Up* Geste hervor. Dies ist insofern für die Vertextung von Begründungen ein hochgradig funktionales Vorgehen, da die Lexeme dadurch nicht nur prosodisch als relevant markiert, sondern auch körperlich-sichtbar prominent gesetzt werden, wie die nachfolgenden Beispiele 1 & 2 illustrieren.

Beispiel 1 – Hervorheben begründungsrelevanter Bedeutungselemente (12 Jahre)

```

109  TO:    =wir könnten ja dann |nämlich in_n |euROpa park fahren;=|
      to-gz -----|zu MU (>Z. 112)
      to-he                                     |nickt leicht----->
      to-bo                                     |Shoulder Shrug re---|

```


Auch mit 15 Jahren nutzt die Sprecherin die multimodale Gestalt, um sowohl durch gestische als auch prosodische Markierungen relevante Bedeutungselemente, hier nun innerhalb einer Aufzählung, präsentierend hervorzuheben.

Beispiel 2 – Hervorheben begründungsrelevanter Bedeutungselemente (15 Jahre)



#1

178 |man |macht |was |zu |SAMmen,|
to-gz |.....|blickt zu MU-----|
to-gs |.....|Palm Up, Beat|,,,,,,|
to-he |senkt Kopf ab----|,,,,,,|

((...))

181 TO: man fährt (.) naja (.) so MITtelmäßig weg;
182 naja;
183 MU: tja;
184 [(du hattest/sonst) AUCH-]



185 TO: [(und) ich denk |man] |SIEHT |auch |(0.8) jeden|
 to-gs |.....|rH Palm Up Ri Blatt|,,,,|
 to-he |bewegt Kopf n. li-----|
 to-gz >>blickt auf Blatt----->>|
 |tag etwas ANderes wahrscheinlich,
 to-gz |blickt zu MU----->>

Im Verlauf der in diesem Transkriptausschnitt nicht vollständig abgebildeten Aufzählung hebt die Sprecherin mehrere Bedeutungselemente durch die gestische Ressource der *Palm Up* Geste kurzzeitig hervor. In der abgebildeten Sequenz präsentiert die Sprecherin in zwei Argumenten jeweils die Lexeme „zuSAMmen“ (Z. 178, vgl. #1) und „SIEHT“ (Z. 184, vgl. #2) auf der geöffneten Handfläche. Diese werden nicht nur gestisch, sondern auch prosodisch durch die Verwendung des Fokusakzents prominent gesetzt und somit innerhalb der Äußerung besonders hervorgehoben. Im Rahmen der gesamten Aufzählung, die hier nur in Ausschnitten abgebildet ist, nutzt die Sprecherin die multimodale Gestalt also, um in aufeinander folgenden Listenelementen einzelne Bedeutungselemente hervorzuheben (vgl. unten, Markieren von Kohäsion bzw. Kohärenz).

Zu beiden Erhebungszeitpunkten setzt die Sprecherin die multimodale Gestalt des kurzzeitigen Präsentierens und Hervorhebens von relevanten Bedeutungselementen für die Begründung ein, um einzelne Lexeme der Begründung aus dem Äußerungskontext prominent hervorzuheben. Während sie die multimodale Gestalt im Alter von 12 Jahren besonders häufig (vgl. Tab. 37) und im Rahmen mehrerer Begründungen einsetzt, nutzt die Sprecherin die multimodale Gestalt im Alter von 15 Jahren ausschließlich innerhalb des *Multi-Unit Turns* der Aufzählung, um einzelne Bedeutungselemente der Listenelemente besonders hervorzuheben. Die Reihung der multimodalen Gestalten trägt auch zu einer internen Strukturierung der Aufzählung bei (vgl. unten, Markieren von Kohäsion bzw. Kohärenz). Indem die Strukturierung für alle Beteiligten sichtbar ist, wird die Verständlichkeit der Äußerung auch für das Gegenüber erhöht. In der Strukturierung drückt sich somit auch eine Adressat*innenorientierung aus.

Einführen von Turnfortsetzungen

In der Vertextung von Begründungen kann durch die Koordinierung von Strukturelementen wie beispielsweise Konnektoren (z. B. konzедierendes *aber*, konditionales *dann*) und *Palm Up* Gesten eine Turnfortsetzung eingeführt werden (vgl. Kap. 8.2.3). Diese multimodale Gestalt nutzt die Sprecherin in den vorliegenden Daten ausschließlich mit 12 Jahren – hier aber mehrfach (vgl. Tab. 37). Das nachfolgende Beispiel 1 demonstriert, wie die Sprecherin diese multimodale Gestalt einsetzt, um im Rahmen eines Konditionalgefüges eine Folge einzuführen.

Beispiel 1 – Einführen von Turnfortsetzungen (12 Jahre)

260 MU: ich (1.1) DENke ma;=
 261 =eigentlich ham wa (---) ham wa schon uns entSCHIEDen;=oder?
 262 TO: hm:: joa;;
 263 (2.1)

264 TO: aber man könnte zum beispiel auch die spiele verKAufen,
 265 |wenn man die gar nicht SPIElen |will,
 to-gz |zu MU-----|auf Blatt->
 mu-gz >>auf Blatt (>Z. 267)
 266 MU: hm_hm,



267 TO: |un::d äh:m: | (-) |DANN |
 to-he |senkt Kopf ab|
 to-gz |zu MU-----|auf Blatt----->
 to-gs |senkt bHde Palm Down auf Tisch ab|...|rH Palm Up---|
 mu-gz |zu TO----->
 |kann man von dem |gEld| (.) in den |frei |zeitpark fahren;|
 to-gs |rotiert-----|hält|.....|tippt m. rIF|
 to-gz ----->|zu MU----->>
 mu-gz -----|

268 (1.4)
 269 MU: DAT is aber nich gemeint;

In diesem Ausschnitt baut die Sprecherin ein kontrafaktisches Szenario (vgl. Brünner und Gülich 2002b) inklusive einem Bedingungsgefüge auf (Z. 264-265), das sie dann in Zeile 267, durch die gedehnte Konjunktion *und* syntaktisch angeschlossen, fortsetzt. Die darin enthaltene Folge wird durch mehrere Ressourcen eingeführt: Zu Beginn des *Turns* entstehen Häsitationen, in deren Pause die Sprecherin bereits beginnt, die rechte Hand aufzudrehen (Z. 267). Die Folge wird anschließend durch die Konjunktion „DANN“ (Z. 267) eingeleitet. Dabei senkt die Sprecherin den Kopf ab, während zeitgleich die Geste ihren Gestenhöhepunkt erreicht (vgl. #1). Die Fortsetzung des Szenarios wird also durch intrapersonell fein aufeinander abgestimmte multimodale Ressourcen eingeführt. Dadurch übernehmen die multimodalen Ressourcen gemeinsam in ihrer multimodalen Gestalt eine wichtige Strukturierungsfunktion und tragen zugleich zur Verständlichkeit der Äußerung bei: Denn durch ihren Einsatz werden neue, abgrenzbare Segmente der Begründung sichtbar.⁹⁹ Dabei setzt die Sprecherin diese multimodale Gestalt nicht nur wiederholt in weiteren Konditionalgefügen, sondern auch während des Konzedierens in Begründungen ein, um einen Einwand, eingeleitet mit dem Adversativjunktorkon *aber*, einzuführen.

Die multimodale Gestalt des Einführens von Turnfortsetzungen wird von der Sprecherin ausschließlich zum ersten Erhebungszeitpunkt, zu dem sie 12 Jahre alt ist, eingesetzt. Dabei führt sie in kataphorischer Funktion Strukturelemente ein, die Aufschluss darüber geben, ob es sich um eine konditionale, konzedierende und adversative Fortsetzung des *Turns* handelt.

Markieren von Kohäsion und Kohärenz

Zur Anzeige zusammenhängender Begründungssegmente, d. h. zur Markierung von Kohäsion und Kohärenz, nutzt die Sprecherin im Alter von 12 Jahren Gestenwiederaufnahmen, in denen durch die Ähnlichkeiten in der Form der Geste sowie ihrer Bewegungsrichtung Äußerungsbestandteile als zusammengehörig identifiziert werden können (vgl. Kap. 8.2.4). Die Sprecherin bindet hier sowohl zeitlich direkt

⁹⁹ Die Geste wird anschließend in eine rotierende Bewegung überführt (Z. 267: kann man von dem), wodurch „the arguments as a continuous sequence“ (Müller 2004: 250) angezeigt werden können (vgl. auch Ladewig 2014).

aufeinanderfolgende Begründungselemente kohäsiv aneinander zurück, markiert aber auch zeitlich entfernte Begründungselemente als kohärent zusammengehörig (vgl. Beispiel 1). Im Alter von *15 Jahren* (vgl. Beispiel 2) werden Wiederholungen der Äußerungsstruktur sowie multimodaler Ressourcen eingesetzt, um innerhalb einer Aufzählung die Listenelemente sowohl kohäsiv als auch kohärent als zusammengehörig zu markieren. Zu beiden Erhebungszeitpunkten markieren die multimodalen Gestalten sowohl Kohäsion als auch Kohärenz und erweisen sich in ihrer Struktur als insgesamt komplex. Vergleichend zeigt sich allerdings, dass sowohl die Komplexität der multimodalen Gestalten als auch ihre Reichweite, d. h. über welche zeitlichen Distanzen sie sich entfalten, im Alter von 15 Jahren ausgeprägter sind.¹⁰⁰ Zwei Belegstellen werden im Nachfolgenden exemplarisch einander gegenübergestellt.

Im vorliegenden Beispiel 1 markiert die Sprecherin im Alter von *12 Jahren* zwei satzförmige Begründungselemente als zusammengehörig, die durch einen Sprecher*innenwechsel in zeitlicher Distanz zueinander geäußert werden.

Beispiel 1 – Markieren von Kohäsion und Kohärenz (12 Jahre)

233 TO: ((lacht)) obwohl DAvon hat man länger was;
((...))
237 MU: von der SPIELEkonsole;
238 TO: mja;;
239 (1.2)
240 MU: aber trotzdem sind |wir ja Eher ne familie die gerne WEGfährt;
mu-gz | zu TO (>Z. 243)
to-gz | zu MU (>Z. 245)
241 und sich irgendwo was ANschaut;=oder?=
242 TO: =ja;=



243 | =aber | DANN | euro|pa park;|
to-he |nickt-----|
to-gs |rH Palm Up Ri. MU, ausgestr. IF|
mu-gz ----->|
244 (--)
245 MU: DANN könnten wir |in den europa park fahren;
to-gz ----->|auf Blatt (>Z. 249)
246 (1.4)
247 MU: DAS wÄr gu[t;]

¹⁰⁰ Der Begriff der Reichweite, der an dieser Stelle zur Beschreibung der Veränderungen der multimodalen Gestalten gewählt wird, lehnt an Quasthoff et al. (2019: 215). Die Autor*innen verwenden den Begriff, um die Entwicklung von Vertextungskompetenzen nachzuzeichnen: Mit zunehmenden Vertextungsfähigkeiten ist es Heranwachsenden möglich, Äußerungspakete nicht mehr ausschließlich lokal, sondern vermehrt global und dadurch mit einer größeren Reichweite zu strukturieren (vgl. auch Kap. 4.1.2).



248 TO: [und] DANN |sehen wir|schnEI |der|s;
 to-gs |.....|rH Palm Up Ri. MU, ausgestr. IF|,,,|
 249 TO: viel|[leicht;]
 to-gz ----|zu MU-- >>
 250 MU: [tja;]

In der vorliegenden Sequenz werden zwei Folgen sowohl syntaktisch als auch gestisch miteinander verbunden (Z. 243: aber DANN europa park; / Z. 248: und DANN sehen wir schnEIders;). Der Beginn der kindseitigen Begründung wird dabei mit dem Adversativitätsmarker *aber* eingeleitet. Allerdings deutet diese Markierung nicht auf eine Opposition der Sprecherin hin, vielmehr stimmt die Tochter dem Einwand der Mutter (Z. 240f.) nickend zu. Sie nennt dann eine Folge der Wahl des Freizeitparks (Z. 243) und markiert diese explizit-sprachlich durch die fokusakzentuierte Konjunktion „DANN“ (Z. 243). Die Tochter führt die Bedingung mit einer auf die Mutter ausgerichteten *Palm Up* Geste (vgl. #1), begleitet von dem fortgesetzten modalisierend verstärkenden Nicken sowie einem auf die Mutter ausgerichteten Blick, ein. Es erfolgt eine Zustimmung seitens der Mutter (Z. 245), in der sie Lexeme des *Turns* der Tochter recycelt (Z. 243, 245: DANN europa park) und die elliptische Äußerung der Tochter dabei syntaktisch vervollständigt. Die Tochter setzt syntaktisch durch die Konjunktion *und* (Z. 248) den *Turn* der Mutter fort. An dieser Stelle der Interaktion, d. h. in Zeile 248, passieren nun mehrere Dinge gleichzeitig: Die Tochter wiederholt das fokusakzentuierte Adverb „DANN“ aus ihrer eigenen Vorgängeräußerung (Z. 243); sie knüpft außerdem an die syntaktische Struktur des *Turns* der Mutter (Z. 245) an und baut ihren *Turn* in Zeile 248 parallel dazu auf; und schließlich nimmt sie die *Palm Up* Geste ihres eigenen *Turns* (Z. 243) wieder auf, die sie erneut in Richtung der Mutter ausrichtet und die nun aber auf dem nebenakzentuierten Lexem „schnEIders“ geäußert wird. Durch dieses Vorgehen schafft die Sprecherin eine Bezugnahme zwischen dem aktuellen *Turn*, ihrer eigenen Vorgängeräußerung (Z. 243) sowie zu dem *Turn* der Mutter (Z. 245): Durch die syntaktische sowie prosodische Parallelität bindet sie ihre Äußerung *kohäsiv* an die Äußerung der Mutter zurück (Z. 245 & Z. 248); die Gestenwiederaufnahme hingegen verweist auf eine *Kohärenz* zwischen den zeitlich entfernt liegenden Äußerungen der Tochter an (Z. 243 & Z. 248).

Im Alter von 15 Jahren, in dem die Sprecherin ihre Begründungen ausschließlich zu global strukturierten Äußerungspaketen aufbaut, werden nun im Rahmen einer Aufzählung (vgl. Selting 2004) die einzelnen Listenelemente sowohl kohäsiv als auch kohärent als zusammengehörig markiert, wie das nachfolgende Beispiel zeigt.

Beispiel 4 – Markieren von Kohäsion und Kohärenz (15 Jahre)

165 MU: also-
 166 (.)
 167 wir halten FEST,
 168 argument EINS (.) wäre?
 ((...))



177 TO: er is <<p>das:> | (.) | LÄNG|ste, |
to-fa | zieht rAB hoch|
to-gs | dreht rH auf-----| lässt rH Palm Down fallen|



178 |man |macht |was |zu |SAMmen, |
to-gz || blickt zu MU-----| blickt auf Blatt (>Z. 182)
to-gs || Palm Up, Beat| , , , , , |
to-he | senkt Kopf ab-----| , , , , , |

179 MU: ((zählt weiter mit der rH an der linken ab))
180 (1.3)



181 TO: man fährt (.) | naja (.) so MITtelmäßig |weg; |
to-gs | rotiert bHde Palm Away vor & zurück|
to-he | bewegt Kopf von re n. li|

182 naja;

183 tja;

184 [(du hattest/sonst) AUCH-]



185 TO: [(und) ich denk |man |] |SIEHT |auch | (0.8) jeden|
to-gs || rH Palm Up Ri Blatt| , , , , , |
to-he | bewegt Kopf n. li-----|
to-gz >>blickt auf Blatt----->>
|tag etwas ANderes wahrscheinlich,
to-gz |blickt zu MU----->>

((...))



188 TO: [|man kann |auch]|in die |NORD |see |
 to-he |Beat n. vorn|
 to-gz |blickt auf Blatt-----|blickt zu MU----->
 to-gs |.....|stellt bHde auf, HF Ri TO|bewegt rH n. vorn|
 übersetzen;

((...))



209 TO: man könnte |auch |Abends |(.) da noch (.)
 to-gs |hebt lH an, rotiert rH leicht auf|
 to-he |.....|senkt Kopf nach vorn ab----|
 e(.)twas am |segelschiff machen;
 to-gz |blickt zu MU----->
 210 MU: ((schnalzt)) oKAY.

Im Kontext der Aufforderung der Mutter, die Argumente für den Segeltörn zu benennen, bringt die Tochter ein größeres, global strukturiertes Äußerungspaket einer Begründung ein. Das Äußerungspaket ist als Aufzählung mit insgesamt sieben Listenelementen vertextet, wobei das erste Listenelement aus einer Wiederholung eines von der Mutter benannten Arguments besteht (Z. (177), 178, 171, 184, 185, 188, 209). Aufgrund ihrer jeweiligen multimodalen Gestalt und internen Struktur können v. a. die letzten sechs Listenelemente, i. S. v. Argumenten, sowohl kohäsiv als auch kohärent als zusammengehörig identifiziert werden: Die einzelnen Äußerungen sind syntaktisch parallel aufgebaut, da sie in der Regel mit der Wiederholung des Indefinitpronomens *man* beginnen, an die das finite Verb anschließt.¹⁰¹ Im Umfeld des finiten Verbs wird dann in jedem Listenelement eine Geste unterschiedlicher Form und Funktion geäußert, die von einer Bewegung des Kopfes (vgl. #2-#7; vgl. McClave 2000: 867ff.) begleitet wird. Die einzelnen Argumente werden dabei nicht durch eine sich wiederholende *spezifische* Hand- oder Kopfgeste als zusammengehörig markiert, sondern das Markieren der Struktur dieses Äußerungspakets mit den insgesamt sieben Listenelementen zeigt sich deutlich subtiler: In der syntaktischen und lexikalischen Parallelität erzeugen prosodische Markierungen wie Fokusakzentuierungen sowie Hand- und Kopfgesten eine Art gemeinsamen Rhythmus. Die Listenelemente werden durch diese Wiederho-

¹⁰¹ In Zeile 186 ist dieser syntaktischen Struktur die Diskursmarkierung *ich denk* vorangestellt, die innerhalb einer Überlappung mit der Rezipientin realisiert wird. Der auf die Diskursmarkierung anschließende Teil der Turnkonstruktionseinheit wird parallel zu den anderen Listenelementen aufgebaut.

lungen jeweils als zusammengehörig wahrnehmbar markiert. Diese Form der Kohärenzmarkierung ermöglicht, dass die Gesten, außerhalb ihrer Funktion der Äußerungsstruktur, noch weitere Funktionen, wie z. B. das Markieren von Unbestimmtheit (vgl. Z. 181; vgl. unten, Epistemische Modalisierung), übernehmen können.

Im Altersvergleich des Einsatzes der multimodalen Gestalten zeigt sich, dass die Sprecherin mit 12 Jahren einen *Turn* sowohl an eine eigene vorangegangene Äußerung sowie auch an die Äußerung der Mutter multimodal zurückbinden kann. Dazu nutzt sie gestische Wiederaufnahmen sowie syntaktische und prosodische Parallelitäten. Im Alter von 15 Jahren entfaltet sich die multimodale Gestalt im Rahmen einer größeren Aufzählung von Argumenten über einen *Multi-Unit Turn*. Auch hier markiert die Sprecherin multimodale Kohäsion und Kohärenz, jedoch werden die einzelnen Listenelemente subtiler aneinander zurückgebunden, indem der Ressourceneinsatz insgesamt eine Art Rhythmus erzeugt, der kohäsions- und kohärenzstiftende Funktionen übernimmt. Insgesamt zeigt sich im Alter von 15 Jahren ein deutlicher Ausbau der Komplexität, in der die multimodalen Ressourcen zu einer multimodalen Gestalt konfiguriert werden; und auch die Reichweite, d. h. über welche zeitliche Distanz sich die multimodale Gestalt innerhalb des global strukturierten Äußerungspakets entfalten, nimmt deutlich zu.

10.3.1.3 Epistemische Modalisierung

Wie in Kapitel 8.6 ausführlich dargelegt worden ist, können in argumentativen Entscheidungsfindungen epistemische Modalisierungen, wie beispielsweise Evidenzmarkierungen, dazu beitragen, den Inhalt einer Begründung als offensichtlich zu markieren und somit den Geltungsanspruch des Arguments zu erhöhen (vgl. Reineke 2016), oder durch die Anzeige von Unsicherheit oder auch Unbestimmtheit diesen Geltungsanspruch abzuschwächen. Die epistemische Modalisierung wird in den vorliegenden Fallstudien als Teil der Ebene der Markierung untersucht, da sie als Modus verstanden wird, auf dessen Boden die Interagierenden ihre argumentativen Beiträge aufbauen.

Wie die nachfolgende Übersicht (vgl. Tab. 38) illustriert, verwendet die heranwachsende Sprecherin im Alter von 12 Jahren multimodale Ressourcen, um geteiltes Wissen¹⁰² und auch epistemische Unsicherheit (vgl. Kap. 8.6.3) anzuzeigen. Besonders häufig verwendet die Sprecherin aber multimodale Gestalten zur Evidenzmarkierung (vgl. Kap. 8.6.1). Im Alter von 15 Jahren erweist sich insbesondere die hohe Anzahl von Evidenzmarkierungen als stark rückläufig, wie die Tabelle 38 zeigt, und auch epistemische Abschwächungen, i. S. des Markierens von Unbestimmtheit sind nur einmal belegt.

Epistemische Modalisierung			
Alter	12 Jahre	13 Jahre	15 Jahre
multimodale Gestalten und ihre Funktionen zur epistemischen Modalisierung	Anzeige geteilten Wissens - Anzeige von geteiltem Wissen (2) Evidenzmarkierung - Markieren von Evidenz - durch <i>Hand Shrug</i> (1) - durch Fallenlassen der Hand (2) - durch <i>Palm Up</i> Gesten mit lateraler Bewegung (1)	–	Evidenzmarkierung - Markieren von Evidenz - durch Fallenlassen der Hand (1)

¹⁰² Die multimodale Gestalt der *Anzeige von geteiltem Wissen* wird an dieser Stelle erstmalig aufgenommen, da sie in dem Datenkorpus der ersten empirischen Untersuchung auftritt.

	– durch kurzzeitiges Präsentieren einer Modalpartikel (1)¹⁰³ Anzeige epistemischer Unsicherheit • Anzeige von epistemischer Unsicherheit (1)		Markieren von Unbestimmtheit • Markieren von Unbestimmtheit (1)

Tabelle 38: Funktionen und Anzahl multimodaler Gestalten zur epistemischen Modalisierung (Fokuskind I)

Anzeige geteilten Wissens

In Interaktionen bauen Sprecher*innen ein gemeinsames Wissen auf, auf das sie sich während des Gesprächs zurückbeziehen können, wie die vorliegende Sequenz (vgl. Beispiel 1) der 12-jährigen Sprecherin zeigt. Die Sprecherin hat zu Beginn der Interaktion bereits die Auswahl des Freizeitparks begründend angeführt, die sie nun elliptisch wieder aufgreift und dabei in der multimodalen Gestalt anzeigt, dass sie auf bestehendes Wissen der Interagierenden referenziert.

Beispiel 1 – Anzeige geteilten Wissens

106 TO: ja;=ein wochenende in einem freizeits(.)PA:RK; ne?
 107 (---)
 108 is ja ganz SCHÖN;=
 109 =wir könnten ja dann nämlich in_n euROpa park fahren;=
 110 =denn da sehen wir SCHNEIders; ne? (.)
 ((...))
 240 MU: aber trotzdem sind |wir ja Eher ne familie die gerne WEGfährt;
 mu-gz | zu TO (>Z. 244)
 to-gz | zu MU (>Z. 245)
 241 und sich irgendwo was ANschaut;=oder?=
 242 TO: =ja;=



243 | =aber |DANN |euro|pa park;|
 to-he |nickt-----|
 to-gs |rH Palm Up Ri. MU, ausgestr. IF|
 mu-gz ----->|
 244 (---)
 245 MU: DANN könnten wir |in den europa park fahren;
 to-gz ----->|auf Blatt (>Z. 250)
 246 (1.4)
 247 MU: DAS wÄr gu[t;]
 248 TO: [und] DANN sehen wir schnEIders;

¹⁰³ Das kurzzeitige Präsentieren von Modalpartikeln ist in der Übersicht zur Vertexung (vgl. Kap. 10.3.1.2) nicht berücksichtigt worden, denn das Lexem übernimmt zwar eine relevante pragmatische Funktion für die jeweilige Begründung, allerdings trägt es nicht zu der inhaltlichen Entfaltung der Begründung bei.

In dem vorliegenden Transkriptauszug greift die Sprecherin eine Begründung von dem Gesprächsanfang wieder auf (Z. 109f., 243, 248) und verstärkt den Inhalt der wieder aufgenommenen Äußerung modalisierend durch ein Nicken (vgl. McClave 2000). Sie führt eine Folge der Wahl des Freizeitparks mit fokusakzentuiertem „DANN“ (Z. 243) ein (vgl. oben, Vertexung). Indem die Sprecherin die dabei geäußerte *Palm Up* Geste auf die Mutter ausrichtet (vgl. #1) und sie auch durch ihren Blick adressiert, markiert die Sprecherin, dass sie inhaltlich Bezug auf geteiltes bzw. bereits etabliertes Wissen nimmt, dass der Mutter bekannt ist. Die geäußerte Geste wird daher von Bavelas et al. (1995: 397) auch „*shared information gesture*“ genannt, denn durch sie können Interagierende anzeigen, „that the addressee is probably already aware of the information being delivered“ (Bavelas et al. 1995: 395). Da es sich um ein wieder aufgenommenes Argument handelt, scheint an dieser Stelle eine elliptische Ausführung zu genügen; auch die Mutter markiert keinen weiteren Elaborierungsbedarf, sondern evaluiert den Vorschlag positiv (Z. 247).

Die Sprecherin verwendet die multimodale Gestalt der Anzeige von geteiltem Wissen ausschließlich im Alter von 12 Jahren. Durch diese kann die Heranwachsende im Rahmen einer interpersonellen Bezugnahme auf bereits etabliertes und beiden bekanntes Wissen verweisen – ohne dies explizit-sprachlich markieren zu müssen.

Evidenzmarkierung

Multimodale Gestalten zur Markierung von Evidenz (vgl. Kap. 8.6.1) setzt die Sprecherin sowohl im Alter von 12 Jahren als auch im Alter von 15 Jahren multimodale Gestalten. Zum ersten Erhebungszeitpunkt wird eine besonders hohe Anzahl von Evidenzmarkierungen erkennbar (vgl. Tab. 38), aus denen nachfolgend exemplarisch ausgewählte vorgestellt werden (vgl. Beispiel 1, 2). Die Anzahl an Evidenzmarkierungen erweist sich dann im Alter von 15 Jahren als rückläufig – die Sprecherin verwendet nun nur noch ein einziges Mal eine multimodale Gestalt zum Markieren von Evidenz (vgl. Beispiel 3).

Im nachfolgenden Transkriptauszug (Beispiel 1) setzt die Sprecherin im Alter von 12 Jahren im Anschluss an eine selbst eingebrachte Proposition einen *Hand Shrug* (vgl. Kap. 8.6.1.3) zum *nachträglichen* Markieren von Evidenz ein. Es folgt anschließend eine Turnexpansion, in der die Sprecherin ihre Proposition nun durch eine Begründung stützt.

Beispiel 1 – Evidenzmarkierung (12 Jahre)

254 TO: |aber is |ja |am wOchenende; (.)
 to-gz ----->
 to-br |senkt AB-----|
 to-ns |kräuselt Nase--|
 to-gs |Palm Up|



255 MU: |ja. |
 to-gz ----->
 to-ey |kneift Augen zusammen--|
 to-he |nickt-----|
 to-gs |öffnet rH-----|

256 (--)
 257 TO: du musst ja auch [ZEIT] haben;
 258 MU: [in-]
 259 in ORdnung;

In der vorliegenden Sequenz bringt die Tochter eine Entkräftung (Z. 254) eines vorangegangenen Einwandes, der nicht im Transkript abgebildet ist, ein. Sie modalisiert die satzförmige Proposition verstärkend, indem sie durch die *Palm Up* Geste die Modalpartikel *ja* präsentierend hervorhebt (Z. 254). Es erfolgt nun eine minimale Zustimmung der Mutter (Z. 255), mit der die Tochter zeitgleich die Augen zusammenkneift, kurzzeitig den Daumen und Indexfinger zu einem *Hand Shrug* anhebt und fallenlässt und begleitend nickt (vgl. #1). Die Geste zeigt hier gewissermaßen nachträglich die eigene Proposition der Sprecherin als evident an und stuft rückwirkend die Relevanz der Äußerung hoch. Durch die nachträgliche multimodale Evidenzmarkierung wird aber auch die Reaktionsrelevanz des Gegenübers, d. h. der Mutter, die bislang der Proposition nur minimal durch die Partikel *ja* zugestimmt, erneuert (vgl. König 2017: 251). Die Tochter markiert hier die Rückmeldung der Mutter als noch nicht ausreichend und übernimmt im direkten Anschluss an die minimale Zustimmung und nachträgliche multimodale Evidenzmarkierung erneut den *Turn*. Darin führt sie eine weitere, die vorherige Äußerung stützende Voraussetzung für die Wahl dieser Option an (Z. 257). Diese stützende Begründung wird nun seitens der Mutter durch eine ausgebaute(re) Zustimmung affirmiert (Z. 259: in ORdnung;). Insgesamt greift die Tochter in der dargelegten Sequenz also auf die multimodale Gestalt der Evidenzmarkierung durch einen *Hand Shrug* zurück, um *nachträglich* Evidenz zu markieren, dadurch die Relevanz ihres eigenen *Turns* zu erhöhen und außerdem die Antwort- bzw. Reaktionsrelevanz ihres Gegenübers zu erneuern.

Die Sprecherin verleiht im Alter von 12 Jahren aber auch einzelnen Argumenten ihrer Begründung durch das geräuschvolle Fallenlassen der Hand, ebenfalls i. S. einer Evidenzmarkierung, besonderen Nachdruck (vgl. Kap. 8.6.1.1), wie das nachfolgende Beispiel 2 zeigt.

Beispiel 2 – Evidenzmarkierung (12 Jahre)

248 TO: [und] DANN sehen wir schnEiders;
 249 viel[leicht;]
 250 MU: [tja;]
 251 (2.7)

252 TO: aber |papa |muss |ja |auch |
 to-gs |rH Palm Up auf Blatt|hebt rH an|,,,,,,|
 to-gz -auf Blatt-----|zu MU->



|FREI |haben; |
 to-gs |öffnet rH Palm Up-|lässt rH geräuschvoll fallen|
 to-gz ----->

253 <<creaky>>oder so;>
 254 aber is ja am WOchenende; (.)
 255 MU: ja.
 256 (--)

In der hier abgebildeten Sequenz bringt die Tochter, nachdem sich die Interagierenden bereits auf die Option des Freizeitparks geeinigt hatten, noch eine Voraussetzung ein, die es bei der Wahl des Freizeitparks zu berücksichtigen gilt (Z. 252: aber papa muss ja auch FREI haben;). Eine Diskontinuität zur eigentlichen Einigung wird durch die Konjunktion *aber* markiert. Auf dem fokusakzentuierten Lexem „FREI“ (Z. 252) dreht die Sprecherin die rechte Hand zu einer *Palm Up* Geste auf und präsentiert das Lexem kurzzeitig, bevor sie dann die Hand geräuschvoll fallenlässt. Das akustische Signal, das in der Rückführung der Geste entsteht, verleiht der Äußerung, deren Inhalt bereits durch die Modalpartikel *ja* als evident markiert worden ist, turnfinal besonderen Nachdruck.

Im Alter von 15 Jahren markiert die Sprecherin – wie auch im Alter von 12 Jahren (vgl. Beispiel 2) – durch das Fallenlassen der Hand am Ende einer Intonationsphrase das benannte Argument im Rahmen einer Aufzählung als evident.

Beispiel 3 – Evidenzmarkierung (15 Jahre)

165 MU: also-
 166 (.)
 167 wir halten FEST,
 168 |argument EINS (.) wäre? |
 |Zählgeste mit rIF an lDaumen|
 ((...))
 171 MU: wenn man von [TAgen geht;]
 172 TO: [fü (.) achso] jetzt (.) ich dachte n (nun) ALLgemein,
 173 O de:r,
 174 MU: ist der|LÄNGS |te ausflug,
 mu-gs ----->|Beat mit Zählgeste|hält Zählgeste (>Z. 179)
 ((...))



177 er is <<p>das:> | (.) |LÄNG|ste, |
 to-fa |zieht rAB hoch|
 to-gs |dreht rH auf-----|lässt rH Palm Down fallen|

Nach der anfänglichen Schwierigkeit der Tochter, dem Zugzwang der Mutter nachzukommen (Z. 172f.), wiederholt sie anschließend zu Beginn einer einsetzenden längeren Begründungssequenz das Argument der Mutter (Z. 174, 177). Mit Einsetzen der Minimalpause (Z. 177) zieht die Tochter modalisierend verstärkend die rechte Augenbraue hoch und dreht die rechte Hand auf (vgl. #1), die sie anschließend zum Ende der Intonationsphrase geräuschvoll fallenlässt und dadurch das wiederholte Argument bzw. den Inhalt dessen als evident markiert (vgl. Kap. 8.6.1.1).

Im Altersvergleich zeigt sich, dass die Sprecherin mit 12 Jahren in hoher Frequenz und großer Varietät multimodalen Gestalten zum Markieren von Evidenz einsetzt. Dabei nutzt sie u. a. die multimodale Gestalt des Markierens von Evidenz durch einen *Hand Shrug rezipientinnenseitig*, um eine eingebrachte Proposition nachträglich in ihrer Relevanz hervorzuheben sowie die Antwortrelevanz des Gegenübers zu erhöhen. Diese Vielzahl und hohe Frequenz der multimodalen Gestalten zur Markierung von Evidenz erweist sich dann im Alter von 15 Jahren als rückläufig – hier verwendet die Sprecherin die multimodale

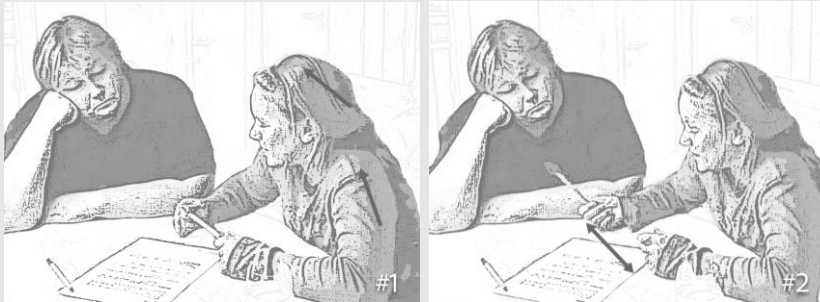
Gestalt nur einmal innerhalb der Interaktion, um ein wiederholtes Listenelement als offensichtlich zu markieren.

Anzeige epistemischer Unsicherheit

Epistemische Modalisierungen können auch für die Interaktion relevante Informationen über die epistemische Unsicherheit der Interagierenden (vgl. Heritage 2012b) beinhalten, „sowohl gegenüber Wissensobjekten, aber auch gegenüber ihren Gesprächspartnern“ (Reineke 2018: 195; vgl. Kap. 8.6.3). Im Alter von 12 Jahren setzt die Sprecherin die multimodale Gestalt der Anzeige von epistemischer Unsicherheit ein und modalisiert ihre Äußerung abschwächend, wie das nachfolgende Beispiel 1 illustriert.

Beispiel 1 – Anzeige epistemischer Unsicherheit (12 Jahre)

171 TO: das sind nur b(.)ücher un[d de vau des; ne?]
 172 MU: [Bücher und] de vau de|s;
 to-gz | zu MU (>Z. 177)
 173 TO: °h JA;=oKAY;
 174 MU: hm_[hm,]
 175 TO: [dann] GEHT das nich;=
 176 =als ()



177 | (.) <<pp>viellei|cht>|kann man |dann ja (.) |SPIELE |
 nehmen;
 to-he |senkt Kopf nach re ab|
 to-sh |hebt Schultern an-----|
 to-gz ----->|auf Blatt----->>>
 to-br |hebt AB an-----|
 to-gs |dreht bHde Palm Up auf| |öffnet bHde|
 178 MU: hm_hm,
 179 TO: ja;=aber;
 180 (---)
 181 JA.
 182 (1.7)
 183 MU: j[a.]

Die Sprecherin benennt in der vorliegenden Sequenz Umstände, unter denen die Wahl des Bücher- und DVD-Pakets in Frage käme (Z. 177: vielleicht kann man dann ja (.) SPIELE nehmen;). Bereits vor der Fortführung des *Turns* hebt die Sprecherin beide Schultern an, senkt den Kopf leicht nach rechts ab (vgl. Debras und Cienki 2012; vgl. #1) und beginnt dann den *Turn* mit dem epistemisch abschwächenden und in geminderter Lautstärke hervorgebrachten Satzadverb *vielleicht*. Die multimodale Gestalt wird nun durch weitere *Shrug*-Komponenten (vgl. Streeck 2009b) wie das Anheben der Augenbrauen, dem Abwenden des Blickes sowie dem seitlichen Aufdrehen beider Hände (vgl. #2) fortgesetzt und zeigt insgesamt die Haltung der Sprecherin als unsicher an (vgl. auch Debras 2017). Da das Hochziehen der Schultern sowie das Absenken des Kopfes bereits vor der Fortsetzung des *Turns* beginnen, können sie als erste, projizierende Hinweise der Unsicherheit interpretiert werden. Eine weitere Elaborierung des *Turns* bleibt an dieser Stelle aus.

Die multimodale Gestalt der Anzeige von epistemischer Unsicherheit wird von der Sprecherin einzig im Alter von 12 Jahren verwendet. Dabei werden gleich mehrere *Shrug*-Komponenten miteinander koordiniert, die gemeinsam in der multimodalen Gestalt zu einer epistemischen Abschwächung der nicht weiter elaborierten Begründung führen.

Markieren von Unbestimmtheit

Wie in Kapitel 8.6.4 dargelegt, können Interagierende durch die eingesetzten multimodalen Ressourcen auch inhaltliche Unbestimmtheit markieren. Die heranwachsende Sprecherin verwendet im Alter von 15 Jahren die multimodale Gestalt der Anzeige von Unbestimmtheit und markiert dadurch eine eingebrachte Entfernungsangabe als un- bzw. unterbestimmt, wie das nachfolgende Beispiel 1 zeigt.

Beispiel 1 – Markieren von Unbestimmtheit (15 Jahre)



181 TO: man fährt (.) |naja (.) so MITtelmäßig |weg; |
to-gs |rotiert bHde Palm Away vor & zurück|
to-he |bewegt Kopf von re n. li.|
mu-gs |zählt an MF ab, hält LH mit Zählgeste (>Z. 208)

182 naja;
183 MU: tja;

In diesem Listenelement (Z. 181) im Kontext einer äußerungsübergreifenden Begründung zeigt die Sprecherin multimodal Vagheit bzw. Unbestimmtheit in Bezug auf die tatsächliche Entfernung der Reise an (vgl. ausführlich Kap. 8.6.4). Dabei markiert die Heranwachsende sowohl durch die nicht akzentuierte Unschärfepartikel *so* (vgl. Weinrich 2003: 687; vgl. auch Ningelgen und Auer 2017: 5) sowie auch durch das Adjektiv „MITtelmäßig“ die Un- bzw. Unterbestimmtheit des Inhalts ihrer Äußerung. Noch vor der explizit-sprachlichen Markierung setzt bereits die Geste der „PDOH [Palm Down Open Hand, N. S.] with clockwise rotation“ (Bressem und Müller 2014a: 1580) ein, mit der Sprecher*innen „events, states, as well as ideas“ als „uncertain and indeterminate“ anzeigen können.

Die Sprecherin nutzt die multimodale Gestalt des Markierens von Unbestimmtheit einzig im Alter von 15 Jahren. Dabei zeigt sich in der Konfiguration der multimodalen Gestalt, dass die körperlich-sichtbaren Ressourcen fein koordiniert mit den sprachlichen Unbestimmtheits- bzw. Unschärfemarkierungen sind.

10.3.1.4 Zusammenfassende Beschreibung der multimodalen Entwicklungslinie

In der längsschnittlichen Untersuchung des diskursstrukturell stärkeren Fokuskindes I zeigen sich in der argumentativen Handlung des Begründens Veränderungen der funktional eingesetzten multimodalen Gestalten auf den analytischen Ebenen der Kontextualisierung, Vertextung und epistemischen Modalisierung. Da die Sprecher im Alter von 13 Jahren keine multimodalen Gestalten zur Realisierung der argumentativen Handlung des Begründens verwendet, konnte das Datum dieses Erhebungszeitpunktes in der Untersuchung nicht berücksichtigt werden. Insgesamt wird im Altersvergleich deutlich, dass die

Sprecherin mit 12 Jahren eine hohe Anzahl von Begründungen sowie eine Vielzahl multimodaler Gestalten einbringt (vgl. Tab. 35). Bis zum Alter von 15 Jahren sind nun sowohl die Begründungssequenzen sowie auch die dafür eingesetzten multimodalen Gestalten rückläufig. Allerdings zeigen sich nun Veränderungen u. a. hinsichtlich der Komplexität und Reichweite der eingesetzten multimodalen Gestalten. Die altersvergleichenden Beobachtungen werden im Folgenden entlang der drei Analyseebenen näher beleuchtet.

Kontextualisierung

- Die Sprecherin kontextualisiert sowohl im Alter von 12 Jahren als auch mit 15 Jahren Begründungen bzw. Begründungssegmente als präsentiert. Während die dabei eingesetzte *Palm Up* Geste im Alter von 12 Jahren nicht nur ein Begründungssegment präsentiert, sondern auch der Desambiguierung sprachlich ambiger Referenzen dient, nutzt die Sprecherin die gestische Ressource im Alter von 15 Jahren zum kurzzeitigen Präsentieren einer Begründung. Dabei wird durch die Ausrichtung der Geste in Richtung des Materials deutlich, dass das hervorgebrachte Argument in Verbindung zur Bearbeitung der kommunikativen Aufgabe steht. Anders als im Alter von 12 Jahren übernimmt die Geste keine Referenzherstellung zur Desambiguierung eines opaken Lexems.
- Mit 12 Jahren verwirft die Sprecherin in einer satzförmigen Proposition einen Gesprächsgegenstand durch eine *Throwing Away* Geste. In der multimodalen Gestalt wird dabei eine deutliche körperliche Orientierung der Sprecherin auf die Mutter erkennbar, durch die die Sprecherin ihre Mutter als Verbündete des Ausschlusses der Option adressiert. Eine Ablehnung des ausschließlich satzförmig als Proposition eingebrachten Arguments, das nicht weiter elaboriert wird, wird durch die Involvierung der Mutter unwahrscheinlich. Die Sprecherin nutzt also im Alter von 12 Jahren diese spezifische Konfiguration der multimodalen Gestalt, um eine Zustimmung der Rezipientin durch ihre verstärkte Involvierung einzuwerben.
- Mit 15 Jahren zeigt sie nun in konzедierenden, aber insgesamt zustimmenden Begründungen das Anhalten einer Opposition retrospektiv durch eine *Holding Away* Geste, die an der Scharnierstelle zwischen Einwand und Zustimmung platziert ist, an. In fein abgestimmter Koordinierung der multimodalen Ressourcen hält die Sprecherin die Opposition durch gestische Ressourcen an, während zugleich das Absenken des Kopfes einen Perspektivwechsel markiert und in die begründende Zustimmung eingeleitet wird.

Vertextung

- Mit 12 Jahren hebt die Sprecherin besonders viele Bedeutungselemente in satzförmigen sowie satzübergreifenden *Turns* kurzzeitig hervor. Mit 15 Jahren hingegen nutzt sie die multimodale Gestalt des kurzzeitigen Präsentierens und Hervorhebens relevanter Bedeutungselemente für die Begründung dann, um einzelne Elemente einer Aufzählung herauszustellen und somit auch die übergeordnete Aufzählung intern zu strukturieren.
- Im Alter von 12 Jahren werden Turnfortsetzungen multimodal eingeführt, die Konnektoren wie Folgenmarkierungen in Konditionalgefügen oder Adversativjunktoren beinhalten. Die *Palm Up* Gesten sind dabei i. d. R. zeitlich fein koordiniert mit den sprachlichen Strukturelementen. Die multimodale Gestalt wird zu den anderen Erhebungszeitpunkten von der Sprecherin nicht (mehr) verwendet.
- Sowohl mit 12 als auch mit 15 Jahren markiert die Sprecherin durch multimodale Kohäsions- und Kohärenzmarkierungen Begründungselemente als zusammengehörig: Mit 12 Jahren ver-

wendet die Sprecherin syntaktische Wiederaufnahmen, um ihre Äußerung kohäsiv an die Vorgängeräußerung der Mutter zurückzubinden, sowie gestische Wiederaufnahmen, um zeitlich entfernte satzförmige Begründungssegmente als kohärent zusammengehörig zu markieren. Mit 15 Jahren nimmt sowohl die Komplexität als auch die Reichweite der multimodalen Kohäsions- und Kohärenzmarkierungen zu: Die Sprecherin markiert innerhalb einer Aufzählung von Argumenten, die sie in einem *Multi-Unit Turn* strukturiert, sowohl die einzelnen Listenelemente als kohäsiv miteinander verbunden. Durch die fortlaufende Kohäsionsmarkierung werden aber in der insgesamt aus sieben Elementen bestehenden Liste auch das erste und das letzte Listenelement *kohärent* miteinander verbunden. Dazu baut die Sprecherin die Listenelemente syntaktisch parallel auf und koordiniert den Einsatz von Kopf- und Handgesten i. S. eines gleichbleibenden ‚Rhythmus‘, um die einzelnen Äußerungen kohärent miteinander zu einer Aufzählung zu verbinden. Vergleichend zeigt sich, dass die Sprecherin mit 15 Jahren die multimodale Gestalt des Markierens von Kohäsion und Kohärenz in einem höheren Grad der Komplexität sowie mit einer größeren Reichweite einsetzt als noch mit 12 Jahren.

Epistemische Modalisierung

- Im Alter von 12 Jahren nutzt die Sprecherin häufig multimodale Evidenzmarkierungen in unterschiedlichen multimodalen Gestalten. Durch den Einbezug ihres Gegenübers gelingt es ihr außerdem, eingebrachte Wissensbestände als geteiltes Wissen anzuzeigen und somit ihre Begründung abzusichern. Epistemische Unsicherheit markiert sie zu diesem Erhebungszeitpunkt hingegen nur einmal multimodal.
- Der Einsatz der multimodalen Gestalten zur epistemischen Modalisierung nimmt dann bis zum Alter von 15 Jahren ab. Sowohl das Markieren eines Begründungsinhalts als fraglos evident, das mit 12 Jahren noch frequent von der Sprecherin eingesetzt wird, sowie auch das Anzeigen epistemischer Unsicherheit erfolgt mit 15 Jahren nur noch selten bzw. gar nicht: Die Sprecherin markiert im Alter von 15 Jahren einmalig noch Evidenz sowie einen Begründungsinhalt als unbestimmt bzw. vage.

Eine Übersicht über die Veränderungen der eingesetzten multimodalen Gestalten sowie der Anzahl der Begründungen wird nachfolgend auch tabellarisch (vgl. Tab. 39) skizziert.

Alter	12 Jahre	13 Jahre	15 Jahre
Kontextualisierung			
Funktionen der multimodalen Gestalten	Präsentieren von Begründungen - Präsentieren von (komplexeren) Begründungen (2) Verwerfen in Begründungen - Verwerfen in Begründungen (1)	–	Präsentieren von Begründungen - kurzzeitiges Präsentieren und Hervorheben relevanter Lexeme der Begründung (2) Anhalten von Oppositionen - retrospektives Anhalten einer vorangegangenen Opposition in Konzessionen (1)
Vertextung			
Funktionen der multimodalen Gestalten	Hervorheben begründungsrelevanter Bedeutungselemente - kurzzeitiges Präsentieren und Hervorheben relevanter Bedeutungselemente der Begründung (6) Einführen von Turnfortsetzungen - multimodales Einführen einer Turnfortsetzung (4) Markieren von Kohäsion und Kohärenz - multimodale Kohäsion und Kohärenz (2)	–	Hervorheben begründungsrelevanter Bedeutungselemente - kurzzeitiges Präsentieren und Hervorheben relevanter Bedeutungselemente der Begründung (3) Markieren von Kohäsion und Kohärenz - multimodale Kohäsion und Kohärenz (1)
Epistemische Modalisierung			
Funktionen der multimodalen Gestalten	Anzeige geteilten Wissens - Anzeige von geteiltem Wissen (2) Evidenzmarkierung - Markieren von Evidenz - durch <i>Hand Shrug</i> (1) - durch Fallenlassen der Hand (2) - durch <i>Palm Up</i> Gesten mit lateraler Bewegung (1) - durch kurzzeitiges Präsentieren einer Modalpartikel (1) Anzeige epistemischer Unsicherheit - Anzeige von epistemischer Unsicherheit (1)	–	Evidenzmarkierung - Markieren von Evidenz - durch Fallenlassen der Hand (1) Markieren von Unbestimmtheit - Markieren von Unbestimmtheit (1)
Anzahl der Begründungen			
GESAMT	13	4	4

Tabelle 39: Übersicht über Funktionen der eingesetzten multimodalen Gestalten für das Begründen (Fokuskind I)

10.3.2 Fokuskind II – Begründen

Im Folgenden werden die Untersuchungsergebnisse der Fallstudie des zweiten Fokuskindes vorgestellt, dessen sprachlich-diskursive Fähigkeiten weniger stark ausgeprägt sind als die des ersten Fokuskindes (vgl. Kapitel 10.2.2). Die Ergebnisse der Fallstudie werden ebenfalls entlang der Fähigkeitsdimensionen der Kontextualisierung und Vertextung sowie, quer liegend dazu, anhand der epistemischen Modalisierungen vorgestellt. In der Fallstudie des zweiten Fokuskindes verwendet die Sprecherin im Alter von 12 Jahren keine multimodalen Gestalten zur Realisierung der argumentativen Handlung des Begründens. Da der Fokus der vorliegenden Untersuchung aber auf der *multimodalen* Entwicklungslinie liegt, kann dieser Erhebungszeitpunkt in den folgenden Analysen nicht berücksichtigt werden.

Um nun die nachfolgenden Befunde des Einsatzes der multimodalen Gestalten hinsichtlich der insgesamt erbrachten argumentativen Handlungen des Begründens einordnen zu können, wird eine Übersicht (vgl. Tab. 40) über die Begründungen des Fokuskindes zu den jeweiligen Erhebungszeitpunkten vorangestellt.

Anzahl der Begründungen	Alter		
	12 Jahre	13 Jahre	15 Jahre
GESAMT	n = 2	n = 4	n = 3

Tabelle 40: Anzahl der Begründungen (Fokuskind II)

10.3.2.1 Kontextualisierung

Das zweite Fokuskind verwendet zur Kontextualisierung, d. h. zur interpretativen Rahmung der Sprechhandlung (vgl. dazu auch 10.3.1.1), im Alter von 13 Jahren sowie im Alter von 15 Jahren multimodale Gestalten zum Präsentieren einer Begründung. Im Alter von 15 Jahren kontextualisiert die heranwachsende Sprecherin außerdem einen Gesprächsgegenstand als verworfen. Die nachfolgende Tabelle 41 gibt einen Überblick über die Funktionen sowie die Häufigkeit der von der Sprecherin für die Kontextualisierung eingesetzten multimodalen Gestalten.

Kontextualisierung			
Alter	12 Jahre	13 Jahre	15 Jahre
multimodale Gestalten und ihre Funktionen zur Kontextualisierung	–	Präsentieren von Begründungen - Präsentieren von (komplexeren) Begründungen (1) - kurzzeitiges Präsentieren und Hervorheben relevanter Lexeme der Begründung (2)	Präsentieren von Begründungen - Präsentieren von (komplexeren) Begründungen (1) Verwerfen in Begründungen - ablehnendes Fallenlassen eines Gesprächsgegenstandes (1)

Tabelle 41: Funktionen und Anzahl multimodaler Gestalten zur Kontextualisierung (Fokuskind II)

Präsentieren von Begründungen

Die Sprecherin kontextualisiert sowohl im Alter von 13 Jahren (vgl. Beispiel 1) als auch im Alter von 15 Jahren (vgl. Beispiel 2) eine Begründung als präsentiert (vgl. Kap. 8.2.1). Allerdings unterscheiden sich die präsentierten Begründungen hinsichtlich ihres sequenziellen Kontextes: Während die Sprecherin im Alter von 13 Jahren eine intonationsphrasenübergreifende Begründungspräsentation unternimmt, nutzt sie die multimodale Gestalt des Präsentierens von Begründungen im Alter von 15 Jahren im Rahmen einer satzförmigen Proposition.

In der nachfolgenden Sequenz baut die Sprecherin eine Begründung, bestehend aus mehreren Intonationsphrasen, auf, die sie auf der geöffneten Hand einer *Palm Up* Geste präsentiert. Allerdings bricht sie anschließend den Turnaufbau ab.

Beispiel 1 – Präsentieren von Begründungen (13 Jahre)

092 MU: also mu|SE |um (0.5) |glaub ich das: (0.8) 2hm| (0.7)
 mu-he |Beat| |Kopfschütteln-----|
 093 TO: |°h (nee/ja) ich MEI[ne;]
 mu-gz |blickt zu TO (>Z. 098)
 094 MU: [ist NICHT] jeder[manns sache;]
 095 TO [die EInen wü] °hh
 096 <<lachend>ja_a>;
 097 MU: WAS?=
 #1



098 TO: =die EInen würden dann ja was ü|ber |IHre |kultur |erfahren;
 to-gz -blickt auf Blatt----->
 to-gs |rH Palm Up auf Blatt|tippt--|.....->
 mu-gz ----->|auf Blatt----->



099 |und die ANder|en |über i |
 to-gz ----->|
 to-gs->|senkt Palm Up seidl. ab-|Palm Up n. außen|
 mu-gz ----->|
 | (0.5) |<<creaky, pp> (unsere)>|
 to-gz |blickt i. d. Raum-----|
 to-gs |Cyclic Gesture-|bHde Palm Up, n. innen--|
 mu-gz |zu TO----->>
 100 ich WEISS es AUch nicht so genau;

In der Begründung (Z. 095, 098ff.) beginnt die 13-jährige Sprecherin selbstständig ein hypothetisches Szenario aufzubauen: In diesem nutzt sie eine *Palm Up* Geste sowohl für die Gegenüberstellung von zwei eingeführten Personengruppen (Z. 098, 099: die EInen, die ANdern) als auch zur Präsentation der Begründung. Mit Bezugnahme auf die erste Schüler*innengruppe (Z. 098: IHre) führt sie die rechte Hand mit der nach oben geöffneten Handfläche als *Palm Up* Geste auf das Aufgabenblatt (vgl. #1).¹⁰⁴ Sie setzt nun die Begründung weiter fort (Z. 099) und führt die *Palm Up* Geste in der Gegenüberstellung (Z. 099: die ANderen) weiter an den Rand des gemeinsamen Wahrnehmungsraumes (vgl. #2, #3). Dabei

¹⁰⁴ In der Aufgabenstellung ist kein inhaltlicher Bezug auf das von der Tochter benannte Kennenlernen der Kulturen enthalten. Es wird lediglich benannt, dass es sich um den Besuch eines Museums zur deutsch-polnischen Geschichte handelt.

hält die Sprecherin die *Palm Up* Geste intonationsphrasenübergreifend konsequent aufrecht, sodass das sich im Aufbau befindliche und anschließend abgebrochene Szenario (Z. 099) als präsentiert kontextualisiert wird. Zeitlich wird das Einsetzen der Geste jedoch nicht mit dem Turnbeginn oder dem einführenden Benennen der ersten Schüler*innengruppe (Z. 098: die ElInen), sondern mit dem Personalpronomen „IHre“ (Z. 098) koordiniert. Die gestische Ressource, durch die das Szenario auf der geöffneten Handfläche zur gemeinsamen Ansicht angeboten wird, setzt also verzögert ein und kontextualisiert erst in der *Mitte* der Intonationsphrase für das Gegenüber die Begründung als präsentiert.

Im Alter von 15 Jahren sind die eingebrachten Begründungen lokal strukturiert. Die Sprecherin präsentiert im Nachfolgenden Beispiel 2 eine satzförmige Begründung, die aus einer nicht weiter elaborierten Proposition besteht.

Beispiel 2 – Präsentieren von Begründungen (15 Jahre)

077 TO: aber ich überLEge zwischen: (.) eins und zwei.



078 ich mein musi|cal|besuch wäre |intere|sSANT.|
 to-gs |...|dreht rH Palm Up auf--|,,,,,,|
 to-he |neigt Kopf nach re-|,,,,,,|

079 MU: |JA.= |
 mu-he |nickt|
 mu-gz |blickt zu TO (>Z. 085)

080 TO: =(na)TÜRlich ne,

081 MU: |geNAU.|
 mu-he |nickt-|

Die Sprecherin leitet die Begründung mit dem Diskursmarker *ich mein* (vgl. Günthner und Imo 2003) ein und dreht mit Benennen der Aktivität ‚Musicalbesuch‘ die rechte Hand zu einer *Palm Up* Geste auf, die sie in einem kurzzeitigen *Post-stroke Hold* und anschließend mit evaluierendem „interessANT“ (Z. 078) zurück in eine Ruheposition führt. Zeitgleich mit dem Öffnen der Hand neigt die Sprecherin ihren Kopf in einem langsamen Nicken nach rechts (vgl. #1, #2) und unterstützt somit verstärkend ihre Proposition. In dieser Sequenz koordiniert die Sprecherin also mit der Benennung des ersten inhaltlichen Elements der Proposition (Z. 078: musicalbesuch) das Öffnen der rechten Hand zu einer *Palm Up* Geste sowie das modalisierende Absenken des Kopfes, deren Rückführung mit der Beendigung des Turns aligniert ist. Es zeigt sich eine feine Abstimmung der koordinierten Ressourcen aufeinander, mit der hier kurzzeitig die satzförmige Begründung auf der geöffneten Handfläche präsentiert wird.

Die von der Sprecherin eingesetzten multimodalen Gestalten unterscheiden sich im Alter von 13 und 15 Jahren hinsichtlich der Reichweite der multimodalen Gestalt sowie der Koordinierung: Während die Sprecherin mit 13 Jahren ein intonationsphrasenübergreifendes Begründungssegment auf der geöffneten Handfläche einer *Palm Up* Geste präsentiert und die Geste über die Intonationsphrasen hinweg hält, präsentiert die Sprecherin mit 15 Jahren eine satzförmige Proposition. Die Reichweite der multimodalen Gestalt wird also mit 15 Jahren reduziert. Zugleich zeigt sich, dass zu diesem Zeitpunkt die Koordinie-

nung der Ressourcen innerhalb der multimodalen Gestalt in der satzförmigen Begründung leichter gelingt, während die temporale Entfaltung der Ressourcen im Alter von 13 Jahren erst verzögert einsetzt. Möglicherweise weisen die Befunde darauf hin, dass der 13-jährigen Sprecherin die zeitliche Koordination der multimodalen Ressourcen in der intonationsphrasenübergreifenden Begründung noch Schwierigkeiten bereitet.

Verwerfen in Begründungen

Interagierende können z. B. durch *Throwing Away* oder *Dropping of Hand* Gesten (vgl. Bressemer und Müller 2014a) Gesprächsgegenstände sichtbar aus der Interaktion ausschließen (vgl. Kap. 8.2.3; 8.6.2). Die heranwachsende Sprecherin kontextualisiert im Alter von 15 Jahren das Verwerfen eines Gesprächsgegenstandes durch die multimodale Gestalt des ablehnenden Fallenlassens (vgl. Kap. 8.6.2), wie das nachfolgende Beispiel 1 illustriert. Dabei projiziert die Sprecherin durch das turninitiale Fallenlassen beider Hände das Ausschließen eines Gesprächsgegenstandes aus der weiteren Interaktion.

Beispiel 1 – Verwerfen in Begründungen (15 Jahre)

082 TO: aber ROM (.) <<p>mit kulturprogramm find ich eigentlich au_noch->
 083 MU: ja find ich [AUCH.]
 084 TO: [SEHR] interessant.
 085 MU: SEHR interessant.
 086 TO: ich (würd) [eigent|lich]
 087 MU: [also |das a|ndere |[ist nur] |
 to-fa |zieht AB hoch & zusammen-----|
 to-gz |blickt zu MU-----|
 088 TO: [m::]
 089 MU: [SHOP]pen oder [sowas-]
 090 TO: |<<creaky, kopfschüttelnd>nei::;>|
 to-gz |blickt auf Blatt----->



090 TO: |['h ja;=] |
 to-gz ->|blickt in Raum->
 to-gs |hebt bHde an|
 091 TO: |=ich |mein kann man ja auch eigentl|ich HIER |mach|en.
 to-gs |lässt bHde fallen| |öffnet lH-|
 to-gz ----->|MU(>Z. 094)
 092 MU: j[a.]|
 093 TO: [a]so|
 to-gz ----->|
 094 |(.) |
 to-gz |blickt n. u.->>
 to-bo |Shoulder Shrug|
 to-fa |facial shrug--|
 095 MU: ja.=
 096 TO: =das find ich jetzt eigentlich UNnötig.
 097 MU: ja.

Während die Mutter die von beiden Sprecherinnen positiv evaluierte Option der Abschlussfahrt nach Rom mit einer anderen Option sprachlich wie auch gestisch kontrastiert (Z. 087), markiert die Tochter überlappend durch das Zusammenziehen der Augenbrauen (Z. 087) sowie die gedehnte Negationspartikel „nei:::“ und das Kopfschütteln (Z. 090) ihre Ablehnung gegenüber der von der Mutter ebenfalls ausgeschlossenen Option. Die Tochter hebt nun turninitial beide Hände an und lässt sie mit Übernahme des *Turns* (Z. 090) fallen: Das Fallenlassen der Hände auf den Tisch, das in einem akustisch wahrnehmbaren Signal resultiert, kontextualisiert projizierend, dass der Gesprächsgegenstand fallengelassen werden wird. In der anschließenden nicht weiter elaborierten Proposition (Z. 091: =ich mein kann man ja auch eigentlich HIER machen.) sowie dem an die Begründung anschließenden Kommentar, „das find ich jetzt eigentlich UNnötig“ (Z. 096), wird der Ausschluss der Option u. a. durch das fokusakzentuierte bewertende Lexem „UNnötig“ (Z. 096) zwar erkennbar. Allerdings expliziert erst die multimodale Gestalt das Fallenlassen des Gesprächsgegenstandes. Die Interagierenden wenden sich daraufhin anderen Gesprächsgegenständen zu.

Die Sprecherin kontextualisiert erstmalig im Alter von 15 Jahren einen Gesprächsgegenstand als verworfen. Durch die multimodale Gestalt des ablehnenden Fallenlassens ist es der heranwachsenden Sprecherin möglich, bereits turninitial das Verwerfen des Gesprächsgegenstandes projizierend anzukündigen – und zwar innerhalb einer nicht weiter elaborierten, satzförmigen Proposition.

10.3.2.2 Vertextung

Auf der Ebene der Vertextung nutzt die heranwachsende Sprecherin unterschiedliche multimodale Ressourcen, die zur Strukturierung ihrer Begründungen beitragen: Im Alter von *13 Jahren* und *15 Jahren* hebt sie einzelne relevante Bedeutungselemente der Begründung kurzzeitig hervor (vgl. Kap. 8.2.1.1) und führt Turnfortsetzungen ein (vgl. Kap. 8.2.3). Wie der nachfolgenden Übersicht 42 zu entnehmen ist, nutzt die Sprecherin mit 13 Jahren außerdem multimodale Ressourcen, um eine Gegenüberstellung innerhalb eines veranschaulichenden Szenarios sichtbar zu machen¹⁰⁵, sowie Gestenwiederaufnahmen, um in einem größeren Äußerungspaket zusammenhängende Begründungselemente zu markieren (vgl. Kap. 8.2.4).

Vertextung			
Alter	12 Jahre	13 Jahre	15 Jahre
multimodale Gestalten und ihre Funktionen zur Vertextung	–	Hervorheben begründungsrelevanter Bedeutungselemente kurzzeitiges Präsentieren und Hervorheben relevanter Bedeutungselemente der Begründung (2)	Hervorheben begründungsrelevanter Bedeutungselemente kurzzeitiges Präsentieren und Hervorheben relevanter Bedeutungselemente der Begründung (1)
		Einführen von Turnfortsetzungen multimodales Einführen einer Turnfortsetzung (1)	Einführen von Turnfortsetzungen multimodales Einführen einer Turnfortsetzung (1)
		Gegenüberstellen in Begründungen kontrastierendes Gegenüberstellen in Begründungen (1)	

¹⁰⁵ Wie die multimodale Gestalt der Anzeige von geteiltem Wissen (vgl. Kap. 10.3.1.3) tritt die multimodale Gestalt des kontrastierenden Gegenüberstellens in Begründungen im Korpus, auf dem die Untersuchung der ersten empirischen Fragestellung fußt, nicht auf. Die multimodale Gestalt wird daher im vorliegenden Kapitel der Fallstudie beschrieben.

		Markieren von Kohäsion und Kohärenz • multimodale Kohäsion und Kohärenz (1)	

Tabelle 42: Funktionen und Anzahl multimodaler Gestalten für das Vertexten (Fokuskind II)

Hervorheben begründungsrelevanter Bedeutungselemente

Zur Strukturierung der eingebrachten Begründungen hebt die Sprecherin sowohl im Alter von *13 Jahren*, im Rahmen satzübergreifender Äußerungspakete (vgl. Beispiel 1), als auch im Alter von *15 Jahren*, während des Einbringens einer satzförmigen Proposition (vgl. Beispiel 2), für die Begründung relevante Bedeutungselemente kurzzeitig hervor (vgl. Kap. 8.2.1.2). Dabei zeigt sich im Vergleich, dass die zeitliche Koordinierung der Ressourcen der Sprecherin im Alter von 13 Jahren noch Schwierigkeiten bereiten, die in dieser Form im Alter von 15 Jahren nicht zu erkennen sind.

Im nachfolgenden Beispiel 1, im Alter von 13 Jahren, verwendet die Sprecherin im Rahmen einer multimodalen Gestalt eine kurzzeitige *Palm Up* Geste, um innerhalb eines sich im Aufbau befindlichen Szenarios ein Bedeutungselement der Äußerung prominent hervorzuheben.

Beispiel 1 – Hervorheben begründungsrelevanter Bedeutungselemente (13 Jahre)

119 TO: (.) ich mein wenn man (.) Abendessen (.) erstmal (.) mit
landestypischen gerichten, (-)
120 dann würde man ja erst mal was über (.)



	ESsens(.)kul tur	er fahren-
to-he	Beat-----	
to-br	zieht AB hoch	
to-gs	 öffnet lH Palm Up	

121 und (.) dann REdet man ja auch während eines essens;=

In der vorliegenden Begründung hebt die Sprecherin das Lexem „ESsenskultur“ als ein für die Begründung relevantes Bedeutungselement hervor: Dies geschieht zum einen prosodisch, denn der Fokusakzent des Wortes liegt auf der ersten Silbe. Zugleich beginnt die Sprecherin aber auch die Vorbereitung einer Gestenäußerung. Unmittelbar nach der Minimalpause innerhalb des Kompositums (Z. 120: ESsens(.)kultur) dreht die Sprecherin dann die Hand zu einer *Palm Up* Geste auf, hebt die Augenbrauen an und führt den Kopf in einem *Beat* nach unten: Die körperlich-sichtbaren Ressourcen erreichen gemeinsam den Höhepunkt ihrer Ausführung am Ende des Lexems. Sie heben an dieser Stelle den Begriff „Essenskultur“ zwar hervor, sind zeitlich aber nicht mit der prosodischen Akzentuierung des Lexems koordiniert.

Im Alter von 15 Jahren zeigt sich nun in der multimodalen Gestalt eine zeitliche Abstimmung der sprachlichen sowie körperlich-sichtbaren Ressourcen aufeinander – allerdings erweist sich die Sichtbarkeit der Geste als eingeschränkt, da sie aus einer Verschränkung der Arme vor dem Körper resultiert, sodass die Handfläche nicht gänzlich nach oben geöffnet, sondern auf die Sprecherin hin orientiert bleibt.

Beispiel 2 – Hervorheben begründungsrelevanter Bedeutungselemente (15 Jahre)

090	TO:		[°h ja;=]
			
091	TO:	=ich	mein kann man ja auch eigent ich HIER mach en.
	to-gs	lässt bHde fallen	öffnet lH-
	to-gz	>>i. d. Raum----->> MU(>Z. 094)	
092	MU:	j[a.]	
093	TO:	[a1]so	
	to-gz	----->	
094		(.)	
	to-gz	blickt n. u.->>	
	to-bo	Shoulder Shrug	
	to-fa	Facial Shrug--	
095	MU:	ja.=	
096	TO:	=das find ich jetzt eigentlich UNnötig.	
097	MU:	ja.	

In der vorliegenden Sequenz wird das Lexem „HIER“ (Z. 091) durch die Koordinierung prosodischer und gestischer Ressourcen hervorgehoben: Die Sprecherin öffnet in zeitlicher Abstimmung mit der Fokusakzentuierung des Lexems die linke Hand (vgl. #1, #2) und betont das für die Begründung relevante Bedeutungselement „HIER“ (Z. 091), mit dem sie darauf Bezug nimmt, dass Aktivitäten wie „SHOpen“ (Z. 089) auch „HIER“ (Z. 091), also in der Heimatstadt der Interagierenden, unternommen werden können.

Die Sprecherin hebt sowohl mit 13 Jahren als auch mit 15 Jahren begründungsrelevante Bedeutungselemente durch den Einsatz multimodaler Gestalten hervor. Es zeigen sich dabei im Alter von 13 Jahren noch Koordinierungsschwierigkeiten in der zeitlichen Alignierung der prosodischen Akzentuierung des Lexems sowie der *Palm Up* Geste – sie entfalten sich in zeitlicher Versetzung zueinander. Mit 15 Jahren gelingt der Sprecherin nun die Koordinierung der Ressourcen: Auf dem Fokusakzent des Lexems erreicht die *Palm Up* Geste ihren Gestenhöhepunkt. Durch die körpernahe Ausführung der Geste bleibt sie in ihrer Sichtbarkeit allerdings begrenzt.

Einführen von Turnfortsetzungen

Wie in Kapitel 8.2.3 dargelegt, können Interagierende durch die Verbindung von gestischen Ressourcen und Strukturelementen wie Konnektoren oder Diskursmarkern eine Turnfortsetzung einführen. Die heranwachsende Sprecherin leitet sowohl im Alter von 13 Jahren als auch mit 15 Jahren das Fortsetzen eines *Turns* mit einer *Palm Up* Geste ein (vgl. Beispiel 1, 2). Zu beiden Erhebungszeitpunkten koordiniert sie dazu i. d. R. die *Palm Up* Geste mit dem Diskursmarker *ich mein*.

Das nachfolgende Beispiel 1 illustriert, wie die Sprecherin im Alter von 13 Jahren die multimodale Gestalt des multimodalen Einführens von Turnfortsetzungen nutzt, um ein größeres Äußerungspaket einzuleiten.

Beispiel 1 – Einführen von Turnfortsetzungen (13 Jahre)

115 TO: °hhh hhh° ich (würd) glaub ich Abendessen machen-
 116 MU: und warum,
 117 TO: |°hh |ich ((lacht)) hh° |
 to-gz >>blickt auf Blatt (>Z. 119)
 to-bo |richtet Oberkörper auf----|
 to-gs |hebt lH, lässt fallen|

118 DANN-



119 |(.) ich mein |wenn man (.) Abendessen (.) erstmal (.)
 to-gz ----->
 to-gs |öffnet lH Palm Up|
 mit landestypischen gerichten, |(-)
 to-gz -----|zu MU (>Z. 122)

120 dann würde man ja erst mal was über (.) ESsens(.)kultur erfahren-
 121 und (.) dann REdet man ja auch während eines essens;=
 122 =ich mein man ISST ja nicht nur; ((räuspert sich))

Nach zwei vorangegangenen Turnabbrüchen (Z. 117-118) setzt die Sprecherin in dem abgebildeten Transkriptausschnitt im Rahmen einer Selbstkorrektur erneut zur Ausführung einer Äußerung an. Sie führt in das anschließende Äußerungspaket mit dem Diskursmarker *ich mein* (vgl. Günthner und Imo 2003) sowie einer *Palm Up* Geste ein (vgl. #1, Z. 119). Durch die zeitliche Synchronisierung des Diskursmarkers und der turninitialen Geste wird das neue Segment ihres *Turns* äußerungsstrukturierend markiert. Ausgerichtet auf das Aufgabenblatt übernimmt die Geste zugleich eine prospektive, referenzierende Funktion auf die Aufgabenstellung. Der Bezug wird mit der Benennung der Option (Z. 119: Abendessen [...]) mit landestypischen gerichten) auch sprachlich expliziert.

Auch im Alter von 15 Jahren koordiniert die Sprecherin den einleitenden Diskursmarker *ich mein* mit einer *Palm Up* Gesten und führt hier in eine satzförmige Begründung ein.

Beispiel 2 – Einführen von Turnfortsetzungen (15 Jahre)

098 MU: und äh (.) SE|geln?
 mu-gz |blickt zu TO (>Z. 107)
 099 (--)
 100 warum NI|CH?|(1.0) |(0.4) |
 to-mo |presst Lippen zusammen|
 to-ns |kräuselt Nase-----|
 to-br |hebt AB zusammengezogen an->
 to-gz |.....|blickt zu MU (>Z. 102)
 101 |(0.4) |(1.0) |(0.5)
 mu-fa |lacht-----|
 to-fa |lacht-|
 to-br ----->|

102 TO: |nee: segeln is nich so.
to-gz |blickt auf Blatt (>Z. 105)

103 nein.



104 al|so |ich mei |n:: |
to-gs |...|dreht rH Palm Up auf Blatt|hebt rH an|
|(..) |
to-gs |lässt rH geräuschvoll fallen|
ich ich persÖNlich mag sowieso nich so |gerne auf_m schiff oder so.
to-gz ----->|blickt zu MU (>Z. 107)

105 MU: ja.
106 TO: glaub ich;
107 MU: okay,

Nachdem die Sprecherin ihre Ablehnung gegenüber der von der Mutter eingebrachten Option (Z. 098) bereits durch faziale Gesten projiziert (Z. 100) und anschließend auch explizit-sprachlich benannt hat (Z. 102), setzt sie nun ihren *Turn* fort und leitet mit dem Ressourcenverbund von Diskursmarker *ich mein* (Z. 104; vgl. Günthner und Imo 2003) und *Palm Up* Geste (vgl. #1) in die Begründung ein. Die Begründung für die Ablehnung der Option wird anschließend satzförmig aufgebaut.

Zu beiden Erhebungszeitpunkten, d. h. sowohl im Alter von 13 Jahren als auch mit 15 Jahren, koordiniert die Sprecherin den initialen Diskursmarker *ich mein* mit einer *Palm Up* Geste und führt mit diesem Ressourcenverbund in Turnfortsetzungen ein. Dabei schließt die Sprecherin im Alter von 13 Jahren ein größeres Äußerungspaket an, während der fortgesetzte Turn der Sprecherin im Alter von 15 Jahren satzförmig verbleibt.

Gegenüberstellen in Begründungen

Im Alter von 13 Jahren, in dem die Sprecherin vermehrt Versuche unternimmt, ihre Begründungen satzübergreifend aufzubauen und global zu strukturieren, macht sie gestisch eine Gegenüberstellung innerhalb eines begonnenen Szenarios sichtbar (vgl. Beispiel 1)

Beispiel 1 – Gegenüberstellen in Begründungen (13 Jahre)



098 TO: =die EInen würden dann ja was ü|ber |IHre |kultur |erfahren;
to-gz >>blickt auf Blatt----->
to-gs |rH Palm Up auf Blatt|tippt--|.....->
mu-gz ----->|auf Blatt----->



099 |und die ANder|en |über i |
 to-gz ----->|
 to-gs->|senkt Palm Up seitl. ab-|Palm Up n. außen|
 mu-gz -----|zu TO (>Z. 102)
 | (0.5) |<<creaky, pp> (unsere)>|
 to-gz |blickt i. d. Raum-----|
 to-gs |Cyclic Gesture-|bHde Palm Up, n. innen--|
 100 ich WEISS es AUch nicht so genau;
 101 MU: ja aber das das [das (-) das eh-]
 102 TO: [ich glaub auch: nicht unbedingt muSE]um;|
 mu-gz ----->|
 103 da GUCKT man ja und da REdet man ja auch nicht und so;
 104 MU: genAU;

Innerhalb eines von der Sprecherin begonnenen Szenarios (vgl. Brünner und Gülich 2002b) stellt die Sprecherin zwei Gruppen von Schüler*innen einander gegenüber (Z. 098-099). Diese werden sprachlich markiert (Z. 098, 099: die EInen / die ANderen) und die Antagonisten durch die jeweilige Fokusakzentuierung prosodisch hervorgehoben. Doch auch gestisch markiert die Sprecherin die Gegenüberstellung: Auf dem Pronomen „IHre“ (Z.098), das Bezug auf „die EInen“ (Z. 098) nimmt, äußert die Sprecherin eine *Palm Up* Geste (vgl. #1), die sie in der Fortsetzung des *Turns* nach außen führt und mit der Benennung der Antagonisten (Z. 099: ANderen) absenkt (vgl. #2): Es zeigt sich, dass die Sprecherin *lexikalisch* und *prosodisch* „die EInen“ (Z. 098) und „die ANderen“ (Z. 099) einander gegenüberstellt – die *Gesten* sind hingegen mit den Lexemen „IHre“ (Z. 099) und „ANDere“ koordiniert.

Die Sprecherin verwendet die multimodale Gestalt des kontrastierenden Gegenüberstellens einzig im Alter von 13 Jahren. In der temporalen Entfaltung der multimodalen Ressourcen werden allerdings noch Koordinierungsschwierigkeiten erkennbar, da die gestischen Ressourcen nicht mit den sich zueinander antagonistisch verhaltenden Inhaltselementen koordiniert werden.

Markieren von Kohäsion und Kohärenz

Zur Strukturierung ihrer Äußerungspakete können Interagierende auch multimodale Kohäsion und Kohärenz markieren (vgl. Kap. 8.2.4). Die Sprecherin der vorliegenden Fallstudie greift im Aufbau global strukturierter Begründungen im Alter von 13 Jahren nun auf Gestenwiederaufnahmen zurück, um Begründungselemente als zusammengehörig zu markieren, wie das nachfolgende Beispiel 1 zeigt.

Beispiel 1 – Markieren von Kohäsion und Kohärenz (13 Jahre)

119 TO: (.) ich mein wenn man (.) Abendessen (.) erstmal (.)
to-gz ----->



|mit landestypischen gerichten, |(-) |
to-gz -----|zu MU (>Z. 122)
to-gs |öffnet LH---|
to-br |zieht AB hoch|

120 dann würde man ja erst mal was über (.)



|ESsens(.)kul|tur |er|fahren-
to-he |Beat-----|
to-br |zieht AB hoch|
to-gs |.....|öffnet LH Palm Up|



121 und (.) dann REdet |man ja auch |während |ei|nes essens;= |
to-gs |öffnet LH---| |öffnet LH Palm Up|
to-sh |zieht Schultern hoch->
to-he |nickt leicht-----|
mu-he |nickt leicht----->



122 =ich mein |man ISST |ja nicht |nur; |((räuspert sich))|
to-gz -----|i. d. Raum-----|zu MU->>
to-gs |.....|öffnet lH|
to-sh ----->|
mu-he -----|
123 MU: ja;
124 gut;

Die jugendliche Sprecherin baut in dem vorliegenden Transkriptauszug ein hypothetisches Szenario in Bezug auf die Option auf, in dem die Austauschschüler*innen mit ihren Gastfamilien zu einem gemeinsamen Abendessen zusammenkommen (Z. 119-122). Dabei reiht sie an die Benennung der Option (Z. 119) zwei Folgen (Z. 120, 121) sowie einen erläuternden Kommentar (Z. 122). Die Sprecherin markiert nun an mehreren Stellen die Kohäsion der einzelnen Begründungssegmente: Sowohl die benannten Folgen (Z. 120, 121) sowie der Kommentar (Z. 122) werden morpho-syntaktisch ähnlich aufgebaut, denn auf äußerungsstrukturierende Markierungen (*dann, und dann, ich mein*), die die Intonationsphrasen jeweils einleiten, folgen finite Verben sowie das generische Pronomen *man* und die Modalpartikel *ja*. Außerdem werden insgesamt fünf *Palm Up* Gesten¹⁰⁶ mit der linken Hand geäußert (vgl. #1-#5), die aufgrund ihrer Platzierung innerhalb der Turnkonstruktionseinheit nun ihrerseits einzelne Begründungselemente als zusammengehörig markieren: So werden in Zeile 119, 120 und 121 jeweils (Anteile von) Nominalphrasen der Begründung (Z. 119-121: mit landestypischen gerichten; Essenskultur; essen), die am Ende der Intonationsphrase eingebracht werden, durch die *Palm Up* Geste hervorhebend präsentiert und kohäsiv miteinander verbunden. Dabei weisen allerdings die Formen der Gesten Unterschiede (vgl. #1, #2) auf. Während sich zwar die erste und die dritte *Palm Up* Geste (vgl. #1, #3) stärker ähneln, erfolgt die erste Geste nicht auf dem Kopf der Nominalphrase (Z. 121: gerichten), sondern auf dem Attribut *landestypischen*. Auch die Zeilen 121 und 122 werden kohäsiv miteinander verbunden, indem die Sprecherin zwei *Palm Up* Gesten (vgl. #3, #5) im Kontext der Evidenz markierenden Modalpartikel *ja* äußert – wenn auch die Koordinierung der zweiten Geste (Z. 122) mit der Partikel *nur* zusammenfällt und nicht mit der Modalpartikel *ja* wie in Zeile 121. Insgesamt zeigt sich in dieser Sequenz, dass die Sprecherin im Alter von 13 Jahren durch unterschiedliche Formen der Markierung die Begründungselemente kohäsiv zu verknüpfen sucht; die Koordinierung der körperlich-sichtbaren Ressourcen mit den jeweiligen Elementen der sprachlichen Äußerung erweist sich allerdings als besonders anforderungsreich, wie die Analyse zeigt. Zu den anderen Erhebungszeitpunkten bleibt die Verwendung der multimodalen Gestalt Markieren von Kohäsion und Kohärenz aus.

¹⁰⁶ Mit Cienki (2021) wird auch das Öffnen der Hand, wie von der Sprecherin der vorliegenden Sequenz ausgeführt, als Form der *Palm Up* Geste verstanden.

10.3.2.3 Epistemische Modalisierung

Epistemische Modalisierungen können dazu beitragen, den Geltungsanspruch eines Arguments zu verstärken oder – durch die Anzeige epistemischer Unsicherheit – auch abzuschwächen (vgl. auch Kap. 8.6). Das zweite Fokuskind setzt im Alter von *13 Jahren* multimodale Ressourcen zur Evidenzmarkierung ein (vgl. Kap. 8.6.1); zeigt aber auch epistemische Unsicherheit an (vgl. Kap. 8.6.3), wie die nachfolgende Tabelle 43 darlegt. Im Alter von *15 Jahren* werden keine multimodalen Gestalten zur epistemischen Markierung eingesetzt.

Epistemische Modalisierung			
Alter	12 Jahre	13 Jahre	15 Jahre
multimodale Gestalten und ihre Funktionen zur epistemischen Modalisierung	–	Evidenzmarkierung - Markieren von Evidenz - durch kurzzeitiges Präsentieren einer Modalpartikel (1)¹⁰⁷ - durch <i>Palm Up</i> Gesten (2) Anzeige epistemischer Unsicherheit - Anzeige von epistemischer Unsicherheit (1)	–

Tabelle 43: Funktionen und Anzahl multimodaler Gestalten zur epistemischen Modalisierung (Fokuskind II)

Evidenzmarkierung

Die Sprecherin markiert ausschließlich im Alter von *13 Jahren* durch den Einsatz multimodaler Ressourcen Evidenz. Das nachfolgende Beispiel 1 illustriert, wie die Sprecherin eine Modalpartikel kurzzeitig präsentierend hervorhebt sowie an weiterer Stelle durch die Verbindung einer *Palm Up* Geste und *Shrug*-Komponenten Evidenz markiert (vgl. Beispiel 1).

Beispiel 1 – Evidenzmarkierung (13 Jahre)

119 TO: (.) ich mein wenn man (.) Abendessen (.) erstmal (.) mit
landestypischen gerichten, |(-)
to-gz >>blickt auf Blatt----->|blickt zu MU (>Z. 122)
120 dann würde man ja erst mal was über (.) ESsens(.)kultur erfahren-



121 und (.) dann REdet |man ja auch |während |ei|nes essens;= |
to-gs |öffnet lH---| |öffnet lH Palm Up|
to-sh |zieht Schultern hoch->
to-he |nickt leicht-----|
mu-he |nickt leicht----->

¹⁰⁷ Das kurzzeitige Präsentieren von Modalpartikeln ist in der Übersicht zur Vertexung (vgl. Kap. 10.3.2.2) nicht berücksichtigt worden, denn das Lexem übernimmt zwar eine relevante pragmatische Funktion für die jeweilige Begründung, allerdings trägt es nicht zu der *inhaltlichen* Entfaltung der Begründung bei.

Anzeige epistemischer Unsicherheit

Beispiel 2 – Anzeige epistemischer Unsicherheit (13 Jahre)

098 TO: =die EInen würden dann ja was über |IHre kultur erfahren;
to-gz >>blickt auf Blatt----->
mu-gz ----->|auf Blatt----->
099 |und die ANderen über i |(0.5) <<creaky, pp> (unsere)>|
to-gz ----->|blickt i. d. Raum-----|
mu-gz ----->|zu TO (i. Z. 102)



271

101 MU: ja aber das das [das (-) das eh-]
 102 TO: [ich glaub auch: nicht unbedingt muSE]um; |
 mu-gz -----> |
 103 da GUCKT man ja und da REdet man ja auch nicht und so;
 104 MU: genAU;

Nach Abbruch des Turns (Z. 099) weist die Sprecherin durch den Kommentar „ich WEISS es AUch nicht so genau;“ (Z. 100) ihre epistemische Unsicherheit aus. Die Äußerung wird mit einer beidhändigen *Palm Up* Geste, dem Hochziehen der Schultern, einer lateralen Bewegung der Hände und intensivierendem Kopfschütteln (vgl. Kendon 2002; McClave 2000) sowie einer Blickausrichtung auf die Rezipientin synchronisiert (vgl. #1, #2).¹⁰⁸ Die Sprecherin zeigt durch die gesamte multimodale Gestalt an, dass es vorerst nicht möglich ist, den abgebrochenen *Turn* fortzusetzen. Durch die Blickzuwendung zur Mutter adressiert die Tochter diese zur Unterstützung in der Fortsetzung der Interaktion, bevor sie anschließend die Begründung wird aufnimmt (Z. 102f.). Zu keinem anderen Zeitpunkt der Erhebung nutzt die Sprecherin diese Form der epistemisch abschwächenden Modalisierung.

10.3.2.4 Zusammenfassende Beschreibung der multimodalen Entwicklungslinie

In der längsschnittlichen Untersuchung des diskursstrukturell schwächeren Kindes zeigt sich insgesamt, dass die Sprecherin mit 13 Jahren im Vergleich zu den anderen Erhebungszeitpunkten eine (nur) leicht erhöhte Anzahl von Begründungen produziert (vgl. Tab. 40), allerdings in der Realisierung der Begründungen ein vielfältiges Repertoire multimodaler Gestalten in unterschiedlichen Funktionen verwendet. Mit 15 Jahren ist dann ein Rückgang in der Anzahl der Begründungen zu verzeichnen. In diesem Alter werden, so wird in den Analysen erkennbar, die eingesetzten multimodalen Gestalten deutlich weniger häufig eingesetzt (vgl. v. a. Kap. 10.3.2.2 & 10.3.2.3). Da die Sprecherin im Alter von 12 Jahren keine multimodalen Gestalten zur Realisierung der argumentativen Handlung des Begründens verwendet, konnte dieser Erhebungszeitpunkt in der Untersuchung der multimodalen Entwicklungslinie nicht berücksichtigt werden. Nachfolgend werden die Befunde der vorangegangenen Kapitel entlang der analytischen Ebenen der Kontextualisierung, Textextung und epistemischen Modalisierung zusammenfassend vorgestellt.

Kontextualisierung

- Die Sprecherin beginnt im Alter von 13 Jahren erstmalig durch die multimodale Gestalt des Präsentierens von (komplexeren) Begründungen einen visuell wahrnehmbaren interpretativen Rahmen ihrer Sprechhandlung aufzubauen. Sie präsentiert Begründungen bzw. Begründungssegmente auf der geöffneten Hand einer *Palm Up* Geste, zeigt allerdings noch Schwierigkeiten in der zeitlichen Koordinierung der Ressourcen. So setzt die gestische Ressource beispielsweise erst mit Verzögerung in der vorangeschrittenen Begründung ein. Die Sprecherin nutzt auch im Alter von 15 Jahren die multimodale Gestalt des Präsentierens von Begründungen. Sie kontextualisiert hier eine satzförmige Begründung als präsentiert. Im Vergleich zum Alter von 13 Jahren zeigt sich hier eine Reduktion der Reichweite der multimodalen Gestalt und zugleich eine feinere zeitliche Koordinierung der sich entfaltenden Ressourcen.
- Zum dritten Erhebungszeitpunkt, als die Sprecherin 15 Jahre alt ist, verwendet sie außerdem eine weitere, vorher noch nicht eingesetzte multimodale Gestalt: das ablehnende Fallenlassen eines Gesprächsgegenstandes. Durch die multimodale Gestalt ist es der Sprecherin möglich,

¹⁰⁸ Debras (2021: 43) nutzt den Terminus der „pragmatic marker [...] as an umbrella term emphasising the multi-functionality of certain discourse items that fulfil interpersonal, modal and/or textual functions in interaction.“

einen Gesprächsgegenstand bereits turninitial und somit projizierend als verworfen zu kontextualisieren, dessen Ausschluss auf der sprachlichen Ebene, d. h. im Rahmen einer satzförmigen Proposition, vornehmlich implizit verbleibt.

Vertextung

- Sowohl im Alter von 13 Jahren als auch im Alter von 15 Jahren hebt die Sprecherin kurzzeitig für die Begründung relevante Bedeutungselemente in ihrem Äußerungskontext hervor. Dabei wird in der mikroanalytischen Untersuchung des Datums der 13-jährigen Sprecherin deutlich, dass es an einer Abstimmung der sprachlichen und körperlich-sichtbaren Ressourcen zum Hervorheben eines Bedeutungselements fehlt, denn diese werden zeitlich versetzt zueinander geäußert: Während die prosodische Markierung durch einen Fokusakzent wortinitial erfolgt, werden die körperlich-sichtbaren Ressourcen erst wortfinal am Ende des Kompositums produziert. Mit 15 Jahren gelingt der Sprecherin nun die Alignierung zwischen der prosodischen Markierung des Lexems und den gestischen Ressourcen: Die *Palm Up* Geste erreicht mit dem Fokusakzent des Lexems ihren Gestenhöhepunkt. Durch die körpernahe Ausführung der Geste wird diese allerdings nicht als vollständig entfaltete und im gemeinsamen Wahrnehmungsraum platzierte *Palm Up* Geste realisiert.
- Zur Strukturierung ihrer Äußerungen führt die Sprecherin sowohl im Alter von 13 Jahren als auch im Alter von 15 Jahren in Turnfortsetzungen durch die Koordinierung des Diskursmarkers *ich mein* und einer *Palm Up* Geste ein. Sie verwendet also die multimodale Gestalt des multimodalen Einführens in Turnfortsetzungen, um mit 13 Jahren in ein größeres Äußerungspaket einzuführen, während sie im Alter von 15 Jahren eine satzförmige Begründung anschließt.
- Im Alter von 13 Jahren nutzt die Sprecherin auch multimodale Ressourcen, um eine kontrastierende Gegenüberstellung innerhalb eines Szenarios zu markieren. Die bereits aufgezeigten Anforderungen der zeitlichen Koordinierung der multimodalen Ressourcen, die sich für die Sprecherin ergeben, werden auch hier sichtbar: Denn es werden nicht die fokusakzentuierten Antagonisten (Z. 098-099: *EInen / ANderen*) gestisch einander gegenübergestellt, sondern die erste *Palm Up* Geste wird mit einem *anderen Bezugselement* (Z. 098: *Ihre*) koordiniert. Die zweite Geste erfolgt dann *nicht in Koordinierung mit dem Fokusakzent* des Lexems, sondern verzögert auf der letzten Silbe. Die Koordinierungsschwierigkeiten ähneln denen, die die Sprecherin auch innerhalb des kurzzeitigen Hervorhebens eines Lexems zeigt – allerdings weist die vorliegende multimodale Gestalt dadurch, dass insgesamt zwei Lexeme der Begründung in zwei Intonationsphrasen kontrastierend hervorgehoben werden, eine größere Reichweite auf.
- Zur Strukturierung ihrer Begründungen nutzt die Sprecherin dann mit 13 Jahren Gestenwiederaufnahmen sowie morpho-syntaktisch ähnlich aufgebaute Turnkonstruktionseinheiten, um die Begründungselemente kohäsiv miteinander zu verbinden. Insgesamt deutet die Koordinierung der Ressourcen darauf hin, dass diese sich als besonders anforderungsreich erweisen: Die von der Sprecherin eingesetzten Gestenwiederaufnahmen werden zum einen an ähnlicher, aber eben nicht gleicher Position in der Intonationsphrase platziert und sie weisen ferner in ihrer äußeren Form durchaus Unterschiede auf. Durch die Anzahl der Gestenäußerungen, die diskrepante Platzierung in der Turnkonstruktionseinheit sowie ihre äußeren Unterschiede wird eine Identifizierung der zusammengehörigen Begründungselemente erschwert. Multimodale Kohäsionsmarkierungen werden dann im Alter von 15 Jahren nicht mehr eingesetzt.

Epistemische Modalisierung

- Im Alter von 13 Jahren nutzt die Sprecherin insgesamt dreimal multimodale Gestalten zur Evidenzmarkierung, die alle innerhalb *einer* satzübergreifenden Begründung produziert werden. Die Sprecherin modalisiert durch die multimodalen Gestalten die eingebrachte Begründung: Der Geltungsanspruch des Arguments wird erhöht, allerdings bleibt das Argument selbst unvollständig und in seiner argumentativen Funktion opak bleibt.
- Zum gleichen Erhebungszeitpunkt, d. h. im Alter von 13 Jahren, setzt die Sprecherin auch eine multimodale Gestalt zur Anzeige von epistemischer Unsicherheit ein. Die multimodale Gestalt erfolgt dabei nach einem Turnabbruch innerhalb des Aufbaus einer global strukturierten Begründung.

Eine Übersicht über die Veränderungen der eingesetzten multimodalen Gestalten sowie der Anzahl der Begründungen wird nachfolgend tabellarisch (vgl. Tab. 44) skizziert.

Alter	12 Jahre	13 Jahre	15 Jahre
Kontextualisierung			
Funktionen der multimodalen Gestalten	–	Präsentieren von Begründungen • Präsentieren von (komplexeren) Begründungen (1) • kurzzeitiges Präsentieren und Hervorheben relevanter Lexeme der Begründung (2)	Präsentieren von Begründungen • Präsentieren von (komplexeren) Begründungen (1) Verwerfen in Begründungen • ablehnendes Fallenlassen eines Gesprächsgegenstandes (1)
Vertextung			
Funktionen der multimodalen Gestalten	–	Hervorheben begründungsrelevanter Bedeutungselemente • kurzzeitiges Präsentieren und Hervorheben relevanter Bedeutungselemente der Begründung (2) Einführen von Turnfortsetzungen • multimodales Einführen einer Turnfortsetzung (1) Gegenüberstellen in Begründungen • kontrastierendes Gegenüberstellen in Begründungen (1) Markieren von Kohäsion und Kohärenz • Multimodale Kohäsion und Kohärenz (1)	Hervorheben begründungsrelevanter Bedeutungselemente • kurzzeitiges Präsentieren und Hervorheben relevanter Bedeutungselemente der Begründung (1) Einführen von Turnfortsetzungen • multimodales Einführen einer Turnfortsetzung (1)
epistemische Modalisierung			
Funktionen der multimodalen Gestalten	–	Evidenzmarkierung • Markieren von Evidenz – durch kurzzeitiges Präsentieren einer Modalpartikel (1) – durch <i>Palm Up</i> Gesten (2) Anzeige epistemischer Unsicherheit • Anzeige von epistemischer Unsicherheit (1)	–
Anzahl der Begründungen			
GESAMT	2	4	3

Tabelle 44: Übersicht über Funktionen der eingesetzten multimodalen Gestalten für das Begründen (Fokuskind II)

10.3.3 Interindividueller Vergleich

In dem vorliegenden Kapitel werden nun die Befunde der Fallstudien der Kapitel 10.3.1 und 10.3.2 vergleichend zusammengeführt. Dabei wird der Fokus auf den *interindividuell* sichtbar werdenden Veränderungen der multimodalen Gestalten liegen, die zur Realisierung der argumentativen Handlung des Begründens eingesetzt werden (vgl. zu den *intraindividuellen* Veränderungen die Zusammenfassungen 10.3.1.4 & 10.3.2.4). Um diese Veränderungen nachvollziehbar darstellen zu können, erfolgt der Vergleich in zwei Schritten: Es wird in einem ersten Schritt dargelegt, *welche* multimodalen Gestalten die Fokuskinder über die drei Erhebungszeitpunkte hinweg jeweils zur Realisierung von Begründungen verwenden. Dies ermöglicht vorerst deskriptiv zu eruieren, ob sich bereits in der Verwendung der multimodalen Gestalten Unterschiede zeigen. In einem zweiten Schritt wird dann in einem mikroanalytisch orientierten Zugriff vorgestellt, welche Unterschiede sich im Einsatz sowie der Realisierung der multimodalen Gestalten interindividuell zeigen. Das Kapitel bereitet somit auch die im Anschluss (vgl. Kap. 10.4) diskutierten Zusammenhänge multimodaler und sprachlich-diskursiver Entwicklungslinien vor.

Fokussiert man auf die erste Frage, welche multimodalen Gestalten die Sprecherinnen insgesamt verwenden, so zeigt sich in der tabellarischen Übersicht 45, dass die Fokuskinder intraindividuell in ihrer Funktion ähnliche multimodale Gestalten zur Realisierung der argumentativen Handlung des Begründens einsetzen und nur drei multimodale Gestalten nicht von beiden Fokuskindern verwendet werden: das retrospektive Anhalten einer vorangegangenen Opposition in Konzessionen, das kontrastierende Gegenüberstellen sowie die Anzeige von Unbestimmtheit.

	Fokuskind I			Fokuskind II		
	12 Jahre	13 Jahre	15 Jahre	12 Jahre	13 Jahre	15 Jahre
Kontextualisierung						
Präsentieren von Begründungen	x		x		x	x
Verwerfen in Begründungen	x					x
Anhalten einer Opposition			x			
Vertextung						
Hervorheben begründungsrelevanter Bedeutungselemente	x				x	x
Einführen von Turnfortsetzungen	x				x	x
Markieren von Kohäsion und Kohärenz	x		x		x	
Gegenüberstellen in Begründungen					x	
Epistemische Modalisierung						
Evidenzmarkierung	x		x		x	
Anzeige geteilten Wissens	x					
Anzeige von epistemischer Unsicherheit	x				x	
Anzeige von Unbestimmtheit			x			

Tabelle 45: Alters- und kindvergleichende Übersicht zu Funktionen der eingesetzten multimodalen Gestalten für das Begründen

Die Übersicht 45 legt insgesamt offen, dass das erste und zweite Fokuskind zur Kontextualisierung, Vertextung und epistemischen Modalisierung multimodale Gestalten in ähnlichen Funktionen einsetzen. Allerdings wird eine zeitliche Differenz hinsichtlich des Gestalteinsatzes ersichtlich, denn die vom ersten Fokuskind bereits im Alter von 12 Jahren frequent eingesetzten multimodalen Gestalten nutzt das zweite Fokuskind erst ab einem Alter von 13 bzw. 15 Jahren. Beide Kinder verwenden multimodale Ressourcen zum Präsentieren sowie Verwerfen von Gesprächsgegenständen (Kontextualisierung); zum kurzzeitigen Präsentieren für die Begründung relevanter Bedeutungselemente, zum präsentierenden

Einführen von Turnfortsetzungen und zur multimodalen Markierung von Kohäsion und Kohärenz (Vertextung) sowie zur Markierung von Evidenz und Anzeige epistemischer Unsicherheit (epistemische Modalisierung). Ein Rückgang in der Frequenz der multimodalen Gestalten, die sie im Alter von 12 Jahren (Fokuskind I) bzw. 13 Jahren (Fokuskind II) noch rege einsetzen, ist dann bei beiden Heranwachsenden bis zum Alter von 15 Jahren zu verzeichnen. Mit dem Rückgang der eingesetzten multimodalen Gestalten im Alter von 15 Jahren wird nun außerdem sichtbar, dass Fokuskind I weiterhin multimodale Gestalt mit einer größeren Reichweite einsetzt, wie Kohäsions- und Kohärenzmarkierung. Multimodale Gestalten wie das Hervorheben relevanter Bedeutungselemente oder das Einführen von Turnfortsetzungen, die eine sehr lokale Reichweite aufweisen, bleiben jedoch aus.¹⁰⁹ Das zweite Fokuskind verwendet hingegen mit 15 Jahren ausschließlich lokal operierende multimodale Gestalten, deren Realisierung weniger hohe Koordinierungsanforderungen birgt, wie die anschließenden qualitativen Auswertungen (vgl. unten, Kontextualisierung, Vertextung) vermuten lassen.

In einem zweiten Schritt werden die Ergebnisse des globalen Vergleichs um mikroanalytische Befunde hinsichtlich der Veränderungen der multimodalen Gestalten (vgl. Kap. 10.3.1 & Kap. 10.3.2) entlang der Analyseebenen der Markierung der Kontextualisierung, der Vertextung und der epistemischen Modalisierung ergänzt.

Hinsichtlich der *Kontextualisierung* zeigt sich in den Analysen beider Fokuskinder, dass sie für ihr Gegenüber mit 12 Jahren (Fokuskind I) bzw. mit 13 Jahren (Fokuskind II) einen interpretativen Rahmen für die Sprechhandlung ihrer eingebrachten Begründungen schaffen. Sie kontextualisieren die Begründungen durch die jeweilige multimodale Gestalt als präsentiert, als verworfen oder setzen multimodale Ressourcen ein, die das Anhalten einer Opposition anzeigen. Es konnte dabei u. a. rekonstruiert werden, dass das Präsentieren in einer satzübergreifenden Begründung für das zweite Fokuskind im Alter von 13 Jahren noch schwierig zu realisieren ist, während das Präsentieren von satzförmigen Begründungen bzw. Begründungssegmenten ohne sichtbare Komplikationen von beiden Sprecherinnen bewältigt wird (vgl. Kap. 10.3.1.1. & Kap. 10.3.2.1). Möglicherweise sind multimodale Gestalten, in denen eine Geste über einen längeren Zeitraum gehalten wird und somit eine größere Reichweite entfaltet, leichter für die heranwachsenden Sprecherinnen mit *lokal* strukturierten Begründungen zu koordinieren (vgl. auch unten, Vertextung). Auch innerhalb des Verwerfens von Begründungen oder Begründungssegmenten zeigen sich Ähnlichkeiten, denn beide Sprecherinnen verwenden Gesten der Familie der *Away* Gesten (vgl. Bressemer und Müller 2014b), um innerhalb einer einzelnen Proposition einen Gesprächsgegenstand als verworfen zu kontextualisieren – elaborierende oder plausibilisierende Begründungen bleiben bei beiden Sprecherinnen aus. Allerdings zeigt der Vergleich beider Fokuskinder, dass sie die multimodale Gestalt sowohl zu unterschiedlichen Zeitpunkten in ihrer Entwicklung verwenden als auch an verschiedenen sequenziellen Positionen im *Turn* platzieren: Fokuskind I nutzt die multimodale Gestalt bereits mit 12 Jahren innerhalb einer eigenaktiv initiierten Begründung. Es benennt dabei zuerst den Gesprächsgegenstand, auf den es sich bezieht, und mit Beginn der Proposition in der anschließenden Turnkonstruktionseinheit verwirft Fokuskind I den Gesprächsgegenstand projizierend. Fokuskind II verwendet die multimodale Gestalt erstmalig im Alter von 15 Jahren. Es verwirft responsiv reagierend einen von der Mutter eingebrachten Gesprächsgegenstand, indem sie ihn turninitial und projizierend, ohne den

¹⁰⁹ Die Bezeichnung der ‚lokal und global operierenden‘ multimodalen Gestalten, die in der vorliegenden Studie gewählt wird, verweist darauf, dass Gesten eine unterschiedlich große Reichweite besitzen: So können Lexeme kurzzeitig zu einem bestimmten Zeitpunkt in einer Intonationsphrase für den Moment präsentierend hervorgehoben werden, während die multimodale Gestalt des Markierens von Kohäsion und Kohärenz sich i. d. R. über mehrere Intonationsphrasen satzübergreifend erst entfaltet. Die Reichweite der multimodalen Gestalt scheint nun unterschiedliche Koordinierungsanforderungen an die Sprecher*innen zu stellen, wie die qualitativen Auswertungen nahelegen.

Gesprächsgegenstand noch einmal zu benennen, mit einer *Dropping of Hand* Geste fallenlässt. Das *eigenaktive* Einbringen des Gesprächsgegenstandes und anschließende Verwerfen, wie es Fokuskind I im Alter von 12 Jahren zeigt, stellt andere Anforderungen an die Kontextualisierungsfähigkeit des Kindes als das responsiv reagierende Verwerfen eines Gesprächsgegenstandes, der von einem*einer anderen Beteiligten bereits einführend benannt worden ist: Fokuskind I schafft somit *selbst* eine kontextuelle Anschlussstelle für das Verwerfen eines Gesprächsgegenstandes. Aus dem interindividuellen Vergleich geht darüber hinaus hervor, dass Fokuskind I im Alter von 15 Jahren eine multimodale Gestalt zur Kontextualisierung einer Sprechhandlung einsetzt, die in ihrer internen Struktur einen hohen Grad an Komplexität aufweist: die multimodale Gestalt des Anhaltens von Oppositionen in Konzessionen (vgl. auch Kap. 8.4.3). Die multimodale Gestalt erweist sich in ihrer internen Struktur als komplex, denn in der Konzession mit dem Muster X, Y, X' (vgl. Couper-Kuhlen und Thompson 2000) wird das Einlenken zur schwach zustimmenden Begründung innerhalb des Äußerungspakets mit einer *Palm Away* Geste koordiniert. Die gestische Ressource wird in fein abgestimmter zeitlicher Alignierung genau an der Scharnierstelle zwischen Einwand und schwacher Zustimmung platziert. Eine multimodale Gestalt mit ähnlicher Komplexität, die funktional für die Kontextualisierung eines Beitrags eingesetzt wird, ist in den Daten von Fokuskind II zu keinem Erhebungszeitpunkt vorzufinden.

Auf der Ebene der *Vertexung* zeigen sich zwischen den Fokuskindern Unterschiede hinsichtlich der Komplexität der eingesetzten multimodalen Gestalten sowie ihrer Koordinierung: So gewinnt die multimodale Kohäsions- bzw. Kohärenzmarkierung des ersten Fokuskindes, das mit 12 Jahren durch die Wiederaufnahme form- und bewegungsgleicher Gesten Begründungselemente als zusammengehörig markiert, im Alter von 15 Jahren deutlich an Komplexität. Denn die Sprecherin platziert in einem *Multi-Unit Turn*, der aus insgesamt sieben Listenelementen besteht, zeitlich aufeinander abgestimmte Hand- und Kopfgesten an jeweils ähnlichen Turnpositionen. Die Bewegungen der körperlich-sichtbaren Ressourcen in den morpho-syntaktisch vergleichbar aufgebauten *Turns* stellen gemeinsam eine Art ‚Rhythmus‘ her, der die Listenelemente als zusammengehörig identifizierbar macht. Da nun nicht die Gesten in ihrer Form und Bewegungsrichtung, sondern ausschließlich ihrer turninternen Platzierung wiederaufgenommen werden, ermöglicht dies, dass die Gesten in ihren weiteren Funktionen flexibel sind, so werden sie beispielsweise in der Aufzählung eingesetzt, um einzelne relevante Bedeutungselemente zu präsentieren oder auch Unbestimmtheit anzuzeigen. Ein Ausbau der Komplexität multimodaler Gestalten ist nun auch bei Fokuskind II insofern zu beobachten, als dass es mit 13 Jahren deutlich den Aufbau global strukturierter Begründungen erprobt und dabei multimodale Kohäsions- bzw. Kohärenzmarkierungen einsetzt. Im Alter von 15 Jahren weisen die Begründungen diese Form der Komplexität jedoch nicht mehr auf. In den durchgeführten mikrostrukturellen Analysen zeigt sich nun ferner, dass die Fokus Kinder unterschiedlich mit der Koordinierung der multimodalen Ressourcen umgehen. V. a. hinsichtlich der multimodalen Gestalten, die zur Vertexung der Begründung beitragen (vgl. aber auch oben, Kontextualisierung), werden Unterschiede sichtbar: Während die multimodalen Ressourcen des ersten Fokuskindes sowohl im Alter von 12 als auch im Alter von 15 Jahren i. d. R. eine feine Abstimmung aufeinander aufweisen, werden bei Fokuskind II v. a. mit 13 Jahren auf der Ebene der Vertexung, als die Sprecherin ihre Begründungen vermehrt satzübergreifend bzw. global strukturiert, hohe Koordinierungsanforderungen bzw. -schwierigkeiten sichtbar. Innerhalb global(er) strukturierter Begründungen hebt sie kurzzeitig für die Begründung relevante Lexeme hervor oder stellt sie kontrastierend einander gegenüber. In der Koordinierung sprachlicher, prosodischer und gestischer Ressourcen werden nun konstant Schwierigkeiten in der zeitlichen Alignierung sichtbar, so laufen beispielsweise prosodische Markierungen den gestischen voraus (vgl. dazu ausführlich Kap. 10.3.2.2). Doch die Anforderungen werden nicht nur hinsichtlich eher global operierender multimodaler Gestalten deutlich, sondern auch in der multimodalen Kohäsionsmarkierung, mit der einzelne Begründungselemente als zusammengehörig

markiert werden. Es zeigt sich, dass in der Koordinierung der Ressourcen die Gesten – anders als bei Fokuskind I – an jeweils abweichenden Stellen in den morpho-syntaktisch parallel aufgebauten Turnkonstruktionseinheiten platziert werden und auch in ihrer Form zueinander Unterschiede aufweisen. Eine Identifizierung der zusammengehörigen Begründungselemente wird, über die aufgezeigten Schwierigkeiten in der Koordinierung hinaus, außerdem durch eine auffällig hohe Anzahl von Gesten erschwert.

Ein weiterer Befund auf der Ebene der Vertextung betrifft nun nicht die Koordinierung der körperlich-sichtbaren Ressourcen im Rahmen der multimodalen Gestalt, sondern mit welchen sprachlichen Ressourcen die multimodale Gestalt – hier: dem multimodalen Einführen von Turnfortsetzungen – erfolgt. Die Fokuskinder unterscheiden sich nun erkennbar darin, ob sie mit Einführen der Turnfortsetzung, bei dem ein Strukturelement auf der geöffneten Handfläche präsentiert wird (vgl. Kap. 8.2.3), bereits sprachlich markieren können, welche *Art* der Turnfortsetzung angeschlossen wird: Beide Kinder setzen die multimodale Gestalt im Rahmen von Begründungen ein. Während Fokuskind I mit dieser multimodalen Gestalt im Alter von 12 Jahren sowohl Folgen in Konditionalgefügen (*dann*) als auch adversative Begründungselemente (*aber*) einleitet und durch die verwendeten sprachlichen Mittel kataphorische Hinweise auf die Art der angeschlossenen Turnkonstruktionseinheit liefert, verwendet Fokuskind II die multimodale Gestalt hingegen mit 13 und 15 Jahren, um den Diskursmarker *ich mein* einzuführen. Die Koordinierung mit dem Diskursmarker *ich mein*, die Fokuskind II einsetzt, verweist nun ausschließlich darauf, *dass* eine neue Konstruktionseinheit eingeführt wird. Eine explizite sprachliche Markierung, die sichtbar machen würde, welche *Art* der Turnfortsetzung erfolgt, bleibt also bei Fokuskind II sowohl im Alter von 13 als auch mit 15 Jahren aus.

Hinsichtlich der Analyseebene der epistemischen Modalisierung zeigen die Fallstudien, dass die Fokuskinder jeweils zu einem Erhebungszeitpunkt *vermehrt* epistemische Modalisierungen verwenden. Bei Fokuskind I betrifft dies den ersten Erhebungszeitpunkt, als die Sprecherin 12 Jahre alt ist; bei Fokuskind II ist der erhöhte Einsatz von epistemischen Modalisierung ein Jahr später, im Alter von 13 Jahren zu beobachten. Beide Sprecherinnen setzen v. a. vermehrt multimodale Ressourcen zur Evidenzmarkierung ein. Im Vergleich erweist sich insgesamt das multimodale Repertoire der multimodalen Gestalten zur epistemischen Modalisierung des ersten Fokuskindes als vielfältiger als das des zweiten (vgl. Kap. 10.3.1.4 & 10.3.2.4). Fokuskind I verstärkt beispielsweise durch einen *Hand Shrug nachträglich* eine vorangegangene eigene Begründung modalisierend oder zeigt an, dass das eingebrachte Wissen in einer Begründung zum *gemeinsamen* Wissen der Interagierenden gehört. Die von Fokuskind II eingesetzten multimodalen Gestalten erweisen sich als weniger variantenreich. Im Verlauf der Entwicklung sinkt dann bei beiden Fokuskindern bis zum Alter von 15 Jahren die Anzahl sowie auch die Varianz der eingesetzten multimodalen Gestalten zur epistemischen Modalisierung.

Insgesamt zeigt der interindividuelle Vergleich also, dass die Fokuskinder sich nicht (maßgeblich) darin unterscheiden, welche multimodalen Gestalten sie zur Realisierung der argumentativen Handlung des Begründens einsetzen, sondern in welchem *Alter* die multimodalen Gestalten eingesetzt werden. Es konnte gezeigt werden, dass *beide* Kinder in jeweils unterschiedlichen Altersstufen eine besonders *hohe Anzahl* multimodaler Gestalten einsetzen sowie dass sich die Fokuskinder im Alter von 15 Jahren in der Verwendung *global bzw. lokal operierender multimodaler Gestalten* sowie der *Komplexität* der multimodalen Gestalten unterscheiden. Ferner konnten Unterschiede in Bezug darauf festgestellt werden, mit welchen *sprachlichen Ressourcen* die multimodalen Gestalten geäußert werden bzw. in welchen *Äußerungskontext* sie eingebettet sind und wie sie *intrapersonell zeitlich koordiniert* werden.

Innerhalb des nachfolgenden Kapitels (vgl. Kap. 10.4) werden diese Befunde nun mit den sprachlich-diskursiven Entwicklungslinien aus Kapitel 10.2 zusammengeführt und eingeordnet; und es wird diskutiert, welche Zusammenhänge zwischen den multimodalen sowie den sprachlich-diskursiven Entwicklungslinien bestehen (können).

10.4 Zusammenhänge zwischen multimodalen und sprachlich-diskursiven Entwicklungslinien der Kinder in der argumentativen Handlung des Begründens

Hinsichtlich der zweiten Forschungsfrage der vorliegenden Studie, *welche Veränderungen sich in einer längsschnittlich angelegten Studie zweier Fokuskinder im Alter von 12, 13 und 15 Jahren hinsichtlich der sprachlich-diskursiven Fähigkeiten sowie des Einsatzes multimodaler Ressourcen zeigen* (vgl. Kap. 5) werden nun die Befunde der Fallstudien der sprachlich-diskursiven und multimodalen Entwicklungslinien zusammengeführt (vgl. Kap. 10.2 & Kap. 10.3). Ziel des Kapitels ist es, die *Zusammenhänge* der sprachlich-diskursiven sowie der multimodalen Entwicklungsverläufe beider Fokuskinder zu rekonstruieren und diese zusammenfassend darzulegen. Der Fokus liegt auf der argumentativen Handlung des Begründens. Einleitend werden zuerst die Anzahl der Begründungen pro Fokuskind und Erhebungszeitpunkt der Anzahl der eingesetzten multimodalen Gestalten vergleichend gegenübergestellt. Dieser Zugang ermöglicht, Veränderungen in der relativen Häufigkeit pro Fokuskind und Erhebungszeitpunkt sichtbar zu machen. Anschließend werden dann entlang der Analyseebenen (Vertextung, Kontextualisierung, epistemische Modalisierung) bedeutsame Zusammenhänge der sprachlich-diskursiven sowie der multimodalen Entwicklungslinien beider Fokuskinder zusammenfassend dargelegt. Es handelt sich hierbei nicht ausschließlich um altersvergleichende Zusammenhänge, sondern es werden auch Befunde präsentiert, die sich aus den unterschiedlichen Diskurskompetenzen der Sprecherinnen ergeben.

Vergleicht man übergeordnet zuerst die eingebrachten Begründungen der Sprecherinnen hinsichtlich ihrer Anzahl sowie der an ihrer Realisierung beteiligten multimodalen Gestalten, so treten, wie die u. s. tabellarische Übersicht 46 zeigt, bestimmte Erhebungszeitpunkte in der Entwicklung der Fokuskinder markant hervor: Es zeigt sich, dass zu den Zeitpunkten, an denen die Fokuskinder die meisten Begründungen produzieren, auch die Anzahl der eingesetzten multimodalen Ressourcen besonders hoch ist. So realisiert Fokuskind I im Alter von 12 Jahren insgesamt 13 Begründungssequenzen und 23 multimodale Gestalten und Fokuskind II im Alter von 13 Jahren 4 Begründungssequenzen mit insgesamt 12 multimodalen Gestalten.

	Fokuskind I			Fokuskind II		
Alter	12 Jahre	13 Jahre	15 Jahre	12 Jahre	13 Jahre	15 Jahre
Anzahl der Begründungen	13	4	4	2	4	3
Anzahl der multimodalen Gestalten	23	-	9	-	12	4

Tabelle 46: Anzahl der Begründungen und multimodalen Gestalten im Vergleich (Fokuskind I & II)

Unter Rückbezug auf Befunde der in Kapitel 4.2.1 vorgestellten Erwerbsstudien zur Entwicklung multimodaler narrativer und explanativer Fähigkeiten lässt sich dieser Befund vorsichtig wie folgt einordnen: Während Kinder zu Beginn des Diskurserwerbs in explanativen und narrativen Diskurspraktiken vornehmlich abbildende Gesten einsetzen, zeigt sich erst mit zunehmendem Diskurserwerb, dass sie auch Gesten mit *pragmatischer* Funktion einsetzen, z. B. zur Äußerungsstrukturierung. Dabei weisen die Erwerbsstudien nach, dass vom vierten bis zum zehnten bzw. elften Lebensjahr ein steter Anstieg von Gesten mit pragmatischen Funktionen zu verzeichnen ist, der sich bis in das Erwachsensein weiterentwickelt (vgl. u. a. für Erklärungen Colletta und Pellenq 2009: 83ff.; vgl. u. a. für (Nach-)Erzählungen

Colletta et al. 2010; Reig Alamillo et al. 2012; Colletta et al. 2015; Graziano 2014b). Die altersvergleichenden Erwerbsstudien zum Erklären und (Nach-)Erzählen zeichnen allerdings diese Entwicklung nur bis zum elften Lebensjahr nach – und vergleichen anschließend die Befunde mit Daten von erwachsenen Sprecher*innen. Die Studien enden also genau zu dem Zeitpunkt, d. h. in dem Alter der Kinder, an dem die vorliegenden erwerbsbezogenen Analysen einsetzen. Wenn nun auf Grundlage der Studien zum Erklären und Erzählen vorsichtig angenommen werden kann, dass Heranwachsende die Verwendung von Gesten mit pragmatischer Funktion nicht nur innerhalb narrativer und explanativer Diskurspraktiken, sondern auch innerhalb des Argumentierens von der Kindheit bis ins adoleszente Alter erst erlernen, kann hinsichtlich der oben angegebenen Befunde (vgl. Tab. 46) angenommen werden, dass sie dabei zunehmend ein Repertoire an sprachlich-diskursiven sowie multimodalen Ressourcen aufbauen (vgl. auch Kap. 4), die dann möglicherweise von den heranwachsenden Sprecher*innen erst erprobt werden, bevor sie sicher bzw. funktional für die argumentative Handlung eingesetzt werden können. Die deutlich erhöhte Anzahl der Begründungen, die die Fokuskinder der vorliegenden Studie im Alter von 12 bzw. 13 Jahren durch ein breites Repertoire an multimodalen Gestalten realisieren, wird daher als ein solches ‚Erproben des Repertoires‘ verstanden.¹¹⁰ Dass nun nicht beide Kinder zum gleichen Zeitpunkt den Anstieg an Begründungen sowie multimodalen Gestalten aufweisen, scheint im individuellen Diskurserwerb der Kinder begründet zu liegen: Die sprachlich-diskursiven sowie multimodalen Fähigkeiten, über die Fokuskind I bereits mit 12 Jahren verfügt, bildet Fokuskind II erst allmählich über die drei Erhebungszeitpunkte aus (vgl. Kap. 10.2). Interessanterweise ist die Anzahl der produzierten Begründungen sowie multimodalen Gestalten dann nach dem beschriebenen Gipfelpunkt bei beiden Fokuskindern rückläufig (vgl. Tab. 46): Bringt man für eine Erklärung nun die sprachlich-diskursiven Fähigkeiten mit den jeweiligen multimodalen Entwicklungslinien in der argumentativen Handlung des Begründens in Zusammenhang, zeichnet sich ab, dass Fokuskind I im Alter von 15 Jahren seine Fähigkeiten dahingehend stark ausbaut, dass es quantitativ weniger Begründungen und multimodale Gestalten hervorbringt, diese aber deutlich an Komplexität gewinnen. Fokuskind II hingegen zeigt im Alter von 15 Jahren sowohl einen Rückgang der Begründungen und der multimodalen Gestalten als auch eine Reduktion der Komplexität. Die Heranwachsende nutzt möglicherweise zum dritten Erhebungszeitpunkt vermehrt lokal strukturierte *Begründungen*, die durch lokal operierende multimodale *Gestalten* realisiert werden, da diese scheinbar leichter für die Sprecherin zu kontextualisieren und zu vertexten sind als die global strukturierten Begründungen zwei Jahre zuvor: Die Reduktion der Komplexität ermöglicht Fokuskind II einen sicheren Umgang mit den kommunikativen Anforderungen der argumentativen Entscheidungsfindung.

Nachfolgend werden nun die Zusammenhänge der multimodalen und sprachlich-diskursiven Entwicklungslinien entlang der bekannten Analyseebenen der Markierung der Kontextualisierung, der Vertexung und der epistemischen Modalisierung dargelegt. Dazu werden auf jeder Analyseebene die Befunde einleitend zusammenfassend benannt und anschließend eingehend erläutert.

Markierung der Kontextualisierung

Auf Ebene der Kontextualisierung legen die Befunde der Fallstudien dar, dass es heranwachsenden Sprecherinnen in der Kontextualisierung ihrer Begründungen vorerst leichter gelingt, sprachliche sowie körperlich-sichtbare Ressourcen miteinander zu koordinieren, die eine geringere Reichweite aufweisen und im Rahmen lokal strukturierter Begründungen entfaltet werden (vgl. dazu auch Kap. 10.3.3). Es

¹¹⁰ Dass Heranwachsende im Erwerb diskursiver Fähigkeiten zu bestimmten Zeitpunkten das erworbene Repertoire z. B. von Mitteln der Vertexung durch/in einem vermehrten Einsatz ‚erproben‘, darauf liefert u. a. auch die Studie von Kinalzik (2024) zu Vertexungsverfahren des Erklärens Hinweise.

zeigt sich ferner, dass die Sprecherinnen die multimodale Gestalt des *Verwerfens* funktional ausschließlich in nicht weiter elaborierten Begründungen einsetzen.

Der Einsatz multimodaler Gestalten schafft einen interpretativen Rahmen für die lokal und global strukturierten begründenden Äußerungen, die somit z. B. als präsentiert kontextualisiert werden können. Beide Fokuskinder verwenden funktional multimodale Gestalten zur Kontextualisierung ihrer Begründungen (vgl. Kap. 10.3.3, Tab. 45). Dabei deuten die Analysen hinsichtlich des Präsentierens von Begründungen bzw. Begründungssegmenten (vgl. Kap. 10.2.2.1 & Kap. 10.3.2.1) darauf hin, dass die Koordinierung gestischer sowie sprachlicher Ressourcen in satzförmigen Begründungen von heranwachsenden Sprecher*innen leichter zu bewältigen ist als in satzübergreifenden global strukturierten Begründungen. Möglicherweise wird in diesem Befund sichtbar, dass „the use of these gestures presupposes the capacity for planning and structuring the discourse as well as the rhetorical management“ (Graziano 2014b: 327), sodass die Koordinierung körperlich-sichtbarer Ressourcen mit satzförmigen Begründungen mit weniger (Planungs-)Aufwand für die Heranwachsenden leichter zu bewältigen ist. Ähnliche Befunde werden auch auf der Ebene der Vertextung sichtbar (vgl. unten, Vertextung). Abgesehen von den Befunden zur Koordinierung multimodaler Ressourcen zeigt sich ferner, dass die Fokuskinder multimodale Gestalten auch funktional zur Kontextualisierung der Sprechhandlung des Verwerfens nutzen. Denn beide Fokuskinder nutzen multimodale Gestalten, um spezifische Gesprächsgegenstände aus der weiteren Interaktion auszuschließen, jedoch ausschließlich im Rahmen nicht weiter elaborierter, satzförmiger Begründungen: Durch die multimodalen Gestalten wird der Ausschluss der Gesprächsgegenstände innerhalb satzförmiger Begründungen also explizit für die weiteren Beteiligten erkennbar. Die Verwendung der multimodalen Gestalt kann nun einerseits ökonomisch begründet sein, da das Verwerfen eines Gesprächsgegenstandes in der argumentativen Handlung des Begründens sich in einer ausschließlich sprachlichen Realisierung aufwändiger gestalten kann. Zugleich kann angenommen werden, dass den genannten multimodalen Gestalten auch für den Erwerb eine wichtige Bedeutung zukommt, denn sie ermöglichen Heranwachsenden, denen der elaborierte Ausschluss eines Gesprächsgegenstandes sprachlich noch nicht gelingt, die Sprechhandlung des Verwerfens durch körperlich-sichtbare Ressourcen visuell wahrnehmbar für alle Beteiligten anzuzeigen.

Markierung der Vertextung

Auf der Ebene der Vertextung zeichnen sich zwischen den sprachlich-diskursiven sowie multimodalen Entwicklungslinien folgende Zusammenhänge ab: Bei Vertextungsschwierigkeiten können multimodale Gestalten strukturschließende Funktionen übernehmen und zum Abschließen des Äußerungspakets beitragen. Es zeigt sich außerdem, dass der Aufbau global strukturierter Äußerungspakete sowohl einer kohärenten Vertextung als auch einer feinen Koordinierung der multimodalen Ressourcen bedarf.

Mit dem Aufbau global strukturierter Begründungen geht auch die Anforderung einher, das Äußerungspaket kohärent und über mehrere Redezüge hinweg aufzubauen. Dies erweist sich für Heranwachsende als durchaus anforderungsreich und wird erst mit zunehmenden sprachlich-diskursiven Fähigkeiten gekonnt (vgl. u. a. Quasthoff 2009; vgl. Kap. 4.1.2). In der vorliegenden Studie weist v. a. Fokuskind II im Alter von 13 Jahren Vertextungsschwierigkeiten im Aufbau globaler Äußerungspakete auf (vgl. Kap. 10.2.2). Nimmt man die sprachlich-diskursiven und multimodalen Zusammenhänge in den Blick, zeigt sich nun, dass diese Schwierigkeiten durch multimodale Gestalten überbrückt werden können: So wird ein Szenario nach mehreren Verzögerungsmarkierungen abgebrochen und dann durch die formelhafte Phrase *ich weiß es auch nicht so genau* sowie eine beidhändige *Palm Up* Geste mit lateraler Bewegung abgeschlossen. Die Sprecherin verwendet hier also die multimodale Gestalt der Anzeige epistemischer Unsicherheit gemeinsam mit einer fallenden Tonhöhenbewegung sowie der Blickzuwendung zur Mutter, sodass eine turnübergaberelevante Stelle entsteht – die Sprecherin schließt also durch eine

multimodale Gestalt, die epistemischer Unsicherheit anzeigt, das sich im Aufbau befindliche Äußerungspaket hier vorerst lokal ab. Im Unterschied zu anderen Formen des Abschließens, wie etwa konkludierenden evaluativen Bewertungen (vgl. Kap. 10.2.1.2), mit denen ein*e Sprecher*in z. B. in Begründungen eine abschließende individuelle Dis-/Präferenz zum Ausdruck bringen kann, übernimmt die multimodale Gestalt der Anzeige epistemischer Unsicherheit hier nun zwar eine *strukturschließende* Funktion, allerdings ermöglicht sie kein *inhaltliches* Abschließen der Begründung. In der vorliegenden Studie trägt die multimodale Gestalt somit nicht nur zur Anzeige epistemischer Unsicherheit bei, sondern sie bewältigt Vertextungsschwierigkeiten, die im Aufbau einer global strukturierten Begründung entstehen, und ermöglicht, die Begründung strukturell kurzzeitig lokal abzuschließen.

Auf der Ebene der Vertextung zeigen sich nun auch Zusammenhänge zwischen der sprachlich-diskursiven Fähigkeit, global strukturierte Begründungen kohärent aufbauen zu können, und der Fähigkeit, in zeitlicher Alignierung multimodale Ressourcen zu multimodalen Gestalten entfalten zu können. Auf Grundlage der Untersuchungen wird deutlich, dass mit sprachlich-diskursiven Vertextungsschwierigkeiten auch *Koordinierungsschwierigkeiten* der multimodalen Ressourcen einhergehen können: So wird für das zweite Fokuskind im Alter von 13 Jahren sichtbar, dass die Sprecherin in den global strukturierten Begründungen hypothetische Szenarios oder Erlebnisberichte entfaltet. Dabei sind im Aufbau der global strukturierten Begründungen vermehrt Häsitationen sowie Abbrüche zu beobachten (vgl. hierzu auch Kap. 10.2.2.2, Beispiel 4, 5), sodass angenommen werden kann, dass die Vertextung der global strukturierten Begründung eine sichtlich hohe Anforderung an die Sprecherin stellt. Diese werden nicht nur auf der sprachlichen Ebene deutlich, sondern zeigen sich auch in intrapersonellen Koordinierungsschwierigkeiten der multimodalen Ressourcen: Lokal operierende multimodale Gestalten, die in den Aufbau der global strukturierten Begründung eingebunden sind, weisen in der Koordinierung der prosodischen Markierungen, d. h. des Fokusakzents, sowie der Ausführung der Geste zeitliche Diskrepanzen auf, wie anhand der Beispiele zum kurzzeitigen Präsentieren von relevanten Bedeutungselementen sowie zum kontrastierenden Gegenüberstellen erkennbar wird (vgl. Kap. 10.3.2.2). Doch auch im Aufbau global operierender multimodaler Gestalten wie dem Markieren multimodaler Kohäsion und Kohärenz kommt es zu Koordinierungsschwierigkeiten, da u. a. die Gestenwiederaufnahmen in morpho-syntaktisch ähnlich aufgebauten Turnkonstruktionseinheiten nicht an vergleichbaren Stellen erfolgen und die Gesten ferner Formunterschiede aufweisen (vgl. Kap. 10.3.2.2). Vertextungsschwierigkeiten betreffen also nicht ausschließlich die sprachliche Ebene, sondern sie scheinen auch mit Schwierigkeiten der Koordinierung multimodaler Ressourcen einherzugehen. Die von Mondada (2014b) beschriebene feinkörnige Abstimmung der Ressourcen erweist sich im Erwerb noch als ausgesprochen anforderungsreich. Diese Koordinierungsschwierigkeiten des zweiten Fokuskindes betreffen allein multimodale Gestalten, die für den Aufbau global strukturierte Äußerungspakete eingesetzt werden – mit 15 Jahren, als die Sprecherin ihre Begründungen ausschließlich lokal aufbaut, gelingt die zeitliche Alignierung der Ressourcen. Kontrastierend ist für das erste Fokuskind festzustellen, dass bei *ausgebauter* Vertextungskompetenz auch die Koordinierung der multimodalen Ressourcen gelingt. So zeigt die Sprecherin mit 12 Jahren bereits die Fähigkeit, sowohl lokale als auch globale Begründungen produzieren und dabei in fein abgestimmter Koordinierung sprachliche und körperlich-sichtbare Ressourcen entfalten zu können. Diese Fähigkeit baut die Sprecherin dann bis zum Alter von 15 Jahren weiter aus, in dem dann die Begründungen sowie die eingesetzten multimodalen Gestalten deutlich an Komplexität und Reichweite gewinnen (vgl. Kap. 10.3.1.2). Diese Befunde deuten darauf hin, dass der Aufbau von Vertextungsfähigkeiten auch die Koordinierung multimodaler Ressourcen einschließt, die in wiederkehrenden Ressourcenverbünden, d. h. multimodalen Gestalten, zur Realisierung argumentativer Handlungen beitragen.

Markierung epistemischer Modalisierung

Hinsichtlich der epistemischen Modalisierung zeigt sich, dass multimodale Evidenzmarkierungen über die drei Erhebungszeitpunkte bis zum Alter von 15 Jahren zurückgehen, aber das diskursstrukturell schwächere Kind noch verbalsprachliche Evidenzmarkierungen in nicht weiter elaborierten Begründungen nutzt, um den Geltungsanspruch eines Arguments zu erhöhen.

Die Befunde der beiden Fallstudien zur epistemischen Modalisierung stützen insgesamt Forschungsergebnisse, die den Zusammenhang zwischen Argumentationskompetenz und epistemischer Modalisierung beleuchten: Diese stellen heraus, dass mit zunehmendem Argumentationserwerb Heranwachsende ihre Beiträge abschwächend modalisieren und diese als verhandelbar zur Disposition stellen können (vgl. Mundwiler et al. 2019; Kreuz und Luginbühl 2020; Morek 2020). In den vorliegenden Daten zeigt sich, dass die Sprecherinnen im Alter von 12 Jahren (Fokuskind I) bzw. 13 Jahren (Fokuskind II) noch eine besonders hohe Anzahl von Evidenzmarkierungen einsetzen, um ihre eingebrachten Begründungen abzusichern: Bei Fokuskind I tragen die Evidenzmarkierungen im Alter von 12 Jahren dazu bei, in satzförmigen Begründungen den Geltungsanspruch des Arguments zu erhöhen. Bei Fokuskind II unterstützen die multimodalen Evidenzmarkierungen im Alter von 13 Jahren Begründungen, die eine fehlende oder wenig explizite argumentative Funktion aufweisen (vgl. Kap. 10.2.2.3), indem die fehlende Explikation durch einen erhöhten Geltungsanspruch ‚ausgeglichen‘ wird (vgl. Kap. 10.3.2.3). Die Sprecherinnen markieren also durch die Verwendung von Evidenzmarkierungen einen Grad an Faktizität ihrer Argumente, v. a. dann, wenn diese nicht weiter plausibilisiert werden oder ihre jeweilige argumentative Funktion nicht expliziert wird. Zum letzten Erhebungszeitpunkt zeigen sich diese Formen der multimodalen Evidenzmarkierung dann rückläufig, wenn auch das zweite Fokuskind im Alter von 15 Jahren in zwei seiner drei nicht elaborierten Begründungen weiterhin verbale Evidenzmarkierungen in der o. g. Funktion nutzt. Fokuskind I hingegen verwendet zum letzten Erhebungszeitpunkt v. a. *verba sentiendi*, um seine Begründungen abgeschwächt als vorläufig explorierend zu rahmen (vgl. Kap. 10.2.2.4), sodass diese in der Interaktion aufgenommen und weiterverhandelt werden können.

Insgesamt konnten in dem vorliegenden Kapitel 10.4 relevante Einsichten in die Zusammenhänge zwischen den sprachlich-diskursiven und den multimodalen Entwicklungslinien gewonnen werden. So wird anhand der explorativen Fallstudien der zwei Fokuskinder grundsätzlich sichtbar, dass der Auf- und Ausbau argumentativer Diskurskompetenz mit der Fähigkeit einhergeht, multimodale Ressourcen zu multimodalen Gestalten koordinieren zu können. Dabei zeigen sich die folgenden globalen Zusammenhänge in den Entwicklungsverläufen:

- In der Entwicklung der Heranwachsenden tritt jeweils ein Erhebungszeitpunkt markant hervor, an dem die Kinder ihr Repertoire erworbener diskursiver sowie multimodaler Fähigkeiten vermehrt erproben. Der Einsatz der multimodalen Gestalten, die zur Realisierung von Begründungen eingesetzt werden, sowie die Anzahl der eingebrachten Begründungen gehen dann bis zum Alter von 15 Jahren zurück (vgl. Fokuskind I & II).
- Diskursstrukturell schwächere Heranwachsende können Schwierigkeiten im Aufbau von global strukturierten Begründungen, d. h. auf der Ebene der Vertextung zeigen: Es gelingt den Heranwachsenden dann nur schwer, ein kohärentes Äußerungspaket zu produzieren und es abzuschließen. Multimodale Gestalten, wie die Anzeige epistemischer Unsicherheit, können bei diesen Jugendlichen funktional für den Abschluss des Äußerungspakets eingesetzt werden. Ferner können multimodale Evidenzmarkierungen verwendet werden, um in Begründungen, oder die nicht weiter elaboriert werden oder deren argumentative Funktion nicht expliziert wird, den Geltungsanspruch der Redebeiträge zu erhöhen und die beschriebenen Aspekte ‚auszugleichen‘.

Letztlich erweist sich für diskursstrukturell schwächere Heranwachsende auch die Koordinierung der multimodalen Ressourcen in global strukturierten Äußerungspaketen als anforderungsreich (vgl. Fokuskind II). Diskursstrukturell versierteren Sprecher*innen hingegen, denen der kohärente Aufbau global strukturierter Äußerungspakete gelingt, zeigen die Fähigkeit, die eingesetzten multimodalen Ressourcen zeitlich fein aufeinander abgestimmt zu koordinieren (vgl. Fokuskind I).

- In lokal operierenden multimodalen Gestalten gelingt sowohl diskursstrukturell stärkeren wie auch diskursstrukturell schwächeren Kindern die zeitliche Koordinierung der multimodalen Ressourcen leichter als in global operierenden multimodalen Gestalten (vgl. Fokuskind I & II). Dieser Befund zeigt sich sowohl auf der Ebene der Vertextung als auch der Kontextualisierung.
- Die Begründungen der sprachlich-diskursiv stärkeren Heranwachsenden gewinnen mit Ausbau der Argumentationskompetenz sowohl sprachlich als auch multimodal an Komplexität (vgl. Fokuskind I), während die Komplexität der Begründungen sowie die Komplexität der dazu eingesetzten multimodalen Gestalten diskursstrukturell schwächerer Heranwachsender mit dem ‚Erproben des Repertoires‘ erst ansteigt und schließlich in der Komplexität deutlich reduziert wird (vgl. Fokuskind II).

Es folgt nun, in Kapitel 10.5, eine zusammenfassende Einordnung der Befunde hinsichtlich ihrer erwerbstheoretischen Relevanz.

10.5 Zusammenfassung und Ausblick: Hypothesen zum Erwerb

Anhand der explorativen Fallstudien zweier Fokuskinder ist das vorliegende Kapitel 10 dem zweiten Forschungsvorhaben der Arbeit nachgegangen, *welche Veränderungen sich in einer längsschnittlich angelegten Studie zweier Fokuskinder im Alter von 12, 13 und 15 Jahren hinsichtlich der sprachlich-diskursiven Fähigkeiten sowie des Einsatzes multimodaler Ressourcen zeigen* (vgl. Kap. 5). Dazu wurden die sprachlich-diskursiven sowie die multimodalen Entwicklungslinien rekonstruiert und ihre Zusammenhänge (vgl. 10.4) dargelegt. Im Folgenden sollen diese Befunde nun hinsichtlich ihrer erwerbstheoretischen Relevanz zusammenfassend eingeordnet sowie daraus resultierende Hypothesen zum multimodalen Diskurserwerb abgeleitet werden.

Das Forschungsvorhaben der vorliegenden Fallstudie knüpft an bisherige Befunde zum multimodalen Diskurserwerb narrativer und explanativer Fähigkeiten an (vgl. zum (Nach-)Erzählen Colletta et al. 2010; Reig Alamillo et al. 2012; Colletta et al. 2015; vgl. zum Erklären Colletta und Pellenq 2009; Reig Alamillo et al. 2012; vgl. zsf. Kap. 4.2.1) und adressierte dabei die in Kapitel 4 dargelegte Forschungslücke, dass qualitative Untersuchungen zum multimodalen Diskurserwerb des Argumentierens sowie Untersuchungen von Kindern höherer Altersklassen noch ausstanden. Aufgrund der Längsschnittdaten konnten in der vorliegenden Untersuchung sowohl die sprachlich-diskursiven als auch multimodalen Entwicklungslinien zweier Fokuskinder in der Altersspanne von 12 bis 15 Jahren rekonstruiert und dabei relevante Zusammenhänge beider Entwicklungslinien herausgearbeitet werden. Der methodische Zugang erlaubte es, einerseits globale Veränderungen zu ermitteln, die aufdecken, wie sich die Anzahl der Begründungssequenzen sowie der verwendeten multimodalen Gestalten verändert. Andererseits ermöglichen dann die mikroanalytischen Untersuchungen, auch einzelne Facetten des multimodalen Argumentationserwerbs zu rekonstruieren, die sich über die Altersspanne von 12 bis 15 Jahren verändert haben und sich daher für nachfolgende Erwerbsforschung als besonders fruchtbar erweisen können: Es handelt sich um die Facetten der *Reichweite*, der *Koordinierung* sowie der *Komplexität* der multimoda-

len Gestalten. Nachfolgend werden in Bezug auf die Facetten die wichtigsten Befunde zusammenfassend vorgestellt und aus ihnen heraus erste Erwerbs hypothesen abgeleitet. Da in der Untersuchung außerdem das *Erproben der Repertoires* besonders markant hervortrat, das nicht mit einer spezifischen Erwerbsfacette assoziiert ist, sondern als übergeordneter Befund verstanden wird, werden auch diesbezüglich die Untersuchungsergebnisse der vorliegenden Studie zusammengefasst und eine erwerbsbezogene Hypothese abgeleitet.

Befund & Erwerbshypothese zur Reichweite: Die qualitativen Analysen der Argumentationsdaten der Fokuskinder decken insgesamt auf, dass sich die multimodal realisierten Begründungen hinsichtlich ihrer Reichweite sowie den dadurch entstehenden Anforderungen an die heranwachsenden Sprecher*innen unterscheiden. Dementsprechend gelingt die Realisierung *satzförmiger* Begründungen mit *lokal* operierenden multimodalen Gestalten scheinbar früher als die Realisierung global operierender multimodaler Gestalten in global strukturierten Begründungen. So zeigt Fokuskind II mit 13 Jahren in dem Versuch, Begründungen zu größeren Äußerungspaketen zu vertexten, Schwierigkeiten, darin multimodale Gestalten mit einer globalen Reichweite zu realisieren wie multimodale Kohäsions- und Kohärenzmarkierungen. Die Realisierung lokal operierender multimodaler Gestalten hingegen wie das Verwerfen eines Gesprächsgegenstandes gelingt Fokuskind II im Alter von 13 Jahren in einer satzförmigen Begründung ohne sichtbare Komplikationen. Möglicherweise erfordert das Einbringen von Begründungen und multimodalen Gestalten mit globaler Reichweite höhere Anforderungen hinsichtlich zu leistender Planungsprozesse (vgl. auch Graziano 2014b). Aus diesem Befund ableitend kann nun als *Erwerbshypothese* angenommen werden, dass mit fortschreitendem multimodalen Diskurserwerb die Reichweite der realisierten Begründungen sowie der eingesetzten multimodalen Gestalten zunimmt.

Befund & Erwerbshypothese zur Koordinierung: Mit den o. g. Befunden gehen ferner wichtige Erkenntnisse hinsichtlich der Koordinierungsanforderungen einher: Die Koordinierung multimodaler Ressourcen erweist sich für Heranwachsende in Äußerungen mit lokaler Reichweite leichter als in global strukturierten Äußerungspaketen. Die zeitliche Koordinierung global operierender multimodaler Gestalten in global strukturierten Äußerungspaketen gelingt Heranwachsenden erst mit ausgebaute(re)n multimodalen diskursiven Fähigkeiten – hier werden die Ressourcen dann zeitlich sowie intrapersonell fein aufeinander abgestimmt, wie die Untersuchungen zu Fokuskind I sowohl im Alter von 12 als auch im Alter von 15 Jahren zeigen. Weisen die Sprecher*innen hingegen noch Schwierigkeiten in der Strukturierung globaler Äußerungspakete auf, wie Fokuskind II im Alter von 13 Jahren, werden diese auch in der Entfaltung der körperlich-sichtbaren Ressourcen erkennbar – so zeigen sich z. B. Verzögerungen in der zeitlichen Alignierung der Ressourcen, die u. a. dazu führen, dass die multimodale Gestalt ihre eigentliche Funktion nur bedingt entfalten kann. Aus diesem Befund kann als *Erwerbshypothese* abgeleitet werden, dass sich lokal operierende multimodale Gestalten zur Realisierung lokal strukturierter Begründungen als weniger anforderungsreich hinsichtlich der zeitlichen *Koordinierung* der Ressourcen erweisen. Ihre Koordinierung, i. S. einer zeitlichen Alignierung der Ressourcen, wird von den Heranwachsenden früher bewältigt als die Koordinierung der Ressourcen in global operierenden multimodalen Gestalten, die im Rahmen global strukturierter Äußerungspakete auftreten. Der Befund zur Koordinierung der multimodalen Ressourcen legt außerdem frei, dass multimodale Gestalten scheinbar nicht als holistische Gestalten erworben werden – denn dann würde die Koordinierung nicht als besondere Erwerbsanforderung in der Studie hervortreten. Insgesamt erweist sich die Fähigkeit einer zeitlich fein abgestimmten Koordinierung in der Entfaltung der multimodalen Ressourcen als eine relevante Kompetenzausprägung heranwachsender Sprecher*innen.

Befund & Erwerbshypothese zur Komplexität: Im Rahmen der qualitativen Untersuchung konnte nachgewiesen werden, dass mit fortschreitendem Diskurserwerb die Begründungen der Heranwachsenden sich auch hinsichtlich ihrer Komplexität verändern können. So konnte u. a. gezeigt werden, dass die Begründungen im Alter von 12 bzw. 13 Jahren (zunehmend) zu größeren Äußerungspaketen aufgebaut werden und sowohl in ihrer sprachlichen Struktur als auch den eingesetzten multimodalen Ressourcen bzw. multimodalen Gestalten an Komplexität gewinnen. Dieser Grad der Komplexität nimmt bei der diskursstrukturell stärkeren Sprecherin im Alter von 15 Jahren weiter zu. In diesem Alter erfolgt nun das Markieren multimodaler Kohäsion bzw. Kohärenz nicht allein durch z. B. lexikalische Wiederholungen sowie Gestenwiederaufnahmen, wie beispielsweise im Alter von 12 bzw. 13 Jahren, sondern in der gesamten Gestalt eine Art ‚Rhythmus‘ erzeugt wird, der die zusammenhängenden Begründungssegmente erkennbar macht. Dies trägt auch zum Verständlichmachen des Äußerungspakets für das Gegenüber bei. Aus diesem Befund lässt sich als *Erwerbshypothese* ableiten, dass sich mit zunehmendem Diskurserwerb, möglicherweise insbesondere für diskursstrukturell stärkere Kinder, der Grad der Komplexität in multimodal realisierten argumentativen Handlungen wie dem Begründen erhöht.

Befund & Erwerbshypothese zum Erproben: Die Untersuchung zeigte außerdem in dem globalen Zugriff über die Veränderungen der Begründungssequenzen sowie multimodalen Gestalten als weiteren *Befund*, dass die Fokuskinder im Alter von 12 bzw. 13 Jahren die ihnen zur Verfügung stehenden sprachlichen sowie körperlich-sichtbaren Ressourcen intensiv einsetzen. Als daraus resultierende *Erwerbshypothese* kann abgeleitet werden, dass Heranwachsende das erworbene Repertoire der Ressourcen ‚erproben‘ und mit fortschreitendem Kompetenzerwerb dann ausgewählt einsetzen.¹¹¹ Möglicherweise greifen dann diskursstrukturell schwächere Kinder, die Koordinierungsschwierigkeiten (s. oben) innerhalb des Aufbaus global strukturierter Äußerungen zeigen, im weiteren Verlauf ihrer Entwicklung (vorerst) auf Begründungen und multimodale Gestalten mit lokaler Reichweite zurück, die leichter zu realisieren sind. Diskursstrukturell stärkere Kinder bauen ihre Begründungen sprachlich sowie multimodal jedoch in der Komplexität weiter aus.

Zusammenfassend konnten anhand der explorativen Fallstudien erste Erwerbshypothesen hinsichtlich eines multimodalen Diskurs- bzw. Argumentationserwerbs abgeleitet werden, die aufgrund der Datengrundlage aber als explorative Befunde eingeordnet werden. Durch den qualitativen Forschungszugang konnten bisherige Erwerbsstudien zu anderen (monologischen) Diskurspraktiken, die die Häufigkeit eingesetzter Gesten in Abhängigkeit zur Länge der Diskurseinheit untersuchen, um detaillierte Einsichten ergänzt werden. Denn durch den analytischen Zugriff über die in Kapitel 8 rekonstruierten multimodalen Gestalten konnte ermittelt werden, welche Ressourcenverbünde die heranwachsenden Sprecherinnen der längsschnittlichen Studie zur Markierung der Kontextualisierung, Vertextung und epistemischen Modalisierung von Begründungen insgesamt einsetzen und wie diese sich über den Erhebungszeitraum verändern. Der mikroanalytische Zugang ermöglichte dann, nicht nur zu ermitteln, ob sich Veränderungen in den eingesetzten multimodalen Gestalten zeigen, sondern auch, ob Veränderungen hinsichtlich der zeitlichen und/oder intrapersonellen Koordinierung sichtbar werden. Indem sowohl die sprachlich-diskursiven als auch die multimodalen Entwicklungslinien anhand des etablierten Modells von Argumentationskompetenz (vgl. Quasthoff 2009; Krah et al. 2013; Morek et al. 2017) untersucht worden sind, konnten die Entwicklungslinien zusammengeführt sowie Erwerbshypothesen hinsichtlich eines multimodalen Diskurserwerbs abgeleitet werden. Diese Befunde bestärken die Annahme, den Diskurserwerb von Kindern als *multimodalen Diskurserwerb* zu verstehen, und zukünftig multimodale Res-

¹¹¹ Interessanterweise zeigt sich in der Spracherwerbsforschung, dass mit Erreichen bestimmter Erwerbsziele die Häufigkeit von kindseitig eingesetzten Gesten rückläufig ist (vgl. Andr n 2014).

sources innerhalb des (argumentativen) Diskurserwerbs konsequent mit zu berücksichtigen. Denn multimodale Ressourcen in Form multimodaler Gestalten können von heranwachsenden Sprecher*innen durchaus hochgradig funktional für das mündliche Argumentieren eingesetzt werden. Wie sich nun aber in der Entwicklung argumentativer Fähigkeiten der Erwerb multimodaler Gestalten vollzieht, kann im Rahmen der vorliegenden Studie nicht beantwortet werden. Diesbezüglich bedarf es weiterer Untersuchungen zum multimodalen Argumentationserwerb.

Im Anschluss an die erwerbsbezogene Untersuchung des vorliegenden Kapitels 10 wird im Folgenden auf Grundlage beider empirischer Kapitel die Funktionalität multimodaler Ressourcen für die kommunikative Gattung des Argumentierens und für den Erwerb argumentativer Diskurskompetenz diskutiert (vgl. Kap. 11).

11 Diskussion: Funktionalität multimodaler Ressourcen für die kommunikative Gattung des Argumentierens und den Erwerb argumentativer Diskurskompetenz

Das Ziel der vorliegenden Arbeit besteht darin, auf der Grundlage eines Datenkorpus von argumentativen Eltern-Kind-Interaktionen die Funktionen multimodaler Ressourcen für das mündliche Argumentieren sowie den Erwerb argumentativer Diskurskompetenz zu untersuchen. Um die Erkenntnisse der empirischen Untersuchungen der Kapitel 8 und 10 abschließend zusammenzufassen, werden im Folgenden die Forschungsfragen (vgl. Kap. 5) entlang übergeordneter Thesen beantwortet. Strukturell folgt der Aufbau des Kapitels dabei der Zweiteilung der empirischen Untersuchung: Es werden zuerst die Befunde der systematischen Beschreibung aus Kapitel 8 und anschließend die erwerbsbezogenen Erkenntnisse aus Kapitel 10 vorgestellt.

Forschungsfrage I: Welche Funktionen übernehmen multimodale Ressourcen für das mündliche Argumentieren?

I.a.) An der Realisierung argumentativer Handlungen sind wiederkehrende multimodale Ressourcenverbünde, i. S. multimodaler Gestalten, beteiligt: Die multimodalen Gestalten prägen die argumentativen Handlungen insofern, als dass sie performativ zu ihrer interpretativen Rahmung beitragen sowie die argumentative Handlung strukturieren und modalisieren. Darüber hinaus können sie interaktionsorganisatorisch auf die Handlung des Gegenübers wirken.

In der vorliegenden Studie konnten vielfältige Funktionen multimodaler Ressourcen für das Argumentieren rekonstruiert werden, die in der bisherigen Forschung nur punktuell für argumentative Interaktionen untersucht worden sind (vgl. u. a. Jacquin 2018; Heller 2018a, 2021b; Marrese et al. 2021; Heller et al. 2023). Indem die *argumentativen Handlungen*, wie das Begründen oder Einwenden (vgl. Kap. 2), mittels derer Interagierende die Entscheidungsaufgabe bearbeiten, zum Ausgangspunkt der Untersuchung gemacht worden sind, war es möglich, die eingesetzten multimodalen Ressourcen hinsichtlich ihrer Funktion für die Gattung des mündlichen Argumentierens zu rekonstruieren. Durch das mikroanalytische Vorgehen konnten multimodale Gestalten ermittelt werden, die wiederkehrend an der Realisierung argumentativer Handlungen beteiligt sind – und mittels derer wiederum die Interagierenden die kommunikativen Aufgaben innerhalb der diskursiven Praktik bearbeiten (vgl. Kap. 2.2). Der Ressourcenverbund, mit dem eine multimodale Gestalt konfiguriert wird, weist trotz der Unterschiede, die sich in der kontextsensitiven Einpassung der multimodalen Gestalt ergeben, ähnliche Charakteristika auf (vgl. Mondada 2014b, 2018; vgl. auch Kap. 3.1.4): Multimodale Gestalten können aufgrund der beteiligten Ressourcen, ihrer systematisch aufeinander abgestimmten temporalen Entfaltung und ihrer kontextuellen sowie sequenziellen Einbettung als zusammengehörig identifiziert werden. Dabei sind die in der vorliegenden Untersuchung rekonstruierten multimodalen Gestalten als eine spezifische *Ausprägung der Erscheinungsform* der jeweiligen argumentativen Handlung zu verstehen (vgl. Kap. 9).

Es zeigte sich, dass die multimodalen Gestalten für die Realisierung der argumentativen Handlungen sowohl interpretative, d. h. performative, modalisierende, aber auch interaktive und äußerungsstrukturierende Funktionen übernehmen. Die multimodalen Gestalten geben beispielsweise *performativ* Hinweise auf das Verwerfen einer Option, sie können *äußerungsstrukturierend* zusammenhängende Begründungssegmente anzeigen, eine argumentative Handlung epistemisch *modalisierend* abschwächen oder verstärken sowie die Rückmeldung des Gegenübers *interaktionsorganisatorisch* in ein Listenformat steuern. Insgesamt konnten 25 multimodale Gestalten, die an der Realisierung argumentativer

Handlungen beteiligt sind, Dyaden übergreifend rekonstruiert werden. Sie wurden entlang ihrer *Funktionen* systematisiert (vgl. Kap. 8). Ohne an dieser Stelle nochmals auf die einzelnen Funktionen der multimodalen Gestalten im Detail einzugehen (vgl. dazu ausführlich Kap. 8. & Tab. 33), soll nachfolgend resümiert werden, welche *Relevanz* den Gestalten für das mündliche Argumentieren zukommt. Dazu wird beleuchtet, welches Potenzial in ihrer ‚Explizitheit bzw. Implizitheit‘ (vgl. Deppermann 2013) hinsichtlich argumentativer Entscheidungsfindungen liegt.

In argumentativen Entscheidungsfindungen, in denen „talk is potentially at risk“ (Schönfelder und Heller 2019: 86), können körperlich-sichtbare Ressourcen u. a. Einstellungen des*der Sprecher*in oder Sprechhandlungen anzeigen, ohne dies explizit-sprachlich zu markieren. So können Sprecher*innen beispielsweise durch Gesten der Familie der *Away Geste*, wie der *Dropping of Hand* oder *Throwing Away Geste* spezifische Gesprächsgegenstände aus der weiteren Interaktion ausschließen und somit individuelle Dispräferenzen für alle Beteiligten sichtbar machen (vgl. Kap. 8.2.5; vgl. Kap. 8.6.1.1). In der Literatur werden dabei die körperlichen *Displays* (vgl. Deppermann 2013: 4) u. a. als ‚implizit‘ und ‚ambig‘ beschrieben, da ihre Bedeutung und Funktion aus einem (multimodalen) konversationsanalytischen Verständnis heraus nicht allein anhand des nächsten Redebeitrags überprüft werden kann, sondern unter Berücksichtigung ethnographischen und praxeologischen Wissens zu rekonstruieren ist (vgl. Deppermann 2013; vgl. Kap. 7.1). Die vorliegende Arbeit bezieht den Begriff der ‚Implizitheit‘ nun nicht auf die *Ambiguität* körperlich-sichtbarer Ressourcen – nicht zuletzt, weil gezeigt werden konnte, dass körperlich-sichtbaren Ressourcen im Rahmen ihrer multimodalen Gestalten durchaus eindeutige Bedeutungen und Funktionen zukommen –, sondern sie versteht Implizitheit als Pendant zu explizit-sprachlichen Markierungen. Die Implizitheit körperlich-sichtbarer Ressourcen ermöglicht es den Interagierenden in argumentativen Entscheidungsfindungen, die sprachliche Angriffsfläche zu reduzieren. Es konnte beispielsweise in der multimodalen Gestalt des Präsentierens in Einwänden (vgl. Kap. 8.4.1, Beispiel 2) herausgearbeitet werden, dass Interagierende Anteile von vorangegangenen Begründungen des Gegenübers auf der geöffneten Handfläche an diesen zur Überarbeitung zurückgeben können: In dem genannten Beispiel wiederholt der Sohn unmittelbar an eine Begründung des Gegenübers anschließend ein einzelnes Lexem der Vorgängeräußerung, richtet seine Hand als *Palm Up Geste* auf sein Gegenüber und gibt durch die multimodale Gestalt einen Aspekt der Begründung als ‚korrekturbedürftig‘ präsentierend an sein Gegenüber zurück. Die sprachliche Oberfläche der Äußerung, die potenziell vom Gegenüber aufgegriffen und widerlegend bearbeitet werden könnte, wird dabei auf ein Minimum reduziert (vgl. zur Reduktion von Konfrontation durch interpersonelle Bezugnahme auch Jacquin 2018), wohingegen die performative Sprechhandlung *explizit* durch die körperlich-sichtbaren Ressourcen ausgedrückt wird. Multimodale Ressourcen übernehmen also wichtige kommunikative Funktionen hinsichtlich der Gattung des mündlichen Argumentierens.

I.b.) Sowohl die temporale Entfaltung und Dauer der eingesetzten multimodalen Ressourcen, die erst in ihrer Verbindung miteinander eine multimodale Gestalt bilden, als auch ihre sequenzielle Einbettung sind für die Bedeutung und Funktion der Gestalt wesentlich.

Multimodale Ressourcen entfalten sich innerhalb der multimodalen Gestalt in zeitlicher Relation und in eng aufeinander abgestimmter intrapersoneller Koordinierung (vgl. u. a. Deppermann und Schmitt 2007; Stukenbrock 2015). Die Bedeutung und Funktion einer multimodalen Gestalt ist nun zum einen von der Entfaltung und Dauer der intrapersonell koordinierten Ressourcen abhängig sowie von den Ressourcen selbst, die in ihrer temporalen Entfaltung als einzelne Schichten, Mondada (2016: 351) spricht auch von „multiple layers of multimodal resources“, aufeinander laminiert sind. Sie werden bedeutsam *in* und *für* den sequenziellen Kontext, in dem sie entstehen.

Die als einzelne Lagen bzw. Schichten eingesetzten Ressourcen tragen wesentlich zu ihrer Bedeutungskonstitution bei. Eine mikroanalytische und den sequenziellen Kontext berücksichtigende Untersuchung ermöglicht es, auch äußerlich scheinbar ähnliche multimodale Gestalten in ihren Funktionsunterschieden zu ermitteln. So konnten beispielsweise zwei multimodale Gestalten des schwachen Zustimmens rekonstruiert werden, in denen die Interagierenden folgende Ressourcen zum Einsatz bringen: ein schwach affirmierendes sprachliches Element; eine Handgeste, die aus der Extension der Finger besteht; eine Kopfgeste, die eine seitliche Bewegung des Kopfes beinhaltet; sowie ein vom Gegenüber abgewandter Blick (vgl. Kap. 8.3.2 & 8.3.3). Erst die detaillierte Beschreibung der Bewegungsrichtung der Ressourcen (hier: der Kopfgeste) sowie die sequenzielle Einbettung der Gesamtgestalt ermöglichen Einsichten in ihre Bedeutungsunterschiede. Denn während die jeweils laminieren Ressourcen in der einen multimodalen Gestalt (vgl. Kap. 8.3.2) durch eine vom Gegenüber abgewandte Kopfgeste bereits einen *Widerspruch* ankündigen, der anschließend eingelöst wird, wird die Kopfgeste der anderen Gestalt affirmierend, d. h. als schwaches Nicken bzw. affirmierendes Kopfschütteln, geäußert, dem sequenziell ein verbales Beipflichten und/oder ein Themenwechsel folgt.

Auch die zeitliche Entfaltung der Ressourcen in der Gesamtgestalt hat Auswirkungen auf ihre Bedeutung und Funktion. Die Ressourcen können *simultan* oder *nacheinander* hervorgebracht und unterschiedlich lang *gehalten* werden. Der Aspekt der Zeitlichkeit erweist sich insofern als bedeutsam für die Funktion der multimodalen Gestalt, als dass zum einen erst über die emergente Entfaltung und Koordinierung der eingesetzten Ressourcen eine Gesamtgestalt entstehen kann. Zum anderen hat die Dauer der Gestalt auch interaktive Konsequenzen für den weiteren Verlauf der Argumentation, wie nachfolgend weiter erläutert wird. Multimodale Ressourcen können *simultan* miteinander koordiniert werden. Das bedeutet, dass die Ressourcen in der multimodalen Gestalt in einem multimodalen Verdichtungsraum zusammenlaufen. Die Ressourcen kulminieren dabei an dem Punkt mit dem höchsten informativen Gehalt (vgl. Stukenbrock 2015: 32, 2018: 40). Diese Simultaneität der Ressourcen erweist sich beispielsweise für die multimodale Gestalt des kurzzeitigen Präsentierens und Hervorhebens relevanter Bedeutungselemente als besonders wichtig, denn die Funktion der multimodalen Gestalt kann nur identifiziert werden, wenn das entsprechende Lexem zeitgleich mit der *Palm Up* Geste (und der prosodischen Markierung des Fokusakzents) geäußert wird (vgl. Kap. 8.2.1.2). Diese multimodalen Gestalten werden aufgrund der Simultaneität ihrer zeitlichen Entfaltung in einem Verdichtungsraum als lokal operierende multimodale Gestalten verstanden. Multimodale Ressourcen können sich auch *nacheinander* und z. T. in deutlicher temporaler Distanz zueinander entfalten, wie z. B. für die multimodale Kohäsions- und Kohärenzmarkierungen gezeigt werden konnte (vgl. Kap. 8.2.4). Entsprechend der internen Logik der multimodalen Gestalt, die in der sich sequenziell entfaltenden sprachlichen Äußerung einzelne Elemente kohäsiv miteinander verbindet bzw. als zusammengehörig markiert, entfalten sich die multimodalen Ressourcen nacheinander. Die multimodale Gestalt erstreckt sich über eine größere Reichweite. Sie wird daher als eine global operierende multimodale Gestalt verstanden. Die zusammengehörigen Elemente der multimodalen Gestalt sind durch ihre emergente Entfaltung sowie die Orientierung der Ressourcen aufeinander und ihre Koordinierung untereinander erkenn- und identifizierbar (vgl. Mondada 2014b, 2018). Multimodale Ressourcen können innerhalb einer multimodalen Gestalt auch *gehalten* werden (vgl. u. a. Streeck 2007). Das bedeutet, dass eine bereits entfaltete multimodale Gestalt, z. B. eine Zählgeste, anhand derer Interagierende Kriterien für die gemeinsame Entscheidungsfindung etabliert haben, über einen längeren Zeitraum und auch über Sprecher*innenwechsel hinweg im gemeinsamen Wahrnehmungsraum der Interagierenden präsent gehalten werden (vgl. Kap. 8.2.1.1). Ihnen kommt somit für die Dauer des Haltens eine relevante interaktive bzw. interaktionsorganisatorische Funktion zu. I. d. S. weisen auch diese multimodalen Gestalten eine globale Reichweite auf.

Die Bedeutungen und Funktionen multimodaler Gestalten, die aus „multiple layers of multimodal resources“ (Mondada 2016: 351) bestehen, sind also geprägt durch die beteiligten Ressourcen, ihre zeitliche Entfaltung sowie die sequenzielle Einbettung der Gesamtgestalt.

*I.c.) Funktional für das mündliche Argumentieren eingesetzte multimodale Gestalten sind nicht per se an die Rolle der ‚Sprecher*innen‘ gebunden.*

Unabhängig von ihrer Sprecher*innenrolle setzen Interagierende in mündlichen Argumentationen multimodale Ressourcen auch zur Rederechtsorganisation, zur Lenkung des Gesprächsthemas und des Turnformats sowie zur Verstehensdokumentation ein.

In mündlichen Argumentationen, in denen die Konkurrenz um das Rederecht relevant werden kann, um z. B. eine Begründung an möglichst geeigneter Stelle in der Interaktion zu platzieren, erweist sich das Präsenthalten eines etablierten Gesprächsgegenstandes als äußerst funktional, um einen begonnenen Beitrag auch nach einer Unterbrechung noch fortsetzen zu können. Dabei bestätigen die Befunde z. T. bestehende Forschungsergebnisse, wie beispielsweise, dass durch das Halten deiktischer Gesten auf das Material Gesprächsgegenstände in Unterbrechungen präsent gehalten werden und zugleich die Fortsetzung des *Turns* angezeigt wird (vgl. Mondada 2007). Doch auch andere Gestenformen, wie z. B. Zählgesten (vgl. Kap. 8.2.2.2), können in dieser Funktion von den Interagierenden eingesetzt werden.

Die Steuerung der Gesprächsthemen sowie des Turnformats erfolgt auf unterschiedlichen Wegen: Interagierende halten beispielsweise etablierte Gesprächsgegenstände über ihren eigenen *Turn* hinaus präsent und setzen sie somit auch relevant für den Gesprächsbeitrag ihres Gegenübers. Interagierende können beispielsweise Kriterien für die Entscheidungsfindung anhand von Zählgesten etablieren, diese halten und somit letztlich eine konditionelle Relevanz etablieren, sodass das Gegenüber nach dem Sprecher*innenwechsel die präsent gehaltenen Kriterien in seinem eigenen Beitrag aufgreift und weiterbearbeitet. Während des Einforderns einer Rückmeldung kann ein Gesprächsgegenstand aber auch durch deiktische Gesten auf das Material etabliert und für die Rückmeldung des Gegenübers als relevante Referenz präsent gehalten werden (vgl. Kap. 8.2.2.2). Gesprächsthemenunabhängig können Interagierende aber auch lenkend in das Turnformat einer eingeforderten Rückmeldung hineinwirken, indem beispielsweise durch Zählgesten markiert wird, dass die elizitierte Rückmeldung des Gegenübers als Aufzählung einzubringen ist (vgl. Kap. 8.5.1). Dabei werden diese Formen der Gesprächssteuerung i. d. R. in den untersuchten Eltern-Kind-Argumentationen elternseitig genutzt (vgl. auch Heller 2012; Quasthoff und Kluger 2021). Es kann angenommen werden, dass sich hierin unabhängig von der Kompetenz des Kindes eine Asymmetrie der Sprecher*innen in ihren familialen Rollen niederschlägt.

Zuletzt können multimodale Ressourcen rezipient*innenseitig auch der Verstehensdokumentation dienen (vgl. Kap. 8.2.2.3). Durch deiktische Gesten auf konkrete Optionen auf dem Material können Rezipient*innen sichtbar machen, welche sprecher*innenseitig benannten Optionen sie in dem Moment nachvollziehen. Es wird deutlich, dass die Rezipient*innen dadurch keinen Anspruch auf das Rederecht erheben – vielmehr zeigen sie durch den interaktiven Nachvollzug auch ihre Involviertheit gegenüber der sich aktuell im Aufbau befindlichen Begründung des*der Sprecher*in an. Die multimodale Gestalt ermöglicht zugleich dem*der aktuellen Sprecher*in, zu überprüfen, ob die Rezipient*innen auf die richtigen Optionen hin orientiert sind, und sie ggf. zu korrigieren (vgl. Kap. 8.5.2).

Im Rahmen der Studie wurde ersichtlich, dass auch unabhängig von der Rolle der ‚Sprecher*innen‘ bzw. ‚Rezipient*innen‘ multimodale Gestalten eingesetzt werden, die wichtige Funktionen für die mündliche Argumentation übernehmen. Diese Beobachtungen bekräftigen, dass auch Interagierende, die gerade nicht als Sprecher*innen in Erscheinung treten, an dem Prozess der Interaktion beteiligt sind

– und dass die von ihnen eingesetzten multimodalen Ressourcen Konsequenzen für den Aufbau der argumentativen Handlung haben.

Aufbauend auf Forschungserkenntnissen des ersten empirischen Kapitels 8 wurden in dem zweiten empirischen Teil der Arbeit (vgl. Kap. 10) erwerbsbezogene Fragen adressiert. Diese Befunde werden nachfolgend vorgestellt.

Forschungsfrage II: Welche Veränderungen zeigen sich in einer längsschnittlich angelegten Studie zweier Fokuskinder im Alter von 12, 13 und 15 Jahren hinsichtlich der sprachlich-diskursiven Fähigkeiten sowie des Einsatzes multimodaler Ressourcen?

II.a.) Heranwachsende nutzen multimodale Ressourcen i. S. komplexer multimodaler Gestalten funktional für das Argumentieren und bearbeiten durch diese die diskursiven Anforderungen des mündlichen Argumentierens.

Anhand der explorativen Fallstudien konnte nachgewiesen werden, dass Heranwachsende in unterschiedlichem Maße multimodale Ressourcen zu multimodalen Gestalten koordinieren, die dann zur Realisierung argumentativer Handlungen beitragen. Dabei verwenden die Heranwachsenden multimodale Markierungsformen i. S. der Gestalten zur Realisierung argumentativer Handlungen (hier: dem Begründen). Wie anhand der Fallstudie gezeigt werden konnte, werden die multimodalen Ressourcen allerdings zu unterschiedlichen Zeitpunkten sowie in Abhängigkeit von dem diskursstrukturellen Profil der Sprecher*innen eingesetzt. Zur Bearbeitung der diskursiven Anforderungen der argumentativen Entscheidungsfindung verwenden die Heranwachsenden v. a. während des ‚Erprobens des Repertoires‘ ein besonders breites, reichhaltiges Repertoire an multimodalen Gestalten zur Markierung der Kontextualisierung, Vertextung sowie epistemischen Modalisierung in ihren Begründungen (vgl. Kap. 10.4). Sie setzen beispielsweise in der Vertextung global strukturierter Äußerungspakete Gestenwiederaufnahmen (z. T. mit lexikalischen und morpho-syntaktischen Wiederaufnahmen) als kohäsionsstiftende Markierungen ein; sie nutzen zur Kontextualisierung ihrer Begründung *Dropping of Hand* oder *Throwing Away* Gesten, um das Verwerfen von Gesprächsgegenständen sichtbar anzuzeigen; und sie markieren u. a. durch das Hervorheben der Modalpartikel *ja* Begründungsinhalte als evident. Auch wenn die Untersuchung der Daten zeigt, dass und wie vielfältig die Heranwachsenden multimodale Gestalten zur Realisierung der argumentativen Handlung des Begründens einsetzen, werden auch erwerbsbezogene Unterschiede sichtbar. Diese liegen in der *Konfiguration* der multimodalen Gestalten selbst und ergeben sich entlang von drei Merkmalen: Multimodale Ressourcen im Rahmen einer argumentativen Handlung zu einer multimodalen Gestalt zu verbinden, erfordert eine präzise zeitliche *Koordinierung* der Ressourcen, wie in Kapitel 8 rekonstruiert werden konnte. Gelingt diese Koordinierung nicht, kann dies Auswirkungen auf die Identifizierbarkeit der multimodalen Gestalt und somit auf ihre Funktion innerhalb der sich im Aufbau befindlichen argumentativen Handlung haben. Neben der Koordinierung legen die Fallstudien zwei weitere Merkmale multimodaler Konfigurationen frei, die in der systematischen Untersuchung (vgl. u. a. Kap. 8.2.2 & Kap. 8.2.4) nur gestreift werden konnten, sich aber als relevant für den multimodalen Diskurserwerb erweisen. Die argumentativen Handlungen sowie die an ihrer Realisierung beteiligten multimodalen Gestalten unterscheiden sich hinsichtlich ihrer *Reichweite*, d. h. ihrer Lokalität bzw. Globalität, sowie ihrer internen *Komplexität*. Diese Merkmale, d. h. die Koordinierung der ausdrucksseitigen, semiotischen Ressourcen sowie die Reichweite und Komplexität der multimodalen Konfigurationen, erweisen sich als fruchtbare Beschreibungskategorien der Erwerbsanforderungen, mit denen Heranwachsende im multimodalen Diskurserwerb umgehen müssen.

II.b.) In der Reichweite, Koordinierung sowie der Komplexität multimodaler Gestalten liegen spezifische Erwerbsanforderungen des multimodalen argumentativen Diskurserwerbs.

Die Anforderungen, die sich an Interagierende in mündlichen Argumentationen im Gespräch ergeben, betreffen grundsätzlich Planungs-, Produktions- und Rezeptionsprozesse (vgl. u. a. Mundwiler et al. 2017: 93). Gemeinsam arbeiten die Beteiligten an der Hervorbringung der argumentativen Gespräche und zwar im Verständnis der vorliegenden Arbeit unter Berücksichtigung unterschiedlicher Ausdrucksressourcen, die sich sowohl nacheinander (*sprachliche Ressource*) als auch simultan (*körperlich-sichtbare Ressourcen*) entfalten können. Aus der Erwerbsstudie ableitend zeigten sich spezifische multimodale Erwerbsanforderungen, die in der Reichweite und Komplexität der multimodalen Gestalten in der Realisierung argumentativer Handlungen zu liegen scheinen und darüber hinaus die Koordinierung der multimodalen Ressourcen betreffen. Dabei wurde sichtbar, dass die *Koordinierungsanforderungen* mit der Zunahme der Reichweite sowie der Komplexität der multimodalen Gestalten ansteigen. Denn in *lokal* operierenden multimodalen Gestalten, bei denen z. B. mittels der *Palm Up* Geste ein Lexem auf der geöffneten Handfläche präsentiert wird, können die Ressourcen eher miteinander funktional koordiniert werden als in komplexen, *global* operierenden multimodalen Gestalten wie dem Markieren von Kohäsion und Kohärenz. Denn für das Markieren von Kohäsion und Kohärenz müssen innerhalb des Produktionsprozesses der argumentativen Handlung kohäsionsstiftende Elemente der Äußerung mit körperlich-sichtbaren Ressourcen, z. B. in Form von *mehreren* aufeinanderfolgenden Gesten, koordiniert werden. Das kurzzeitige Präsentieren und Hervorheben erfordert hingegen, ein *einzelnes* Lexem durch eine gestische Ressource prominent zu setzen. Es scheint für Heranwachsende also anforderungsreicher zu sein, global operierende, komplexe multimodale Gestalten, die keine ‚feste‘ temporale Alig-nierung der Ressourcen aufweisen, miteinander zu koordinieren als lokal operierende multimodale Gestalten.

Vergleicht man nun diese induktiv-rekonstruktiv ermittelten Erwerbsanforderungen mit bisherigen Erwerbsstudien zum multimodalen Diskurserwerb, zeigt sich, dass die dort beschriebene Zunahme an Komplexität, die u. a. über die Anzahl der verwendeten Sätze zur Umsetzung der Diskurseinheit, über die Häufigkeit der verwendeten Konnektoren sowie über die Anzahl an Gesten ermittelt worden ist (vgl. u. a. Colletta und Pellenq 2009: 70f.; Reig Alamillo et al. 2012), deutlich spezifiziert werden konnte: Denn die Komplexität wurde in der vorliegenden Studie sowohl an die explizit-sprachliche Vertex-tung, d. h. den Aufbau und die Struktur der argumentativen Handlung, als auch an die dabei eingesetzten körperlich-sichtbaren Ressourcen i. S. multimodaler Gestalten zurückgebunden. Dies ermöglichte zu ermitteln, inwiefern diese in einem Zusammenhang zueinanderstehen (vgl. Kap. 10.4). Außerdem konnten mit der *Reichweite* und *Koordinierung* zwei weitere Erwerbsanforderungen dargelegt werden, die analytisch getrennt voneinander herauspräpariert worden sind, allerdings Verbindungslinien zueinander aufweisen (vgl. dazu die Ausführung zur Koordinierung lokal operierender multimodaler Gestalten).

II.c.) Mit zunehmenden diskursiven Kompetenzen kann die Komplexität und Reichweite argumentativer Handlungen sowie die zeitliche Koordinierung der Ressourcen ausgebaut werden. Können dabei entstehende Anforderungen nicht bewältigt werden, kommt es in der weiteren Entwicklung zu einer Reduktion der Komplexität und Reichweite.

Die zuvor beschriebenen Erwerbsanforderungen der Reichweite und Komplexität (vgl. II.b.)), die in der vorliegenden Studie herausgearbeitet werden konnten, spiegeln Aspekte der Modellierung von Diskurskompetenz wider, wie sie von Quasthoff (2009) dargelegt worden sind. Denn das Diskurskompetenzmodell erfasst auf der Ebene der Vertex-tung zum einen die Reichweite, d. h. die Strukturierung lokaler bzw. globaler Äußerungen (vgl. Quasthoff et al. 2019: 215). Dabei zeigt sich in der Entwicklungsst-

denz sprachlich-diskursiver Kompetenzen, dass heranwachsende Sprecher*innen zuerst lokale, eher reichende Beiträge mit zunehmendem Diskurserwerb zu global strukturierten, logisch-hierarchisierenden Äußerungspaketen ausbauen (vgl. zsf. Kap. 4.1.2). Über die Beschreibung der internen Struktur (reihend vs. logisch-hierarchisch) können zum anderen Einblicke in die Komplexität der Äußerungsstrukturierung gewonnen werden. Hinsichtlich des Nachvollzugs *multimodaler* Diskurserwerbstendenzen tritt als dritte Anforderung die *Koordinierung* hinzu (vgl. Kap. 10.4 & 10.5).

In der explorativen Fallstudie wurden nun in der Untersuchung des multimodalen Diskurserwerbs die rekonstruierten sprachlich-diskursiven und multimodalen Entwicklungen der Fallstudien in einen Zusammenhang (vgl. Kap. 10.4) gesetzt. Dabei zeigte sich, dass die oben beschriebene Entwicklungstendenz für das diskursstrukturell stärkere *Fokuskind I* auch im Rahmen des multimodalen Diskurserwerbs bestätigt werden kann: Denn das diskursstrukturell stärkere Fokuskind I baut seine Fähigkeiten, die es im Alter von 12 Jahren zeigt, bis zum Alter von 15 Jahren weiter aus. Das bedeutet, dass die argumentative Handlung des Begründens, auf der der Fokus der Fallstudien lag, sowohl in ihrer Struktur als auch in den eingesetzten multimodalen Gestalten komplexer wird und die Begründungen an Globalität, d. h. an Reichweite gewinnen (vgl. dazu Kap. 10.3.1.2, Argumentaufzählung mit 15 Jahren). Mit zunehmendem Diskurserwerb gelingt der Sprecherin dabei auch die Koordinierung der multimodalen Ressourcen in komplexen, global strukturierten Begründungen.

Im Hinblick auf das diskursstrukturell schwächere *Fokuskind II* zeigt sich allerdings eine andere Entwicklungstendenz: Ein Anstieg der Reichweite und Komplexität der Begründungen ist bei Fokuskind II nur im Alter von 12 bis 13 Jahren festzustellen. Produziert es im Alter von 12 Jahren noch satzförmige *Brauchen*-Argumente ohne den Einsatz multimodaler Gestalten, baut es im Alter von 13 Jahren selbstständig Szenarios und Belegerzählungen auf – die Begründungen gewinnen an Globalität und Komplexität. Die Sprecherin erprobt dabei nicht nur sprachlich-diskursive, sondern auch multimodale Fähigkeiten und setzt lokal und global operierende multimodale Gestalten ein. Allerdings wird ersichtlich, dass in global-strukturierten Äußerungen die Koordinierung der eingesetzten Ressourcen nicht gelingt – auch die lokal operierenden multimodalen Gestalten, wie das präsentierende Hervorheben eines Lexems, bereiten der Sprecherin in diesem sequenziellen Kontext Schwierigkeiten. Während nun das diskursstrukturell stärkere *Fokuskind I* seine Fähigkeiten bis zum Alter von 15 Jahren weiter ausbaut, *reduziert* das diskursstrukturell schwächere *Fokuskind II* die Reichweite und Komplexität der multimodal realisierten Begründungen. Es zeigt also eine Entwicklung entgegen der vermuteten Erwerbstendenzen, die einen weiteren Ausbau der Diskurskompetenzen erwartbar machen würden. Doch es wird an dieser Stelle noch ein anderer Befund der explorativen Fallstudie sichtbar: Mit der Reduktion der Reichweite und Komplexität gelingt Fokuskind II im Alter von 15 Jahren die Koordinierung der Ressourcen. Daraus lässt sich vorsichtig ableiten: Wenn heranwachsenden Sprecher*innen während des ‚Erprobens‘ der erworbenen Repertoires an sprachlichen und multimodalen Ressourcen Schwierigkeiten in der Vertextung global strukturierter Begründungen sowie in der Koordinierung der dazu eingesetzten Ressourcen zeigen, kann die Globalität zugunsten einer besseren Koordinierung zurückgebaut werden.

Aus den hier zusammenfassend vorgestellten zentralen Befunden der Studie ergeben sich nun Forschungsdesiderate und -perspektiven, die im folgenden Kapitel 12 aufgezeigt werden.

12 Ausblick

Das mündliche Argumentieren gilt als wichtige Diskurspraktik, um als Mitglied einer Gesellschaft divergente Geltungsansprüche oder Wissensbestände sowie Probleme aus- bzw. verhandeln zu können (vgl. Quasthoff et al. 2021a; vgl. Kap. 2). Im Lichte der dargelegten Befunde wird die Relevanz multimodaler Ressourcen in argumentativen Aushandlungsprozessen offensichtlich. Aufbauend auf der vorliegenden Studie werden im folgenden Kapitel 12 aus der Untersuchung hervorgehende, weiterführende Desiderate benannt, die sich als fruchtbar für zukünftige Forschungsvorhaben in den Bereichen des mündlichen Argumentierens und multimodalen Diskurserwerbs erweisen. Angesichts der erwerbsbezogenen Befunde, die diese Studie zutage gefördert hat, schließt die Arbeit sodann mit sprachdidaktischen Implikationen und aus diesen hervorgehenden Forschungsdesideraten.

Desiderate zu Funktionen multimodaler Ressourcen für das mündliche Argumentieren

Dass multimodale Ressourcen als relevante Ausdrucksressourcen argumentative Handlungen prägen, konnte im Rahmen der vorliegenden Studie auf Grundlage eines Datenkorpus von Eltern-Kind-Interaktionen gezeigt werden. Bedingt durch die Datengrundlage sind die Befunde durch spezifische kontextuelle Voraussetzungen geprägt: Die Interagierenden befinden sich in sitzender, nicht-dynamischer Position am Tisch und zwischen ihnen liegt das Material der Entscheidungsaufgabe (vgl. Kap. 6), auf das sie sich in der Interaktion hin orientieren können. Die Interaktion weist somit spezifische *ökologische Bedingungen* auf (vgl. Mondada 2014b), die auch in den rekonstruierten multimodalen Gestalten, wie dem Etablieren und Präsenthalten eines Gesprächsgegenstandes, sichtbar werden. Wenn wir mit Depermann (2003: 23) berücksichtigen, dass aufgrund der Vielgestaltigkeit des Argumentierens unterschiedliche Gattungen des Argumentierens existieren, nimmt auch die Spezifik der *argumentativen Entscheidungsfindung* Einfluss auf die eingesetzten multimodalen Gestalten. So liegen der Entscheidungsfindung vier Auswahloptionen zugrunde, aus denen eine von den Interagierenden gemeinsam begründet ausgewählt werden soll. Da die Optionen den Beteiligten außerdem in Form des Arbeitsblattes (vgl. oben, ökologische Bedingungen) materiell zur Verfügung stehen, verwundert es nicht, dass die Interagierenden gestische Bezüge zu den Auswahlmöglichkeiten herstellen und diese für entsprechende argumentative Handlungen funktional einsetzen, wie beispielsweise für das nachdrückliche Zustimmung (vgl. Kap. 8.3.1) gezeigt werden konnte. Nicht zuletzt nehmen auch die *familialen Rollen* der Interagierenden Einfluss auf die verwendeten multimodalen Gestalten. Denn in den argumentativen Entscheidungsfindungen zeigt sich, dass die Eltern steuernde Funktionen in der Interaktion übernehmen können (vgl. auch Quasthoff und Kluger 2021; Heller 2012). In den vorliegenden Daten setzen die Elternteile u. a. multimodale Ressourcen zur Steuerung von Gesprächsthemen sowie der Turnfortsetzung (vgl. Kap. 8; vgl. zsf. Kap. 11) ein. Wenn die vorliegende Studie nun danach fragt, welche Funktionen multimodale Ressourcen für das mündliche Argumentieren übernehmen, dann sind diese kontextuellen Faktoren in der Beantwortung der Frage zu berücksichtigen. Aus diesem Grund erscheint es gewinnbringend für zukünftige Forschung, auch weitere Kontexte im Hinblick auf die Funktion multimodaler Ressourcen für das Argumentieren zu untersuchen, in denen z. B. die Rollen der Interagierenden symmetrisch(er) verteilt sind und andere Gattungen des Argumentierens in den Blick genommen werden, wie z. B. die Dissensbearbeitung unter Peers in Cliquengesprächen. Dabei können Forschungsfragen adressiert werden, die auf potenzielle Unterschiede multimodaler Ausdrucksressourcen fokussieren, die mit der kontextuellen Rahmung der Interaktion einhergehen können: Bearbeiten beispielsweise Heranwachsende Dissense stärker *konfrontativ* (vgl. Morek 2015) und werden in den dann eingesetzten opponierenden Handlungen weitere bzw. andere multimodale Gestalten funktional eingesetzt als solche, die in der vorliegenden Studie rekonstruiert werden konnten? Führt eine stärkere Konfrontation möglicherweise auch

zu einem erhöhten Einsatz sowie einer qualitativen Veränderung, d. h. einer erhöhten Intensität in der Ausführung der multimodalen Ressourcen?

Aus diesem v. a. auf kontextuelle Bedingungen der Argumentation fokussierenden Forschungsdesiderat resultieren auch forschungsmethodische Implikationen zur Untersuchung multimodaler Ressourcen beim Argumentieren.

Forschungsmethodische Implikation zur Untersuchung multimodaler Ressourcen beim Argumentieren

Die Konzeptualisierung multimodaler Gestalten ist von Mondada (2014b) aus einem multimodalen conversationsanalytischen Zugriff entwickelt worden (vgl. Kap. 3.1.4). Aufgrund dieser Prägung scheint es nachvollziehbar, dass die Studien, die mit der Konzeptualisierung multimodaler Gestalten arbeiten, diese v. a. hinsichtlich genuin gesprächsanalytischer Fragestellungen untersuchen, z. B. im Rahmen der Rederechtsorganisation (vgl. Mondada 2007), des Abschließens von Redebeiträgen und *Actions* (vgl. Mondada 2006, 2015a; Stefani und Mondada 2021) oder auch in Frage-Antwort-Paarsequenzen (vgl. Kärkkäinen und Thompson 2018). Konkrete Bezüge zu *Gattungen* oder *komplexeren Diskurspraktiken* werden nicht hergestellt. Die vorliegende Arbeit bietet einen ersten Ansatz, mit dem die Konzeptualisierung multimodaler Gestalten fruchtbar gemacht werden kann, um auch Untersuchungen größerer Diskurseinheiten vorzunehmen: Multimodale Gestalten werden darin als Verbünde von Ausdrucksressourcen verstanden, mit denen *gattungsspezifische* Handlungen realisiert werden, die sich aus pragmatischen Mitteln zusammensetzen, wie sie in GLOBE (vgl. Kap. 2.2.2) beschrieben werden, und zur Bearbeitung der kommunikativen Aufgabe beitragen. Zukünftige Arbeiten können diesen Zugriff nutzen, um in anderen kommunikativen Gattungen, wie beispielsweise dem Erklären oder Erzählen, für die das GLOBE-Modell ebenfalls als Beschreibungsinstrument vorliegt (vgl. Hausendorf und Quasthoff 1996/2005; Morek 2012; Ohlhus 2014), zu untersuchen, wie Interagierende diese *multimodal* bearbeiten. Denn bisherige *gesprächsanalytische* Studien fokussieren in diesen Gattungen vornehmlich auf die sprachliche Ebene, während Untersuchungen zur *multimodalen* Realisierung v. a. Gesten in abbildender Funktion in den Blick nehmen und die Funktion der körperlich-sichtbaren, meist gestischen Ressource unabhängig von ihrem sequenziellen Kontext ermitteln (vgl. zsf. Kap. 4; vgl. aber Ohlhus 2016; Heller 2018b). Der in dieser Arbeit vorgeschlagene Ansatz könnte bisherige Untersuchungen dementsprechend sinnvoll ergänzen.

Öffnet man an dieser Stelle die Blende etwas weiter – in Vorausschau auf die gleich anschließenden Desiderate zum multimodalen Diskurserwerb –, scheint eine neu perspektivierte Untersuchung kommunikativer Gattungen auch deshalb sinnvoll, da jüngst erwerbsbezogene Befunde zutage fördern, dass Kinder innerhalb ihres explanativen Diskurserwerbs nicht zuerst mit illustrierenden, veranschaulichenden Verfahren ihre Erklärungen durchführen, sondern mit *abstrahierenden*, d. h. vom konkreten Gegenstand abgelösten Verfahren (vgl. Kinalzik 2024). Dieser Befund kontrastiert mit bisherigen Forschungsergebnissen zum multimodalen Erklärerwerb, die in der frühen Entwicklung explanativer Fähigkeiten beobachten, dass die jungen Sprecher*innen vornehmlich Gesten mit abbildender Funktion einsetzen. Mit zunehmendem Diskurserwerb geht die Verwendung dieser Gesten dann zurück und die Heranwachsende bringen vermehrt Gesten mit pragmatischer Funktion in ihre Erklärungen ein (vgl. u. a. Colletta und Pellenq 2009; Reig Alamillo et al. 2012; vgl. Kap. 4.2.1). In den genannten Studien zum multimodalen Diskurserwerb explanativer Fähigkeiten sind i. d. R. keine qualitativen gesprächsanalytischen Untersuchungen durchgeführt worden, die die Funktionen multimodaler Ressourcen in ihrem sequenziellen Kontext rekonstruieren (vgl. Kap. 4). Der in der vorliegenden Arbeit vorgeschlagene forschungsmethodische Zugriff kann in zukünftiger Erwerbsforschung genutzt werden, um der Frage nachzugehen, ob mit dem früheren Erwerb abstrahierender Vertextungsverfahren auch ein früher Erwerb multimodaler

Gestalten mit pragmatischer Funktion einhergeht. Hieran anschließend sollen im Folgenden nun *Desiderate* zum multimodalen Diskurserwerb benannt werden.

Desiderate zum multimodalen Diskurserwerb

Der Erwerb argumentativer Fähigkeiten beginnt, so zeigen Erwerbsstudien, bereits früh in der kindlichen Entwicklung, sodass das multimodale Markieren eines Dissenses und das Einbringen von Begründungen bereits im Alter von 2 Jahren zu beobachten ist (vgl. u. a. Völzing 1982; Arendt 2019; vgl. ausführlich Kap. 4). Im Alter von 4 Jahren – frühere Studien hierzu liegen m. W. nicht vor – verwenden Kinder dann gestische Ressourcen in pragmatischer Funktion in der kommunikativen Gattung des (Nach-)Erzählens (vgl. u. a. Graziano 2014b) und zwischen 4 und 6 Jahren in argumentativen Interaktionen (vgl. Ladewig und Hotze 2023). Die vorliegende erwerbsbezogene Fallstudie adressiert nun Heranwachsende im Alter zwischen 12 und 15 Jahren und zeigt, dass diese multimodale Ressourcen in Form multimodaler Gestalten funktional für das mündliche Argumentieren einsetzen. Anhand der multimodalen Gestalten können kindseitig eingebrachte Beiträge *kontextualisiert* werden, indem z. B. Gesprächsgegenstände durch eine Geste der Familie der *Away* Gesten verworfen werden. Äußerungspakete können durch multimodale Ressourcen, wie z. B. durch Gestenwiederaufnahmen, *strukturiert* werden. Heranwachsende können außerdem ihre Beiträge *epistemisch modalisieren*, z. B. indem sie durch beidhändige *Palm Up* Gesten mit lateraler Bewegung und im Verbund mit weiteren *Shrug*-Komponenten Unsicherheit anzeigen (vgl. Kap. 10). Die verwendeten multimodalen Gestalten weisen nun Unterschiede hinsichtlich ihrer Konfigurationen in Abhängigkeit vom diskursstrukturellen Profil des Kindes auf: Das diskursstrukturell stärkere Fokuskind I kann mit zunehmenden Diskurserwerb die multimodalen Gestalten in ihrer Reichweite und Komplexität weiter ausbauen. Die dabei eingesetzten multimodalen Ressourcen erweisen sich in einem hohen Grad fein temporal aufeinander abgestimmt. Das diskursstrukturell schwächere Fokuskind II erprobt im Alter von 13 Jahren den Aufbau größerer Äußerungspakete sowie den Einsatz multimodaler Gestalten, die an Komplexität und Reichweite gewinnen; allerdings treten dabei Koordinierungsschwierigkeiten auf. Mit der Reduktion der Globalität der eingebrachten Beiträge im Alter von 15 Jahren gelingt Fokuskind II dann die Koordinierung der Ressourcen – allerdings ausschließlich in lokal strukturierten Äußerungen. Auf Grundlage dieser Befunde kann angenommen werden, dass mit fortschreitendem Ausbau diskursiver Fähigkeiten, die bei Fokuskind I bis ins Alter von 15 Jahren und bei Fokuskind II bis zum Alter von 13 Jahren zu beobachten sind, Heranwachsende nicht nur ihre Äußerungspakete global(er) strukturiert aufbauen, sondern auch die Reichweite und Komplexität der dabei eingesetzten multimodalen Gestalten zunehmen. Es ist v. a. mit Blick auf die Entwicklung des zweiten Fokuskindes außerdem zu vermuten, dass multimodale Ressourcen in lokal operierenden multimodalen Gestalten leichter zu koordinieren sind als in global operierenden multimodalen Gestalten.

Es stellt sich nun die Frage, wie sich der multimodale Diskurserwerb argumentativer Fähigkeiten vor dem in der vorliegenden Studie anvisierten Alter von 12 bis 15 Jahren vollzieht – denn Fokuskind I setzt beispielsweise bereits mit 12 Jahren ein reichhaltiges Repertoire multimodaler Gestalten ein. Um nachvollziehbar zu machen, wie Heranwachsende lernen, multimodale Ressourcen in argumentativen Handlungen zu multimodalen Gestalten zu konfigurieren und in diesen zu koordinieren, erweist es sich als notwendig, dass zukünftige Forschung Daten jüngerer Kinder berücksichtigt. Dabei könnten die Zusammenhänge der sprachlich-diskursiven und multimodalen Entwicklungslinien weiter beleuchtet und folgende Fragen in den Blick genommen werden: Ist es denkbar, dass Kinder, die z. B. ihre Beiträge sprachlich noch nicht in den sequenziellen Kontext einpassen können (vgl. Morek 2018) oder denen erst wenige explizit-sprachliche Mittel zur Strukturierung eines Äußerungspakets zur Verfügung stehen (vgl.

Ohlhus 2016), auf gestische Ressourcen zurückgreifen, um eigene Beiträge zu platzieren bzw. strukturiert aufzubauen? Und inwiefern können körperlich-sichtbare Ressourcen somit als eine *situative* Unterstützung fungieren, um aktuelle Anforderungen zu bewältigen? Oder handelt es sich nicht um situative Unterstützungsverfahren, sondern läuft die Fähigkeit, körperlich-sichtbare Ressourcen funktional für die Kontextualisierung sowie auch Vertextung einzusetzen, den sprachlich-diskursiven Fähigkeiten möglicherweise voraus? Es gilt ferner zu klären, ob sich auch in weiterer Erwerbsforschung beobachten lässt, dass, wenn Kindern der Ausbau der Reichweite und Komplexität ihrer multimodalen Äußerungspakete nicht gelingt (wie bei Fokuskind II), diese zugunsten der Koordinierung reduziert werden. Zuletzt gilt in zukünftiger Forschung zu klären, ob die in der vorliegenden Studie formulierte hypothetische Erwerbslogik, dass lokal operierende multimodale Gestalten in lokal operierenden argumentativen Handlungen leichter zu realisieren sind und somit möglicherweise zuerst erworben werden, unter Berücksichtigung jüngerer Alterskohorten bestätigen lässt.

Da das mündliche Argumentieren über den familialen Kontext hinaus u. a. auch in der Schule eine wichtige Diskurspraktik darstellt, die es Heranwachsenden ermöglicht, an unterrichtlichen Bildungs- und Lernprozessen aktiv teilzuhaben (vgl. u. a. Heller 2012; Morek und Heller 2012; Quasthoff et al. 2021a), wird nachfolgend eine sprachdidaktische Perspektive eröffnet.

Sprachdidaktische Implikationen und Desiderate

Schüler*innen verfügen, wie die Heranwachsenden der vorliegenden Studie, über unterschiedlich ausgebaute Repertoires an sprachlich-diskursiven Fähigkeiten, die es ihnen ermöglichen, mit den sprachlich-diskursiven Anforderungen des Unterrichts umzugehen (vgl. Morek und Heller 2021). Welche Ressourcen die Heranwachsenden zur Bewältigung der Anforderungen mitbringen und wie weit ihre Fähigkeiten ausgebaut sind, ist sowohl abhängig von den internen Ressourcen selbst, die in der vorliegenden Arbeit durch die Kontextualisierung, Vertextung und Markierung erfasst worden sind, als auch von den Erfahrungen, die sie in den unterschiedlichen Erwerbskontexten gewinnen (vgl. Kap. 4). Abhängig von dem Ausbau der individuellen Fähigkeiten der Heranwachsenden sowie abhängig von der interaktiven Unterstützung, die sie in der Bewältigung der Anforderung erfahren, können sie am Unterrichtsgeschehen sowie den darin stattfindenden Bildungsprozessen teilhaben (vgl. Quasthoff et al. 2021a). Doch mündliche Fähigkeiten sind nicht nur *Voraussetzung* zur Teilhabe, sondern bilden im schulischen Kontext auch einen eigenen Kompetenzbereich (Mündlichkeit), in dessen Rahmen die schüler*innenseitigen Fähigkeiten auch bewertet werden. Es stellt sich angesichts der vorliegenden Arbeit die Frage, inwieweit dieser Kompetenzbereich nun auch körperlich-sichtbare Ressourcen als zugehörig zur Mündlichkeit konzipiert. Es zeigt sich, dass die Beschlüsse der Kultusministerkonferenz (2004 i. d. F. von 2022) in der Primarstufe Aspekte der Körperlichkeit in den Kompetenzbereich integrieren, wie z. B. dass die Schüler*innen „ausgewählte Wirkungen ihrer Redeweise situations- und adressatengerecht [berücksichtigen]“ und dabei „Lautstärke, Betonung, Sprechtempo, Pausen, Körpersprache (Gestik/Mimik), Körperhaltung“ (Kultusministerkonferenz 2004 i. d. F. von 2022: 11) relevant werden. Wenngleich aber die Multimodalität im Kompetenzbereich der Mündlichkeit abgebildet ist, fehlt bislang noch, wie auch in den Desideraten der vorliegenden Studie (vgl. Kap. 4 & 5) deutlich wurde,

[e]ine systematische Beschreibung medial mündlicher Markierungsformen, in der ggf. ein kompetenzbezogenes Anwachsen eines Ressourceninventars vielfältiger sprachlicher, parasprachlicher und nonverbaler Mittel genutzt und eingesetzt werden kann [...]. (Grundler 2022: 581)

Dabei wird in der vorliegenden Arbeit deutlich, dass multimodale Ressourcen sowohl hinsichtlich der Markierung der Kontextualisierung und Vertexung als auch der epistemischen Modalisierung von den Sprecher*innen relevant gemacht werden. Multimodale Ressourcen bilden also wichtige Ausdrucksressourcen, die in der Bearbeitung kommunikativer Aufgaben von Heranwachsenden eingesetzt werden (vgl. Kap. 10). Ohlhus (2016: 61) fordert deshalb, körperlich-sichtbare Ressourcen in die Beurteilung von Diskurskompetenzen mit einzubeziehen und damit zu berücksichtigen, dass, wie er zeigt, auch körperlich-sichtbare Ressourcen beispielsweise zur Strukturierung größerer Diskurseinheiten beitragen können. Mit Blick auf die Befunde der vorliegenden erwerbsbezogenen Untersuchung, in der die Interagierenden u. a. auch performative multimodale Gestalten einsetzen, um eine Sprechhandlung erkennbar zu machen, die in der sprachlichen Realisierung noch implizit verbleibt (vgl. 10.4, Kontextualisierung), scheint eine solche Forderung sinnvoll und notwendig, wenn die Diskurskompetenz eines Kindes umfassend eingeschätzt werden soll. Aus dieser Implikation, auch multimodale Ressourcen in der Beurteilung schüler*innenseitiger argumentativer Fähigkeiten anzuerkennen, resultiert das Desiderat, in zukünftiger Forschung zu ermitteln, wie eine sinnvolle Berücksichtigung multimodaler Ressourcen in der Leistungsbeurteilung aussehen kann. Ein möglicher Zugang könnte über die Modellierung von Diskurskompetenz nach Quasthoff (2009) gelingen, wenn in diese multimodale Ausdrucksressourcen als Markierungsformen der Kontextualisierung und Vertexung systematisch einbezogen werden.

Betrachtet man nun hinsichtlich der in der vorliegenden Studie untersuchten Diskurspraktik des Argumentierens, wie die explizite Vermittlung und Förderung mündlichen Argumentierens als Lerngegenstand (vgl. Mundwiler et al. 2017) im Unterricht erfolgt, wird ersichtlich, dass auch hier der Aspekt der Körperlichkeit bislang fehlt. Mündliches Argumentieren wird im Unterricht, wenn es als Lerngegenstand thematisiert wird, i. d. R. theoretisch eingeführt (vgl. Mundwiler et al. 2017) und dann in *isolierten* Übungssettings wie z. B. in Form von Fishbowl-Diskussionen geübt (vgl. Heller und Morek 2015; Morek 2017). Die isoliert erprobten Argumentationen weisen deutliche Unterschiede zu Alltagsargumentationen auf, da u. a. die Interagierenden feste Positionen einnehmen und die Rederechtszuweisung über Moderator*innen erfolgt (vgl. Morek 2017; vgl. zsf. Kap. 4). Dabei orientiert sich die theoretische Vermittlung argumentativer Fähigkeiten v. a. an Formen der sprachlichen Markierung, beispielsweise der Einführung einer Begründung durch die Konjunktion *weil*, sowie dem Aufbau von Argumenten, in dem eine Behauptung durch eine Begründung gestützt und durch ein Beispiel veranschaulicht wird (vgl. Mundwiler et al. 2017: 98f.). Die Interagierenden sollen dann in der Diskussion eigene Beiträge mit denen von Vorredner*innen kohärent verknüpfen. Wie in der vorliegenden Studie gezeigt werden konnte und an anderer Stelle bereits von Jacquin (2018) beschrieben worden ist, können diese gegenseitigen Bezugnahmen hochgradig funktional auch gestisch hergestellt werden (vgl. Kap. 8.4.1, Beispiel 2). Wenn nun in der Vermittlung argumentativer Fähigkeiten (wie auch in der Beurteilung) auf die *sprachliche* Ressource fokussiert wird, wird das Bild mündlichen Argumentierens, das als Lerngegenstand bereits eine gewisse Ferne zu Alltagsargumentationen aufweist, noch stärker verengt. Daraus resultiert nun ein zweites sprachdidaktisches Desiderat, das an das o. g. anschließt: Wird mündliches Argumentieren im Unterricht zum Lerngegenstand, sind in der Vermittlung grundsätzlich multimodale Ausdrucksressourcen zu berücksichtigen, die funktional für das Argumentieren eingesetzt werden. Denkbar wäre i. d. S. beispielsweise, videographierte mündliche Argumentationen zum Ausgangspunkt der Vermittlung zu nehmen, sodass Schüler*innen dann über einen ersten deskriptiven Zugang beschreiben, welche Ressourcen sie in den argumentativen Handlungen seitens der Interagierenden beobachten können, die sie dann interpretierend hinsichtlich ihrer argumentativen Funktion einordnen (vgl. auch Vorschläge zur Erzähldidaktik Ohlhus und Stude 2009). Dieser Vermittlungsansatz würde zum einen eine schüler*innenseitige Auseinandersetzung mit Alltagsinteraktionen ermöglichen und zum anderen

der körperlichen Präsenz von Interagierenden im Gespräch gerecht werden, die bislang aus der Vermittlung ausgespart wird.

Resultierend aus dem zentralen Anliegen der vorliegenden Arbeit, Funktionen multimodaler Ressourcen für das mündliche Argumentieren sowie für den Erwerb argumentativer Fähigkeiten zu rekonstruieren, können die hier gewonnenen Erkenntnisse als Grundlage zukünftiger Erforschung multimodaler Gestalten in diskursiven Gattungen dienen.

13 Literatur

- Ambrazaitis, Gilbert & David House. 2017. Multimodal prominences: Exploring the patterning and usage of focal pitch accents, head beats and eyebrow beats in Swedish television news readings. *Speech Communication* 95(3). 10.1016/j.specom.2017.08.008.
- Andrén, Mats. 2010. *Children's gestures from 18 to 30 months* (Travaux de l'Institut de Linguistique de Lund 50). Lund: Dep. of Linguistics and Phonetics Centre for Languages and Literature Lund Univ.
- Andrén, Mats. 2014. 89. Gestures in Northern Europe: Children's gestures in Sweden. In Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill & Jana Bressemer (Hrsg.), *Body - Language - Communication: An international handbook on multimodality in human interaction* (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 38/2), 1282–1289. Berlin & Boston: De Gruyter Mouton.
- Antaki, Charles & Margaret Wetherell. 1999. Show Concessions. *Discourse Studies* 1(1). 10.1177/1461445699001001002.
- Arcidiacono, Francesco & Antonio Bova. 2015. Activity-bound and activity-unbound arguments in response to parental eat-directives at mealtimes: Differences and similarities in children of 3–5 and 6–9 years old. *Learning, culture and social interaction* 6(2). 10.1016/j.lcsi.2015.03.002.
- Arendt, Birte. 2019. *Argumentieren mit Peers: Erwerbsverläufe und -muster bei Kindergartenkindern* (Stauffenburg Linguistik Band 110). Tübingen: Stauffenburg Verlag.
- Arendt, Birte, Vivien Heller & Antje Krah. 2015. Zur Einführung. In Birte Arendt, Vivien Heller & Antje Krah (Hrsg.), *Kinder argumentieren: Interaktive Erwerbskontexte und -mechanismen*, 1–4. Göttingen: V&R.
- Arendt, Birte & Sara Zadunaisky Ehrlich. 2020. Literacy-related features in repetitions: Using the example of argumentative events of German- and Hebrew-speaking pre-schoolers. *Research on Children and Social Interaction* 4(1). 10.1558/rcsi.12414.
- Auer, Peter. 2007. Syntax als Prozess. In Heiko Hausendorf (Hrsg.), *Gespräch als Prozess: Linguistische Aspekte der Zeitlichkeit verbaler Interaktion*, 95–124. Tübingen: Gunter Narr Verlag.
- Auer, Peter. 2009. On-line syntax: Thoughts on the temporality of spoken language. *Language Sciences* 31(1). 1–13.
- Auer, Peter. 2013. *Sprachliche Interaktion: Eine Einführung anhand von 22 Klassikern*. Berlin & Boston: De Gruyter.
- Auer, Peter & Margret Selting. 2001. Der Beitrag der Prosodie zur Gesprächsorganisation. In Klaus Brinker, Gerd Antos, Wolfgang Heinemann & Sven F. Sager (Hrsg.), *Text- und Gesprächslinguistik: Ein internationales Handbuch zeitgenössischer Forschung* (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 16/2), 1122–1132. Berlin & Boston: De Gruyter Mouton.
- Auer, Peter & Susanne Uhmann. 1982. Aspekte der konversationellen Organisation von Bewertungen. *Deutsche Sprache* 1. 1–32.
- Auer, Peter & Elisabeth Zima. 2021. On word searches, gaze, and co-participation. *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion* 22. 390–425.
- Augst, Gerhard, Katrin Disselhoff, Alexandra Henrich, Thorsten Pohl & Paul-Ludwig Völzing. 2007. *Text - Sorten - Kompetenz: Eine echte Longitudinalstudie zur Entwicklung der Textkompetenz im Grundschulalter* (Theorie und Vermittlung der Sprache 48). Frankfurt am Main & Berlin: Peter Lang.
- Barth-Weingarten, Dagmar. 2006. fuzzy boundaries – Überlegungen zu einer Grammatik der gesprochenen Sprache nach konversationsanalytischen Kriterien. In Arnulf Deppermann, Reinhard Fiehler & Thomas Spranz-Fogasy (Hrsg.), *Grammatik und Interaktion. Untersuchungen zum Zusammenhang von grammatischen Strukturen und Gesprächsprozessen*, 67–93. Radolfzell: Verlag für Gesprächsforschung.
- Bavelas, Janet B. 1994. Gestures as Part of Speech: Methodological Implications. *Research on Language & Social Interaction* 27(3). 10.1207/s15327973rlsi2703_3.

- Bavelas, Janet B., Nicole Chovil, Linda Coates & Lori Roe. 1995. Gestures Specialized for Dialogue. *Personality and Social Psychology Bulletin* 21(4). 394–405.
- Bavelas, Janet B., Nicole Chovil, Douglas A. Lawrie & Allan Wade. 1992. Interactive gestures. *Discourse Processes* 15(4). 10.1080/01638539209544823.
- Bayer, Klaus. 1982. Mit Sprache bewerten. *Praxis Deutsch. Zeitschrift für den Deutschunterricht*(53). 15–25.
- Beaupoil-Hourdel, Pauline & Camille Debras. 2017. Developing Communicative postures: The emergence of shrugging in child communication. *Language, Interaction and Acquisition* 8(1). 10.1075/lia.8.1.05bea.
- Beaupoil-Hourdel, Pauline, Aliyah Morgenstern & Dominique Boutet. 2016. A Child’s Multimodal Negations from 1 to 4: The Interplay Between Modalities. In Pierre Larrivée & Chungmin Lee (Hrsg.), *Negation and Polarity: Experimental Perspectives* (Language, Cognition, and Mind 1), 95–123. Cham: Springer International Publishing.
- Becker-Mrotzek, Michael. 2009. B3 Mündliche Kommunikationskompetenz. In Michael Becker-Mrotzek (Hrsg.), *Mündliche Kommunikation und Gesprächsdidaktik*, 66–83. Baltmannsweiler: Schneider-Verlag Hohengehren.
- Bergmann, Jörg R. 2001. Das Konzept der Konversationsanalyse. In Klaus Brinker, Gerd Antos, Wolfgang Heinemann & Sven F. Sager (Hrsg.), *Text- und Gesprächslinguistik: Ein internationales Handbuch zeitgenössischer Forschung* (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 16/2), 919–927. Berlin & Boston: De Gruyter Mouton.
- Bergmann, Jörg R. & Thomas Luckmann. 1995. Reconstructive Genres of Everyday Communication. In Uta M. Quasthoff (Hrsg.), *Aspects of oral communication* (Research in text theory 21), 289–304. Berlin & New York: De Gruyter.
- Betz, Emma & Arnulf Deppermann. 2021. OKAY in responding and claiming understanding. In Emma Betz, Arnulf Deppermann, Lorenza Mondada & Marja-Leena Sorjonen (Hrsg.), *OKAY across languages: Toward a comparative approach to its use in talk-in-interaction* (Studies in Language and Social Interaction 34), 55–92. Amsterdam & Philadelphia: John Benjamins Publishing Company.
- Birkner, Karin, Peter Auer, Angelika Bauer & Helga Kotthoff. 2020. *Einführung in die Konversationsanalyse*. De Gruyter.
- Blum-Kulka, Shoshana. 2008. Language Socialization and Family Dinnertime Discourse. In Nancy H. Hornberger (Hrsg.), *Encyclopedia of Language and Education* (Springer eBook Collection Humanities, Social Science and Law), 2661–2673. Boston, MA: Springer Science+Business Media LLC.
- Bohle, Ulrike. 2014a. 108. Contemporary classification systems. In Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill & Jana Bressemer (Hrsg.), *Body - Language - Communication: An international handbook on multimodality in human interaction* (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 38/2), 1453–1461. Berlin & Boston: De Gruyter Mouton.
- Bohle, Ulrike. 2014b. 97. Gesture and conversational units. In Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill & Jana Bressemer (Hrsg.), *Body - Language - Communication: An international handbook on multimodality in human interaction* (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 38/2), 1360–1367. Berlin & Boston: De Gruyter Mouton.
- Borràs-Comes, Joan, Evangelia Kiagia & Pilar Prieto. 2019. Epistemic intonation and epistemic gesture are mutually co-expressive: Empirical results from two intonation-gesture matching tasks. *Journal of Pragmatics* 150(3). 10.1016/j.pragma.2019.07.004.
- Bressemer, Jana & Cornelia Müller. 2014a. 119. A repertoire of German recurrent gestures with pragmatic functions. In Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill & Jana Bressemer (Hrsg.), *Body - Language - Communication: An international handbook on multimodality in human interaction* (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 38/2), 1462–1473. Berlin & Boston: De Gruyter Mouton.

- dality in human interaction* (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 38/2), 1575–1591. Berlin & Boston: De Gruyter Mouton.
- Bressem, Jana & Cornelia Müller. 2014b. 120. The family of Away gestures: Negation, refusal, and negative assessment. In Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill & Jana Bressem (Hrsg.), *Body - Language - Communication: An international handbook on multimodality in human interaction* (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 38/2), 1592–1604. Berlin & Boston: De Gruyter Mouton.
- Brünner, Gisela & Elisabeth Gülich (Hrsg.). 2002a. *Krankheit verstehen. Interdisziplinäre Beiträge zur Sprache in Krankheitsdarstellungen*. Bielefeld: Aisthesis.
- Brünner, Gisela & Elisabeth Gülich. 2002b. Verfahren der Veranschaulichung in der Experten-Laien-Kommunikation. In Gisela Brünner & Elisabeth Gülich (Hrsg.), *Krankheit verstehen. Interdisziplinäre Beiträge zur Sprache in Krankheitsdarstellungen*, 17–93. Bielefeld: Aisthesis.
- Bublitz, Wolfram. 2001. Formen der Verständnissicherung in Gesprächen. In Klaus Brinker, Gerd Antos, Wolfgang Heinemann & Sven F. Sager (Hrsg.), *Text- und Gesprächslinguistik: Ein internationales Handbuch zeitgenössischer Forschung* (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 16/2), 1330–1340. Berlin & Boston: De Gruyter Mouton.
- Bücker, Jörg. 2015. Selbstbeantwortete Fragen in der Mündlichkeit - ein Fall für die Konstruktionsgrammatik? In Jörg Bücker, Susanne Günthner & Wolfgang Imo (Hrsg.), *Konstruktionsgrammatik V: Konstruktionen im Spannungsfeld von sequenziellen Mustern, kommunikativen Gattungen und Textsorten* (Stauffenburg Linguistik 77), 221–246. Tübingen: Stauffenburg Verlag.
- Capirci, Olga & Virginia Volterra. 2008. Gesture and speech: The emergence and development of a strong and changing partnership. *Gesture* 8(1). 10.1075/gest.8.1.04cap.
- Cienki, Alan. 2021. From the finger lift to the palm-up open hand when presenting a point: A methodological exploration of forms and functions. *Languages and Modalities* 1. 10.3897/lamo.1.68914.
- Ciolek, T. M. & Adam Kendon. 1980. Environment and the Spatial Arrangement of Conversational Encounters. *Sociological Inquiry* 50(3-4). 10.1111/j.1475-682X.1980.tb00022.x.
- Clark, Eve V. & Josie Bernicot. 2008. Repetition as ratification: how parents and children place information in common ground. *Journal of Child Language* 35(2). 10.1017/S0305000907008537.
- Clark, Herbert H. 1996. *Using language*. Cambridge: Cambridge University Press.
- Clark, Herbert H. 2003. Pointing and Placing. In Sotaro Kita (Hrsg.), *Pointing: Where language, culture, and cognition meet*, 243–268. New York: Psychology Press.
- Clark, Herbert H. & Catherine R. Marshall. 1981. Definite reference and mutual knowledge. In Aravind K. Joshi, Bonnie L. Webber & Ivan A. Sag (Hrsg.), *Elements of discourse understanding*, 10–63. Cambridge & New York: Cambridge University Press.
- Clift, Rebecca. 2020. Stability and visibility in embodied action: The 'Palm Up' in interaction. *Journal of Pragmatics*(169). 190–205.
- Clift, Rebecca. 2021. Embodiment in Dissent: The Eye Roll as an Interactional Practice. *Research on Language and Social Interaction* 54(3). 10.1080/08351813.2021.1936858.
- Colletta, Jean-Marc. 2022. On the Codevelopment of Gesture and Monologic Discourse in Children. In Aliyah Morgenstern & Susan Goldin-Meadow (Hrsg.), *Gesture in language: Development across the lifespan* (Language and the human lifespan series), 205–242. Washington & Berlin: American Psychological Association; De Gruyter.
- Colletta, Jean-Marc, Michèle Guidetti, Olga Capirci, Carla Cristilli, Ozlem E. Demir, Ramona N. Kunene-Nicolas & Susan Levine. 2015. Effects of age and language on co-speech gesture production: an investigation of French, American, and Italian children's narratives. *Journal of child language* 42(1). 10.1017/S0305000913000585.
- Colletta, Jean-Marc & Catherine Pellenq. 2009. The Development of Multimodal Explanations in French Children. In Marilyn A. Nippold & Cheryl M. Scott (Hrsg.), *Expository Discourse in Children, Adolescents, and Adults: Development and Disorders*, 63–99. Erlbaum: Taylor & Francis.

- Colletta, Jean-Marc, Catherine Pellenq & Michèle Guidetti. 2010. Age-related changes in co-speech gesture and narrative: Evidence from French children and adults. *Speech Communication* 52(6). 10.1016/j.specom.2010.02.009.
- Condon, William S. & William D. Ogston. 1966. Sound film analysis of normal and pathological behavior patterns. *The Journal of nervous and mental disease* 143(4). 10.1097/00005053-196610000-00005.
- Condon, William S. & William D. Ogston. 1971. Speech and Body Motion Synchrony of the Speaker-Hearer. In Paul M. Kjeldergaard & David L. Horton (Hrsg.), *Perception of language: Proceedings of a symposium of the Learning Research and Development Center. Conference chairman: Paul Myron Kjeldergaard. Edited by David L. Horton and James J. Jenkins* (Merrill's international psychology series), 150–173. Columbus, Ohio: Merrill.
- Cooperrider, Kensy, Natasha Abner & Susan Goldin-Meadow. 2018. The Palm-Up Puzzle: Meanings and Origins of a Widespread Form in Gesture and Sign. *Frontiers in Communication* 3. 10.3389/fcomm.2018.00023.
- Couper-Kuhlen, Elizabeth. 2014. What does grammar tell us about action? *Pragmatics. Quarterly Publication of the International Pragmatics Association (IPrA)* 37(3). 10.1075/prag.24.3.08cou.
- Couper-Kuhlen, Elizabeth & Sandra A. Thompson. 2000. Concessive patterns in conversation. In Elizabeth C. Traugott, Bernd Kortmann & Elizabeth Couper-Kuhlen (Hrsg.), *Cause - Condition - Concession - Contrast* (Topics in English Linguistics 33), 381–410. Berlin, New York: Mouton de Gruyter.
- Cristilli, Carla. 2014. How gestures help children to track reference in narrative. In Mandana Seyfedinipur & Marianne Gullberg (Hrsg.), *From Gesture in Conversation to Visible Action as Utterance*, 331–350. Amsterdam: John Benjamins Publishing Company.
- Debras, Camille. 2017. The Shrug: Forms and Meanings of a Compound Enactment. *Gesture* 16(1). 10.1075/gest.16.1.01deb.
- Debras, Camille. 2021. Multimodal profiles of *je (ne) sais pas* in spoken French. *Journal of Pragmatics* 182. 10.1016/j.pragma.2021.05.018.
- Debras, Camille & Alan Cienki. 2012. Some Uses of Head Tilts and Shoulder Shrugs during Human Interaction, and Their Relation to Stancetaking. *International proceedings of the ASE 2012 International Conference of Social Computing*. 10.1109/SocialCom-PASSAT.2012.136.
- Deppermann, Arnulf. 2003. Desiderata einer gesprächsanalytischen Argumentationsforschung. In Arnulf Deppermann & Martin Hartung (Hrsg.), *Argumentieren in Gesprächen: Gesprächsanalytische Studien* (Stauffenburg Linguistik 28), 10–26. Tübingen: Stauffenburg Verlag.
- Deppermann, Arnulf. 2008. *Gespräche analysieren: Eine Einführung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Deppermann, Arnulf. 2009. Verstehensdefizit als Antwortverpflichtung: Interaktionale Eigenschaften der Modalpartikel denn in Fragen. In Susanne Günthner & Jörg Bücker (Hrsg.), *Grammatik im Gespräch: Konstruktionen der Selbst- und Fremdpositionierung* (Linguistik - Impulse & Tendenzen 33), 23–56. De Gruyter.
- Deppermann, Arnulf. 2013. Multimodal interaction from a conversation analytic perspective. *Journal of Pragmatics* 46(1). 10.1016/j.pragma.2012.11.014.
- Deppermann, Arnulf. 2018. 3. Sprache in der multimodalen Interaktion. In Arnulf Deppermann & Silke Reineke (Hrsg.), *Sprache im kommunikativen, interaktiven und kulturellen Kontext* (Germanistische Sprachwissenschaft um 2020), 51–86. Berlin/Boston: De Gruyter.
- Deppermann, Arnulf, Helmuth Feilke & Angelika Linke. 2016. Sprachliche und kommunikative Praktiken: Eine Annäherung aus linguistischer Sicht. In Arnulf Deppermann, Helmuth Feilke & Angelika Linke (Hrsg.), *Sprachliche und kommunikative Praktiken*, 1–24. Berlin & Boston: De Gruyter.
- Deppermann, Arnulf & Gabriele Lucius-Hoene. 2003. Argumentatives Erzählen. In Arnulf Deppermann & Martin Hartung (Hrsg.), *Argumentieren in Gesprächen: Gesprächsanalytische Studien* (Stauffenburg Linguistik), 130–143. Tübingen: Stauffenburg Verlag.

- Deppermann, Arnulf & Reinhold Schmitt. 2007. Koordination. Zur Begründung eines neuen Forschungsgegenstandes. In Reinhold Schmitt (Hrsg.), *Koordination: Analysen zur multimodalen Interaktion* (Studien zur Deutschen Sprache 38), 15–54. Tübingen: Gunter Narr Verlag.
- Deppermann, Arnulf & Reinhold Schmitt. 2008. Verstehensdokumentationen: Zur Phänomenologie von Verstehen in der Interaktion. *Deutsche Sprache* 36(3). 220–245.
- Domenech, Madeleine, Vivien Heller & Inger Petersen. 2018. Argumentieren mündlich, schriftlich, zweitsprachlich: Verfahren und Anforderungen. *Beiträge zur Fremdsprachenvermittlung*(26). 15–35.
- Efron, David. 1941/1972. *Gesture, race and culture: a tentative study of some of the spatio-temporal and "linguistic" aspects of the gestural behavior of Eastern Jews and Southern Italians in New York City, living under similar as well as different environmental conditions*. The Hague [u.a.]: Mouton.
- Eggs, Frederike. 2003. „Weiß sowieso jeder“ Eine funktional-grammatische Analyse der Ausdrücke sowieso, eh, ohnedies und ohnehin. In Ludger Hoffmann (Hrsg.), *Funktionale Syntax: Die pragmatische Perspektive*, 270–306. Berlin & Boston: De Gruyter.
- Ehlich, Konrad. 2014. Argumentieren als sprachliche Ressource des diskursiven Lernens. In Antonie Hornung, Gabriella Carobbio & Daniela Sorrentino (Hrsg.), *Diskursive und textuelle Strukturen in der Hochschuldidaktik: Deutsch und Italienisch im Vergleich* (Sprach-Vermittlungen 12), 41–54. Münster & New York: Waxmann.
- Ehlich, Konrad & Jochen Rehbein. 1986. *Muster und Institution: Untersuchungen zur schulischen Kommunikation* (Kommunikation und Institution 15). Tübingen: Narr.
- Eisenberg, Ann R. & Catherine Garvey. 1981. Children's use of verbal strategies in resolving conflicts. *Discourse Processes* 4(2). 10.1080/01638538109544512.
- Ekman, Paul. 1979. About brows: emotional and conversational signals. In Mario von Cranach, Klaus Foppa, Wolf Lepenies & Detlev Ploog (Hrsg.), *Human ethology: Claims and limits of a new discipline ; contributions to the Colloquium sponsored by the Werner-Reimers-Stiftung*, 169–249. Paris & Cambridge.: Éd. de la Maison des Sciences de l'Homme; Cambridge University Press.
- Ekman, Paul & Wallace V. Friesen. 1969. The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding. *Semiotica* 1(1). 10.1515/semi.1969.1.1.49.
- Ekman, Paul & Wallace V. Friesen. 1972. Hand Movements. *Journal of Communication* 22(4). 10.1111/j.1460-2466.1972.tb00163.x.
- Enfield, Nick J. 2009. *The anatomy of meaning: Speech, gesture, and composite utterances* (Language, culture and cognition 8). Cambridge: Cambridge University Press.
- Enfield, Nick J. 2013. 44. A “Composite Utterances” approach to meaning. In Ellen Fricke, Alan Cienki, David McNeill, Silva Ladewig & Sedinha Tessoroff (Hrsg.), *Body - Language - Communication: An international Handbook on Multimodality in Human Interaction* (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science [HSK] 38/1), 689–707. Berlin & Boston: De Gruyter Mouton.
- Enfield, Nick J., Sotaro Kita & Jan P. d. Ruiter. 2007. Primary and secondary pragmatic functions of pointing gestures. *Journal of Pragmatics* 39(10). 10.1016/j.pragma.2007.03.001.
- Feilke, Helmuth. 2012. Was sind Textroutinen? - Zur Theorie und Methodik des Forschungsfeldes. In Helmuth Feilke & Katrin Lehnen (Hrsg.), *Textroutinen: Theorie, Erwerb und didaktisch-mediale Modellierung*, 1–31. Frankfurt am Main: Lang.
- Ferré, Gaëlle. 2012. Functions of three open-palm hand gestures. *Journal Multimodal Communication* 1(1). 10.1515/mc-2012-0002.
- Fricke, Ellen. 2007. *Origo, Geste und Raum: Lokaldeixis im Deutschen* (Linguistik – Impulse & Tendenzen 24). Berlin: De Gruyter.
- Garfinkel, Harold. 1967. *Studies in Ethnomethodology*. Englewood Cliffs: Prentice-Hall.
- Gätje, Olaf, Sara Rezat & Torsten Steinhoff. 2012. Positionierung. Zur Entwicklung des Gebrauchs modalisierender Prozeduren in argumentativen Texten von Schülern und Studenten. In Helmuth Feilke & Katrin Lehnen (Hrsg.), *Textroutinen: Theorie, Erwerb und didaktisch-mediale Modellierung*, 125–153. Frankfurt am Main: Lang.

- Givens, David B. 1977. Shoulder Shrugging: A Densely Communicative Expressive Behavior. *Semiotica* 19(1/2). 10.1515/semi.1977.19.1-2.13.
- Goetz, Peggy J. 2010. The development of verbal justifications in the conversations of preschool children and adults. *First Language* 30(3-4). 10.1177/0142723710370522.
- Goetz, Peggy. J. & Marilyn Shatz. 1999. When and how peers give reasons: justifications in the talk of middle school children. *Journal of child language* 26(3). 10.1017/s0305000999003980.
- Goffman, Erving. 1963. *Behavior in public places: notes on the social organization of gatherings*. New York: Free Press.
- Gohl, Christine. 2006. *Begründen im Gespräch: Eine Untersuchung sprachlicher Praktiken zur Realisierung von Begründungen im gesprochenen Deutsch* (Reihe Germanistische Linguistik 267). Tübingen: Max Niemeyer Verlag.
- Golato, Andrea. 2018. Chapter 14. Turn-initial naja in German. In John Heritage & Marja-Leena Sorjonen (Hrsg.), *Between turn and sequence: Turn-initial particles across languages* (Studies in Language and Social Interaction 31), 413–444. Amsterdam & Philadelphia: John Benjamins Publishing Company.
- Goodwin, Charles. 1979. The Interactive Construction of a Sentence in Natural Conversation. *Everyday Language: Studies in Ethnomethodology*. 97–121.
- Goodwin, Charles. 1981. *Conversational organization: interaction between speakers and hearers*. New York & London: Academic Press.
- Goodwin, Charles. 1986. Gestures as a resource for the organization of mutual orientation. *Semiotica* 62(1-2). 10.1515/semi.1986.62.1-2.29.
- Goodwin, Charles. 2000. Action and embodiment within situated human interaction. *Journal of Pragmatics* 32(10). 10.1016/S0378-2166(99)00096-X.
- Goodwin, Charles. 2007. Participation, stance and affect in the organization of activities. *Discourse & Society* 18(1). 53–73.
- Goodwin, Charles. 2011. Contextures of Action. In Jürgen Streeck, Charles Goodwin & Curtis LeBaron (Hrsg.), *Embodied Interaction: Language and body in the material world*, 182–193. Cambridge [u. a]: Cambridge University Press.
- Goodwin, Charles. 2018. *Co-Operative Action*. Cambridge: Cambridge University Press.
- Goodwin, Charles & Marjorie H. Goodwin. 2004. Participation. In Alessandro Duranti (Hrsg.), *A companion to linguistic anthropology* (Blackwell companions to anthropology 1), 222–244. Malden, Mass.: Blackwell.
- Goodwin, Charles & John Heritage. 1990. Conversation Analysis. *Annual Review of Anthropology* 19(1). 283–307.
- Goodwin, Marjorie H. 1980. Processes of Mutual Monitoring Implicated in the Production of Description Sequences. *Sociological Inquiry* 50(3-4). 10.1111/j.1475-682X.1980.tb00024.x.
- Goodwin, Marjorie H. 1983. Aggravated Correction and Disagreement in Children's Conversations. *Journal of Pragmatics* 7(6). 657–677.
- Goodwin, Marjorie H. 1990. *He-said-she-said: Talk as social organization among black children*. Bloomington: Indiana University Press.
- Goodwin, Marjorie H. & Asta Cekaite. 2018. *Embodied Family Choreography: Practices of Control, Care, and Mundane Creativity*. London & New York: Routledge.
- Goodwin, Marjorie H. & Charles Goodwin. 1986. Gesture and coparticipation in searching for a word. *Semiotica* 62(1-2). 51–76.
- Goodwin, Marjorie H. & Charles Goodwin. 1987. Children's Arguing. In Susan U. Philips, Susan Steele & Christine Tanz (Hrsg.), *Language, Gender, and Sex in Comparative Perspective*, 200–248. Cambridge: Cambridge University Press.
- Goodwin, Marjorie H. & Amy Kyratzis. 2007. Children Socializing Children: Practices for Negotiating the Social Order Among Peers. *Research on Language and Social Interaction* 40(4). 10.1080/08351810701471260.
- Graziano, Maria. 2014a. Gestures in Southern Europe: Children's pragmatic gestures in Italy. In Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill & Jana Bressemer (Hrsg.),

- Body - Language - Communication: An international handbook on multimodality in human interaction* (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 38/2), 1253–1258. Berlin & Boston: De Gruyter Mouton.
- Graziano, Maria. 2014b. The Development of Two Pragmatic Gestures of the So-Called Open Hand Supine Family in Italian Children. In Mandana Seyfeddinipur & Marianne Gullberg (Hrsg.), *From Gesture in Conversation to Visible Action as Utterance*, 311–328. Amsterdam: John Benjamins Publishing Company.
- Grundler, Elke. 2011. *Kompetent argumentieren: Ein gesprächsanalytisch fundiertes Modell*. Tübingen: Stauffenburg Verlag.
- Grundler, Elke. 2015. Interaktive Argumententwicklung in schulischen Argumentationen. *Mitteilungen des Deutschen Germanistenverbandes* 62(1). 47–60.
- Grundler, Elke. 2022. Mündliches Argumentieren. In Steffen Gailberger & Frauke Wietzke (Hrsg.), *Handbuch Kompetenzorientierter Deutschunterricht: Mit E-Book inside und Online-Materialien*, 562–595. Weinheim: Julius Beltz GmbH & Co. KG.
- Grundler, Elke, Sara Rezat & Sabine Schmölzer-Eibinger. 2020. Positionierungen in argumentativen Gesprächen und Briefen der Schule. *Zeitschrift für Angewandte Linguistik* 72(1). 10.1515/zfal-2020-2026.
- Guidetti, Michèle, Kateřina Fibigerová & Jean-Marc Colletta. 2014. Gestures and multimodal development: Some key issues for language acquisition. In Mandana Seyfeddinipur & Marianne Gullberg (Hrsg.), *From Gesture in Conversation to Visible Action as Utterance*, 351–370. Amsterdam: John Benjamins Publishing Company.
- Gülich, Elisabeth & Lorenza Mondada. 2008. *Konversationsanalyse: Eine Einführung am Beispiel des Französischen* (Romanistische Arbeitshefte 52). Berlin & Tübingen: De Gruyter; Max Niemeyer Verlag. <https://www.degruyter.com/isbn/9783110974058>.
- Günthner, Susanne. 1999. Entwickelt sich der Konzessivkonnekter obwohl zum Diskursmarker? Grammatikalisierungstendenzen im gesprochenen Deutsch. *Linguistische Berichte* 180. 409–446.
- Günthner, Susanne. 2015. Ko-Konstruktionen im Gespräch: Zwischen Kollaboration und Konfrontation. In Ulrich Dausendschön-Gay, Elisabeth Gülich & Ulrich Krafft (Hrsg.), *Ko-Konstruktion in der Interaktion: Die gemeinsame Arbeit an Äußerungen und anderen sozialen Ereignissen*, 55–74. Bielefeld: transcript.
- Günthner, Susanne & Wolfgang Imo. 2003. Die Reanalyse von Matrixsätzen als Diskursmarker: ich mein-Konstruktionen im gesprochenen Deutsch. *InLiSt* 37. 1–35.
- Günthner, Susanne & Hubert A. Knoblauch. 1996. Die Analyse kommunikativer Gattungen in Alltagsinteraktionen. In Susanne Michaelis (Hrsg.), *Texte - Konstitution, Verarbeitung, Typik* (Edition Linguistik 13), 35–57. München: Lincom.
- Haddington, Pentti, Tiina Keisanen, Lorenza Mondada & Maurice Nevile (Hrsg.). 2014. *Multiactivity in Social Interaction: Beyond multitasking*. Amsterdam & Philadelphia: John Benjamins Publishing Company.
- Halliday, Michael A. K., M. A. K. Halliday & Ruqaiya Hasan. 1980. *Cohesion in English* (English language series 9). London: Longman.
- Hannken-Illjes, Kati & Ines Bose. 2019. Frozen Children in Argumentation Between the Agonistic and Cooperation. *Informal Logic* 39(4). 465–495.
- Harrison, Simon. 2010. Evidence for node and scope of negation in coverbal gesture. *Gesture* 10(1). 10.1075/gest.10.1.03har.
- Harrison, Simon. 2018. *The Impulse to Gesture: Where Language, Minds, and Bodies Intersect*. Cambridge: Cambridge University Press.
- Harrison, Simon & Pierre Larrivée. 2016. Morphosyntactic Correlates of Gestures: A Gesture Associated with Negation in French and Its Organisation with Speech. In Pierre Larrivée & Chungmin Lee (Hrsg.), *Negation and Polarity: Experimental Perspectives* (Language, Cognition, and Mind 1), 75–94. Cham: Springer International Publishing.
- Hausendorf, Heiko, Lorenza Mondada & Reinhold Schmitt (Hrsg.). 2016. *Raum als interaktive Ressource* (Studien zur Deutschen Sprache 62). Tübingen: Narr.

- Hausendorf, Heiko & Uta M. Quasthoff. 1996/2005. *Sprachentwicklung und Interaktion: Eine linguistische Studie zum Erwerb von Diskursfähigkeiten*. Wiesbaden & s.l.: VS Verlag für Sozialwissenschaften.
- Hausendorf, Heiko & Uta M. Quasthoff. 2005. Konversations-/Diskursanalyse: (Sprach-) Entwicklung durch Interaktion. In Günter Mey (Hrsg.), *Handbuch qualitative Entwicklungspsychologie*, 585–618. Köln: KSV Kölner Studien-Verl.
- Hauser, Stefan & Martin Luginbühl. 2017. Wenn Kinder argumentieren: Grundlagen und erste Befunde einer Studie zur mündlichen Argumentationskompetenz von Schulkindern. In Iris Meißner & Eva L. Wyss (Hrsg.), *Begründen - Erklären - Argumentieren: Konzepte und Modellierungen in der Angewandten Linguistik* (Stauffenburg Linguistik 93), 89–105. Tübingen: Stauffenburg Verlag.
- Heath, Christian. 1984. Talk and reciprocity: sequential organization in speech and body movement. In J. M. Atkinson & John Heritage (Hrsg.), *Structures of Social Action: Studies in Conversation Analysis* (Studies in emotion and social interaction), 247–265. Cambridge: Cambridge University Press.
- Heath, Christian. 1985. The consultation's end: the coordination of speech and body movement. *International Journal of the Sociology of Language* 51. 10.1515/ijsl.1985.51.27.
- Heath, Christian & Paul Luff. 2013. Embodied Action and Organizational Activity. In Jack Sidnell & Tanya Stivers (Hrsg.), *The Handbook of Conversation Analysis*, 283–307. Chichester: Wiley-Blackwell.
- Heller, Vivien. 2012. *Kommunikative Erfahrungen von Kindern in Familie und Unterricht: Passungen und Divergenzen*. Tübingen: Stauffenburg Verlag.
- Heller, Vivien. 2017. Lerngelegenheiten für Erklären und Argumentieren: Wie partizipieren mehrsprachige Schülerinnen und Schüler an bildungssprachlichen Praktiken? In Isabel Fuchs, Stefan Jeuk & Werner Knapp (Hrsg.), *Mehrsprachigkeit: Spracherwerb, Unterrichtsprozesse, Seiteneinstieg*, 173–190. Stuttgart: Fillibach bei Klett.
- Heller, Vivien. 2018a. Embodying epistemic responsibility. *Research on Children and Social Interaction* 2(2). 10.1558/rcsi.37391.
- Heller, Vivien. 2018b. Jenseits des Hier und Jetzt: Multimodale Praktiken der Versetzung in Erzählinteraktionen kleiner Kinder. *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 19. 242–274.
- Heller, Vivien. 2019. Embodied displacements in young German children's storytelling. *Research on Children and Social Interaction* 3(1-2). 10.1558/rcsi.37311.
- Heller, Vivien. 2021a. Die sprachlich-diskursive Darstellung komplexer Zusammenhänge im Fachunterricht. In Uta Quasthoff, Vivien Heller & Miriam Morek (Hrsg.), *Diskurserwerb in Familie, Peer-group und Unterricht* (Reihe Germanistische Linguistik 324), 303–346. Berlin & Boston: De Gruyter.
- Heller, Vivien. 2021b. Embodied Displays of "Doing Thinking." Epistemic and Interactive Functions of Thinking Displays in Children's Argumentative Activities. *Frontiers in psychology* 12. 10.3389/fpsyg.2021.636671.
- Heller, Vivien & Antje Krah. 2015. Wie Eltern und Kinder argumentieren: Interaktionsmuster und ihr erwerbssupportives Potenzial im längsschnittlichen Vergleich. *Mitteilungen des Deutschen Germanistenverbandes* 62(1). 5–20.
- Heller, Vivien & Miriam Morek. 2015. Unterrichtsgespräche als Erwerbskontext: Kommunikative Gelegenheiten für bildungssprachliche Praktiken erkennen und nutzen. *Online-Plattform für Literalität*(3). 1–23.
- Heller, Vivien, Uta Quasthoff, Anna Vogler & Susanne Prediger. 2017. Bildungssprachliche Praktiken aus professioneller Sicht. : Wie deuten Lehrkräfte Erklärungen und Begründungen von Kindern? In Bernt Ahrenholz, Britta Hövelbrinks & Claudia Schmellentin-Britz (Hrsg.), *Fachunterricht und Sprache in schulischen Lehr-/Lernprozessen*, 139–160. Tübingen: Narr Francke Attempto Verlag.
- Heller, Vivien, Nora Schönfelder & Denise Robbins. 2023. Displaying a Critical Stance: Eyebrow Contractions in Children's Multimodal Oppositional Actions. *Social Interaction Video-Based Studies of Human Sociality* 6(3).

- Helmer, Henrike & Arnulf Deppermann. 2017. ICH WEIß NICHT zwischen Assertion und Diskursmarker: Verwendungsspektren eines Ausdrucks und Überlegungen zu Kriterien für Diskursmarker. In Hardarik Blühdorn, Arnulf Deppermann, Henrike Helmer & Thomas Spranz-Fogasy (Hrsg.), *Diskursmarker im Deutschen: Reflexionen und Analysen*, 131–156. Göttingen: Verlag für Gesprächsforschung.
- Helmer, Henrike, Arnulf Deppermann & Silke Reineke. 2017. Antwort, epistemischer Marker oder Widerspruch? Sequenzielle, semantische und pragmatische Eigenschaften von ICH WEISS NICHT. In Arnulf Deppermann, Nadine Proske & Arne Zeschel (Hrsg.), *Verben im interaktiven Kontext. Bewegungsverben und mentale Verben im gesprochenen Deutsch* (Studien zur Deutschen Sprache 74), 377–405. Tübingen: Narr Francke Attempto Verlag.
- Heritage, John. 1984. A change-of-state token and aspects of its sequential placement. In J. M. Atkinson & John Heritage (Hrsg.), *Structures of Social Action: Studies in Conversation Analysis* (Studies in emotion and social interaction), 299–345. Cambridge: Cambridge University Press.
- Heritage, John. 2011. Conversation Analysis: Practices and Methods. In David Silverman (Hrsg.), *Qualitative Research: Theory, Method and Practice*, 208–230. Los Angeles: Sage.
- Heritage, John. 2012a. Epistemics in Action: Action Formation and Territories of Knowledge. *Research on Language & Social Interaction* 45(1). 10.1080/08351813.2012.646684.
- Heritage, John. 2012b. The Epistemic Engine: Sequence Organization and Territories of Knowledge. *Research on Language & Social Interaction* 45(1). 10.1080/08351813.2012.646685.
- Heritage, John. 2013. Epistemics in Conversation. In Jack Sidnell & Tanya Stivers (Hrsg.), *The Handbook of Conversation Analysis*, 370–394. Chichester: Wiley-Blackwell.
- Heritage, John & Geoffrey Raymond. 2005. The Terms of Agreement: Indexing Epistemic Authority and Subordination in Talk-in-Interaction. *Social Psychology Quarterly* 68(1). 10.1177/019027250506800103.
- Heritage, John & Geoffrey Raymond. 2012. Navigating epistemic landscapes: acquiescence, agency and resistance in responses to polar questions. In Jan-Peter de Ruiter (Hrsg.), *Questions: Formal, functional and interactional perspectives* (Language culture and cognition 12), 179–192. Cambridge: Cambridge University Press.
- Hindmarsh, Jon & Christian Heath. 2000. Embodied reference: A study of deixis in workplace interaction. *Journal of Pragmatics* 32(12). 10.1016/S0378-2166(99)00122-8.
- Hofstetter, Emily & Leelo Keevallik. 2020. Embodied interaction. In Jan-Ola Östman & Jef Verschueren (Hrsg.), *Handbook of Pragmatics: 23rd Annual Installment* (Handbook of Pragmatics), 111–138. Amsterdam & Philadelphia: John Benjamins Publishing Company.
- Holler, Judith, Kobin H. Kendrick & Stephen C. Levinson. 2018. Processing language in face-to-face conversation: Questions with gestures get faster responses. *Psychonomic bulletin & review* 25(5). 10.3758/s13423-017-1363-z.
- Houtkoop, Hanneke. 1987/2019. *Establishing agreement: An analysis of proposal-acceptance sequences* (Studies of argumentation in pragmatics and discourse analysis 4). Berlin & Boston: De Gruyter Mouton.
- Hrncal, Christine. 2018. Bewertungsinteraktion. In Jan Gerwinski, Stephan Habscheid, Erika Linz, Marit Besthorn, Mareike Hesse, Christine Hrncal & Eva Schlinkmann (Hrsg.), *Theater im Gespräch: Sprachliche Publikumspraktiken in der Theaterpause*, 235–300. Berlin: De Gruyter.
- Hrncal, Christine. 2020. *Bewertungsinteraktionen in der Theaterpause: Eine gesprächsanalytische Untersuchung von Pausengesprächen im Theaterfoyer*. Berlin & Boston: De Gruyter.
- Hübscher, Iris & Pilar Prieto. 2019. Gestural and Prosodic Development Act as Sister Systems and Jointly Pave the Way for Children's Sociopragmatic Development. *Frontiers in psychology* 10. 10.3389/fpsyg.2019.01259.
- Hübscher, Iris, Laura Vincze & Pilar Prieto. 2019. Children's Signaling of Their Uncertain Knowledge State: Prosody, Face, and Body Cues Come First. *Language Learning and Development* 15(4). 10.1080/15475441.2019.1645669.

- Imo, Wolfgang. 2006. Zur Anwendung der Construction Grammar auf die gesprochene Sprache – der Fall „ich mein(e)“. In Vilmos Ágel & Mathilde Hennig (Hrsg.), *Zugänge zur Grammatik der gesprochenen Sprache* (Reihe Germanistische Linguistik 269), 3–34. Berlin: De Gruyter.
- Inbar, Anna. 2022. The Raised Index Finger gesture in Hebrew multimodal interaction. *Gesture* 21(2–3). 10.1075/gest.21001.inb.
- Inbar, Anna. 2025. Interactional practices accomplished by index-finger pointing directed at the addressee in Hebrew face-to-face interaction. *Frontiers in psychology* 15. 10.3389/fpsyg.2024.1463449.
- Jacobs, Scott & Sally Jackson. 1982. Conversational Argument: A Discourse Analytic Approach. In J. R. Cox & Charles A. Willard (Hrsg.), *Advances in argumentation theory and research*, 205–237. Carbondale: Published for the American Forensic Association by Southern Illinois University Press.
- Jacquin, Jérôme. 2015. Multimodal Counter-Argumentation in the Workplace: The Contribution of Gesture and Gaze to the Expression of Disagreement. *Colloque Gesture and speech in interaction (GESPIN 4)*. 155–160.
- Jacquin, Jérôme. 2017. Embodied argumentation in public debates: The role of gestures in the segmentation of argumentative moves. In Assimakis Tseronis & Charles Forceville (Hrsg.), *Multimodal argumentation and rhetoric in media genres* (Argumentation in context (AIC) 14), 239–261. Amsterdam & Philadelphia: John Benjamins Publishing Company.
- Jacquin, Jérôme. 2018. Multimodal positioning and reference in argumentative talk-in-interaction. *International Review of Pragmatics* 10(2). 10.1163/18773109-01002004.
- Jefferson, Gail. 1979. A technique for inviting laughter and its subsequent acceptance/declination. In George Psathas (Hrsg.), *Everyday language: Studies in ethnomethodology*, 79–96. New York: Irvington.
- Jehoul, Annelies, Geert Brône & Kurt Feyaerts. 2017. The shrug as marker of obviousness. *Linguistics Vanguard* 3(1). 10.1515/lingvan-2016-0082.
- Kärkkäinen, Elise & Sandra A. Thompson. 2018. Language and bodily resources: ‘Response packages’ in response to polar questions in English. *Journal of Pragmatics* 123. 10.1016/j.pragma.2017.05.003.
- Katz-Bernstein, Nitza & Uta M. Quasthoff. 2007. Diskursfähigkeiten. In Manfred Grohnfeldt (Hrsg.), *Lexikon der Sprachtherapie*, 73–76. Stuttgart: Kohlhammer Verlag.
- Kaukomaa, Timo, Anssi Peräkylä & Johanna Ruusuvuori. 2014. Foreshadowing a problem: Turn-opening frowns in conversation. *Journal of Pragmatics* 71(2). 10.1016/j.pragma.2014.08.002.
- Kendon, Adam. 1967. Some functions of gaze-direction in social interaction. *Acta Psychologica* 26. 22–63.
- Kendon, Adam. 1970. Movement coordination in social interaction: Some examples described. *Acta Psychologica* 32. 10.1016/0001-6918(70)90094-6.
- Kendon, Adam. 1972a. Review of the book *Kinesics and Context: Essays on Body Motion Communication* by Ray Birdwhistell. *The American Journal of Psychology* 85(3). 10.2307/1420845.
- Kendon, Adam. 1972b. Some Relationships Between Body Motion and Speech: An Analysis of an Example. In Aron W. Siegman & Benjamin Pope (Hrsg.), *Studies in dyadic communication* (Pergamon general psychology series PGPS-7), 177–210. New York: Pergamon Press.
- Kendon, Adam. 1980. Gesticulation and Speech: Two Aspects of the Process of Utterance. In Mary R. Key (Hrsg.), *The Relationship of Verbal and Nonverbal Communication* (Contributions to the Sociology of Language [CSL] 25), 207–227. The Hague & Paris: De Gruyter Mouton.
- Kendon, Adam. 1990a. Some context for Context Analysis: a view of the origins of structural studies of face-to-face interaction. In Adam Kendon (Hrsg.), *Conducting interaction: Patterns of behavior in focused encounters* (Studies in interactional sociolinguistics 7), 15–49. Cambridge: Cambridge University Press.
- Kendon, Adam. 1990b. Spatial organization in social encounters: the F-formation system. In Adam Kendon (Hrsg.), *Conducting interaction: Patterns of behavior in focused encounters* (Studies in interactional sociolinguistics 7), 209–237. Cambridge: Cambridge University Press.

- Kendon, Adam. 1995a. Gestures as illocutionary and discourse structure markers in Southern Italian conversation. *Journal of Pragmatics* 23(3). 10.1016/0378-2166(94)00037-F.
- Kendon, Adam. 1995b. *The Open Hand. Paper presented at the conference 'Gestures Compared Cross-Linguistically', Linguistic Institute, University of New Mexico, Albuquerque, June 1995.*
- Kendon, Adam. 1997. Gesture. *Annual Review of Anthropology* 26. 109–128.
- Kendon, Adam. 2000. Language and gesture: unity or duality? In David McNeill (Hrsg.), *Language and gesture* (Language, culture, and cognition 2), 47–63. Cambridge: Cambridge University Press.
- Kendon, Adam. 2002. Some uses of the head shake. *Gesture* 2(2). 10.1075/gest.2.2.03ken.
- Kendon, Adam. 2004. *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Kendon, Adam. 2013. 1. Exploring the utterance roles of visible bodily action: A personal account. In Ellen Fricke, Alan Cienki, David McNeill, Silva Ladewig & Sedinha Tessendorf (Hrsg.), *Body - Language - Communication: An international Handbook on Multimodality in Human Interaction* (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science [HSK] 38/1), 7–28. Berlin & Boston: De Gruyter Mouton.
- Kendon, Adam. 2017. Pragmatic functions of gestures: Some observations on the history of their study and their nature. *Gesture* 16(2). 10.1075/gest.16.2.01ken.
- Kendrick, Robin H., Judith Holler & Stephen C. Levinson. 2023. Turn-taking in human face-to-face interaction is multimodal: gaze direction and manual gestures aid the coordination of turn transitions. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 378(1875). 10.1098/rstb.2021.0473.
- Kern, Friederike. 2003. Bedeutung und Interaktion: Spielerklärungen bei Kindern. In Stefanie Haberzettl & Heide Wegener (Hrsg.), *Die Rolle der Konzeptionalisierung im Spracherwerb*, 257–270. Frankfurt: Lang.
- Kern, Friederike. 2011. Der Erwerb kommunikativer Praktiken und Formen – Am Beispiel des Erzählens und Erklärens. In Stephan Habscheid (Hrsg.), *Textsorten, Handlungsmuster, Oberflächen: Linguistische Typologien der Kommunikation* (de Gruyter Lexikon), 231–254. Berlin: De Gruyter.
- Kern, Friederike, Ulrich Boden, Anne Nemeth, Sofia Koutalidis, Olga Abramov, Stefan Kopp & Katharina J. Rohlfing. 2024. Preschool children's discourse competence in different genres and how it relates to iconic gestures. *Journal of child language*. 10.1017/S030500092300065X.
- Kinalzik, Noelle. 2024. *Mündliches Erklären von der Grundschule bis zur Sekundarstufe – Varianzen zwischen Kindern heterogener Ausgangslagen*. Göttingen: Verlag für Gesprächsforschung.
- Klatt, Marie. 2024. *Multimodaler Spracherwerb: Die Entwicklung der kooperativen Nichtzustimmung bei deutschen und französischen Kindern*. Göttingen: Verlag für Gesprächsforschung.
- Klein, Josef. 2009. ERKLÄREN-WAS, ERKLÄREN-WIE, ERKLÄREN-WARUM: Typologie und Komplexität zentraler Akte der Welterschließung. In Rüdiger Vogt (Hrsg.), *Erklären: Gesprächsanalytische und fachdidaktische Perspektiven* (Stauffenburg Linguistik Band 52), 25–36. Tübingen: Stauffenburg Verlag.
- Klein, Josef. 2017. Beginnende ARGUMENTIER- und ERKLÄR-Kompetenz im Vorschulalter (1;9 - 3;0): Vorformen und Topik. In Iris Meißner & Eva L. Wyss (Hrsg.), *Begründen - Erklären - Argumentieren: Konzepte und Modellierungen in der Angewandten Linguistik* (Stauffenburg Linguistik 93), 65–88. Tübingen: Stauffenburg Verlag.
- Klein, Wolfgang. 1980. Argumentation und Argument. *Zeitschrift für Literaturwissenschaft und Linguistik*. 9–57.
- Kline, Susan L. 1998. Influence Opportunities and the Development of Argumentation Competencies in Childhood. *Argumentation* 12(3). 10.1023/A:1007780507894.
- Knoblauch, Hubert. 1995. *Kommunikationskultur: Die kommunikative Konstruktion kultureller Kontexte*. Berlin & New York: De Gruyter.
- Knoblauch, Hubert. 2021. Von kommunikativen Gattungen zu kommunikativen Formen: Konsequenzen des kommunikativen Konstruktivismus. In Beate Weidner, Katharina König, Wolfgang Imo & Lars Wegner (Hrsg.), *Verfestigungen in der Interaktion: Konstruktionen, sequenzielle Muster, kommunikative Gattungen*, 19–38. Berlin: De Gruyter.

- Komor, Anna. 2010. *Miteinander kommunizieren - Kinder unter sich: Eine empirische diskursanalytische Untersuchung zur Ausbildung kindlicher Kommunikationsfähigkeit* (Sprach-Vermittlungen 8). Münster, München & Berlin: Waxmann.
- König, Katharina. 2017. Question tags als Diskursmarker? – Ansätze zu einer systematischen Beschreibung von ne im gesprochenen Deutsch. In Hardarik Blühdorn, Arnulf Deppermann, Henrike Helmer & Thomas Spranz-Fogasy (Hrsg.), *Diskursmarker im Deutschen: Reflexionen und Analysen*, 233–258. Göttingen: Verlag für Gesprächsforschung.
- König, Katharina. 2020. Prosodie und epistemic stance: Konstruktionen mit finalem oder. In Wolfgang Imo & Jens P. Lanwer (Hrsg.), *Prosodie und Konstruktionsgrammatik* (Empirische Linguistik / Empirical Linguistics 12), 167–200. Berlin & Boston: De Gruyter.
- Koshik, Irene. 2002. A conversation analytic study of yes/no questions which convey reversed polarity assertions. *Journal of Pragmatics* 34(12). 10.1016/S0378-2166(02)00057-7.
- Koshik, Irene. 2003. Wh-questions used as challenges. *Discourse Studies* 5(1). 10.1177/14614456030050010301.
- Kotthoff, Helga. 1993. Disagreement and concession in disputes: On the context sensitivity of preference structures. *Language in Society* 22(2). 193–216.
- Kraft, Barbara & Christina Lanzen. 1995. Zur Entwicklung des Widersprechens bei Vorschulkindern: Formen der Etablierung von Opposition. *Der Deutschunterricht* 47(1). 94–99.
- Krah, Antje, Uta Quasthoff, Vivien Heller, Elke Wild, Jelena Hollmann & Nantje Otterpohl. 2013. Die Rolle der Familie beim Erwerb komplexer sprachlicher Fähigkeiten in der Sekundarstufe I. In Angelika Redder & Sabine Weinert (Hrsg.), *Sprachförderung und Sprachdiagnostik: Interdisziplinäre Perspektiven*, 68–88. Münster & New York: Waxmann.
- Krelle, Michael. 2014. *Mündliches Argumentieren in leistungsorientierter Perspektive: Eine empirische Analyse von Unterrichtsdiskussionen in der neunten Jahrgangsstufe*. Baltmannsweiler: Schneider-Verlag Hohengehren.
- Krelle, Michael. 2015. Leistungsprofile und -niveaus von Schülerinnen und Schülern in Unterrichtsdiskussionen. Ergebnisse einer gesprächslinguistischen Studie zum mündlichen Argumentieren am Ende der neunten Jahrgangsstufe. *Mitteilungen des Deutschen Germanistenverbandes* 62(1). 61–73.
- Kreuz, Judith. 2021. *Ko-konstruiertes Begründen unter Kindern: Eine gesprächsanalytische Studie von Kleingruppeninteraktionen in der Primarschule*. Tübingen: Stauffenburg Verlag.
- Kreuz, Judith & Martin Luginbühl. 2020. From flat propositions to deep co-constructed and modalized argumentations: Oral argumentative skills among elementary school children from grades 2 to 6. *Research on Children and Social Interaction* 4(1). 93–114.
- Kuhn, Deanna & Wadiya Udell. 2003. The Development of Argument Skills. *Child Development* 74(5). 1245–1260.
- Kultusministerkonferenz. 2004 i. d. F. von 2022. Bildungsstandards für das Fach Deutsch. Primarbereich.
- Kyrtziz, Amy, Tamara S. Ross & S. B. Koymen. 2010. Validating justifications in preschool girls' and boys' friendship group talk: implications for linguistic and socio-cognitive development. *Journal of child language* 37(1). 10.1017/S0305000908009069.
- Ladewig, Silva H. 2014a. 121. The cyclic gesture. In Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill & Jana Bressemer (Hrsg.), *Body - Language - Communication: An international handbook on multimodality in human interaction* (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 38/2), 1605–1618. Berlin & Boston: De Gruyter Mouton.
- Ladewig, Silva H. 2014b. Recurrent gestures. In Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill & Jana Bressemer (Hrsg.), *Body - Language - Communication: An international handbook on multimodality in human interaction* (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 38/2), 1558–1574. Berlin & Boston: De Gruyter Mouton.
- Ladewig, Silva H. & Lena Hotze. 2023. From action to performative gesture: the Slapping movement used by children at the age of four to six. *Semiotica* 2023(250). 10.1515/sem-2022-0033.

- Leßmann, Ann-Christin. 2021. Diskursive Anforderungen in der Grundschule. In Uta Quasthoff, Vivien Heller & Miriam Morek (Hrsg.), *Diskurserwerb in Familie, Peergroup und Unterricht* (Reihe Germanistische Linguistik 324), 279–302. Berlin & Boston: De Gruyter.
- Levy, Elena T. & David McNeill. 2013. Narrative development as symbol formation: Gestures, imagery and the emergence of cohesion. *Culture & Psychology* 19(4). 10.1177/1354067X13500328.
- Levy, Elena T. & David McNeill. 2015. *Narrative development in young children: Gesture, imagery, and cohesion*. Cambridge: Cambridge University Press.
- Li, Xiaoting. 2021. Multimodal practices for negative assessments as delicate matters: Incomplete syntax, facial expressions, and head movements. *Open Linguistics* 7(1). 10.1515/opli-2020-0164.
- Lucius-Hoene, Gabriele & Arnulf Deppermann. 2002. *Rekonstruktion narrativer Identität: Ein Arbeitsbuch zur Analyse narrativer Interviews*. Wiesbaden: VS Verlag für Sozialwissenschaften GmbH. <https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=6294886>.
- Luckmann, Thomas. 1986. Grundformen der gesellschaftlichen Vermittlung des Wissens: Kommunikative Gattungen. In Friedhelm Neidhardt (Hrsg.), *Kultur und Gesellschaft: René König, dem Begründer der Sonderhefte, zum 80. Geburtstag gewidmet* (Kölner Zeitschrift für Soziologie und Sozialpsychologie Sonderhefte 27), 191–211. Opladen: Westdeutscher Verlag.
- Marrese, Olivia H., Chase W. Raymond, Barbara A. Fox, Cecilia E. Ford & Megan Pielke. 2021. The Grammar of Obviousness: The Palm-Up Gesture in Argument Sequences. *Frontiers in Communication* 6. 10.3389/fcomm.2021.663067.
- Maury-Ruan, Claire. 2011. 'Voices' and Bodies: Investigating Nonverbal Parameters of the Participation Framework. In Gale Stam & Mika Ishino (Hrsg.), *Integrating gestures: The interdisciplinary nature of gesture* (Gesture studies v. 4), 309–319. Amsterdam & Philadelphia: John Benjamins Publishing Company.
- Maynard, Douglas W. 1985. How children start arguments. *Language in Society* 14(1). 1–29.
- Mazeland, Harrie. 2021. Die Aktivität als eigenständiger interaktionsorganisierender Bereich. In Beate Weidner, Katharina König, Wolfgang Imo & Lars Wegner (Hrsg.), *Verfestigungen in der Interaktion: Konstruktionen, sequenzielle Muster, kommunikative Gattungen*, 159–194. Berlin: De Gruyter.
- Mazur-Palandre, Audrey, Audrey Colletta & Kristine Lund. 2014. Context sensitive 'how' explanation in children's multimodal behavior. *Journal of Multimodal Communication Studies* 2. 1–17.
- Mazur-Palandre, Audrey & Kristine Lund. 2016. Explanatory content and visibility effects on the young child's verbal and gestural behavior in free dialogues. *Language, Interaction and Acquisition* 7(2). 10.1075/lia.7.2.02maz.
- McClave, Evelyn Z. 2000. Linguistic functions of head movements in the context of speech. *Journal of Pragmatics* 32(7). 10.1016/S0378-2166(99)00079-X.
- McNeill, David. 1985. So you think gestures are nonverbal? *Psychological Review* 92(3). 10.1037/0033-295X.92.3.350.
- McNeill, David. 1986. Iconic gestures of children and adults. *Semiotica* 62(1-2). 10.1515/semi.1986.62.1-2.107.
- McNeill, David. 1992. *Hand and mind: What gestures reveal about thought*. Chicago & London: The University of Chicago Press.
- McNeill, David. 2005. *Gesture and Thought*. Chicago: University of Chicago Press.
- McNeill, David & Elena T. Levy. 1982. Conceptual Representations in Language Activity and Gesture. In Robert J. Jarvella & Wolfgang Klein (Hrsg.), *Speech, place, and action: Studies in deixis and related topics*, 271–295. Chichester: Wiley.
- Mercer, Neil. 2009. Developing Argumentation: Lessons Learned in the Primary School. In Nathalie Muller Mirza & Anne-Nelly Perret-Clermont (Hrsg.), *Argumentation and Education: Theoretical Foundations and Practices*, 177–194. Boston: Springer US.
- Miller, Max. 1980. Ontogenese moralischer Argumentation. *Zeitschrift für Literaturwissenschaft und Linguistik* 39. 58–107.
- Miller, Max. 1984. Ontogenese des koordinierten Dissens. In Wolfgang Edelstein & Jürgen Habermas (Hrsg.), *Soziale Interaktion und soziales Verstehen: Beiträge zur Entwicklung der Interaktionskompetenz*, 220–250. Frankfurt am Main: Suhrkamp.

- Miller, Max & Wolfgang Klein. 1981. Moral argumentation among children. A case-study. *Linguistische Berichte* 74. 1–19.
- Mondada, Lorenza. 2006. Participants' online analysis and multimodal practices: projecting the end of the turn and the closing of the sequence. *Discourse Studies* 8(1). 10.1177/1461445606059561.
- Mondada, Lorenza. 2007. Multimodal resources for turn-taking: pointing and the emergence of possible next speakers. *Discourse Studies* 9(2). 10.1177/1461445607075346.
- Mondada, Lorenza. 2011. Understanding as an embodied, situated and sequential achievement in interaction. *Journal of Pragmatics* 43(2). 10.1016/j.pragma.2010.08.019.
- Mondada, Lorenza. 2013. Displaying, contesting and negotiating epistemic authority in social interaction: Descriptions and questions in guided visits. *Discourse Studies* 15(5). 10.1177/1461445613501577.
- Mondada, Lorenza. 2014a. Pointing, talk, and the bodies: Reference and joint attention as embodied interactional achievements. In Mandana Seyfeddinipur & Marianne Gullberg (Hrsg.), *From Gesture in Conversation to Visible Action as Utterance*, 95–124. Amsterdam: John Benjamins Publishing Company.
- Mondada, Lorenza. 2014b. The local constitution of multimodal resources for social interaction. *Journal of Pragmatics* 65(2). 10.1016/j.pragma.2014.04.004.
- Mondada, Lorenza. 2015a. Multimodal completions. In Arnulf Deppermann & Susanne Günthner (Hrsg.), *Temporality in interaction* (Studies in Language and Social Interaction 27), 267–307. Amsterdam & Philadelphia: John Benjamins Publishing Company.
- Mondada, Lorenza. 2015b. The facilitator's task of formulating citizens' proposals in political meetings: Orchestrating multiple embodied orientations to recipients. *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion* 16. 1–62.
- Mondada, Lorenza. 2016. Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics* 20(3). 10.1111/josl.1_12177.
- Mondada, Lorenza. 2018. Multiple Temporalities of Language and Body in Interaction: Challenges for Transcribing Multimodality. *Research on Language and Social Interaction* 51(1). 10.1080/08351813.2018.1413878.
- Mondada, Lorenza. 2021a. How Early can Embodied Responses be? Issues in Time and Sequentiality. *Discourse Processes* 58(4). 10.1080/0163853X.2020.1871561.
- Mondada, Lorenza. 2021b. *Sensing in social interaction: The taste for cheese in gourmet shops* (Learning in doing). Cambridge: Cambridge University Press.
- Mondada, Lorenza & Reinhold Schmitt (Hrsg.). 2010. *Situationseröffnungen: Zur multimodalen Herstellung fokussierter Interaktion* (Studien zur Deutschen Sprache 47). Tübingen: Gunter Narr Verlag.
- Mondada, Lorenza, Hanna Svensson & Nynke van Schepen. 2015. "Why that not now" : participants' orientations towards several organizational layers in social interaction. *Bulletin suisse de linguistique appliquée* 101. 10.5169/SEALS-978600.
- Mondada, Lorenza & Burak S. Tekin. 2020. Arranging bodies for photographs: Professional touch in the photography studio. *Social Interaction. Video-Based Studies of Human Sociality* 3(1). 10.7146/si.v3i1.120254.
- Morek, Miriam. 2012. *Kinder erklären: Interaktionen in Familie und Unterricht im Vergleich* (Stauffenburg Linguistik Bd. 60). Tübingen: Stauffenburg Verlag.
- Morek, Miriam. 2015. Dissensbearbeitung unter Gleichaltrigen – (k)ein Kontext für den Erwerb argumentativer Gesprächsfähigkeiten? In Birte Arendt, Vivien Heller & Antje Krah (Hrsg.), *Kinder argumentieren: Interaktive Erwerbskontexte und -mechanismen*, 34–46. Göttingen: V&R.
- Morek, Miriam. 2017. Was lernt man bei einer Fishbowl- Diskussion im Unterricht?: Zur Normativität und pragmatischen Situiertheit des Argumentierens. In Martin Luginbühl & Stefan Hauser (Hrsg.), *Gesprächskompetenz in schulischer Interaktion - normative Ansprüche und kommunikative Praktiken*, 65–90. Bern: hep.
- Morek, Miriam. 2018. Multimodalität als Ressource der Platzierung von Erzählungen in Peer-Interaktionen von Kindern. *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 19. 304–345.

- Morek, Miriam. 2020. Learning to modalize is learning to reason: On the role of epistemic modalizations in parent–child-talk and in written argumentation of secondary school students. *Research on Children and Social Interaction* 4(1). 115–141.
- Morek, Miriam. 2021. Familien- und Peer-Interaktionen als Erwerbsressource für Diskurskompetenzen: Empirische Befunde zu Variabilität und Erwerbspotenzial außerschulischer Diskurspraktiken von Präadoleszenten. In Uta Quasthoff, Vivien Heller & Miriam Morek (Hrsg.), *Diskurserwerb in Familie, Peergroup und Unterricht* (Reihe Germanistische Linguistik 324), 185–240. Berlin & Boston: De Gruyter.
- Morek, Miriam & Vivien Heller. 2012. Bildungssprache – Kommunikative, epistemische, soziale und interaktive Aspekte ihres Gebrauchs. *Zeitschrift für Angewandte Linguistik* 57(1). 10.1515/zfal-2012-0011.
- Morek, Miriam & Vivien Heller. 2021. Individualisierter Zuschnitt diskursiver Anforderung und Unterstützung: Finetuning diskurserwerbsförderlichen Lehrerhandelns in der Unterrichtsinteraktion. In Uta Quasthoff, Vivien Heller & Miriam Morek (Hrsg.), *Diskurserwerb in Familie, Peergroup und Unterricht* (Reihe Germanistische Linguistik 324), 381–424. Berlin & Boston: De Gruyter.
- Morek, Miriam, Vivien Heller & Uta Quasthoff. 2017. Erklären und Argumentieren: Modellierungen und empirische Befunde zu Strukturen und Varianzen. In Iris Meißner & Eva L. Wyss (Hrsg.), *Begründen - Erklären - Argumentieren: Konzepte und Modellierungen in der Angewandten Linguistik* (Stauffenburg Linguistik 93), 11–45. Tübingen: Stauffenburg Verlag.
- Morek, Miriam & Uta Quasthoff. 2018. Sprachliche und diskursive Praktiken unter Kindern. In Peter Schlobinski & Eva Neuland (Hrsg.), *Handbuch Sprache in sozialen Gruppen*, 255–275. Berlin: De Gruyter.
- Morgenstern, Aliyah. 2014. 140. The blossoming of children's multimodal skills from 1 to 4 years old. In Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill & Jana Bressemer (Hrsg.), *Body - Language - Communication: An international handbook on multimodality in human interaction* (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 38/2), 1848–1857. Berlin & Boston: De Gruyter Mouton.
- Morgenstern, Aliyah & Susan Goldin-Meadow (Hrsg.). 2022. *Gesture in language: Development across the lifespan* (Language and the human lifespan series). Washington & Berlin: American Psychological Association; De Gruyter.
- Morris, Desmond. 1994. *Bodytalk: The meaning of human gestures*. New York: Crown Trade Paperbacks.
- Müller, Cornelia. 1998. *Redebegleitende Gesten: Kulturgeschichte, Theorie, Sprachvergleich* (Körper, Zeichen, Kultur 1). Berlin: Berlin-Verlag Spitz.
- Müller, Cornelia. 2004. Forms and uses of the Palm Up Open Hand: A case of a gesture family? In Cornelia Müller & Roland Posner (Hrsg.), *The semantics and pragmatics of everyday gestures: Proceedings of the Berlin conference April 1998* (Körper, Zeichen, Kultur 9), 233–256. Berlin: Weidler Buchverlag Berlin.
- Müller, Cornelia. 2017. How recurrent gestures mean: Conventionalized contexts-of-use and embodied motivation. *Gesture* 16(2). 10.1075/gest.16.2.05mul.
- Müller, Cornelia & Ulrike Bohle. 2007. Das Fundament fokussierter Interaktion: Zur Vorbereitung und Herstellung von Interaktionsräumen durch körperliche Koordination. In Reinhold Schmitt (Hrsg.), *Koordination: Analysen zur multimodalen Interaktion* (Studien zur Deutschen Sprache 38), 129–165. Tübingen: Gunter Narr Verlag.
- Mundwiler, Vera, Judith Kreuz, Stefan Hauser, Brigit Eriksson & Martin Luginbühl. 2017. Mündliches Argumentieren als kommunikative Praktik: Schulbuchübungen und empirische Befunde im Vergleich. In Martin Luginbühl & Stefan Hauser (Hrsg.), *Gesprächskompetenz in schulischer Interaktion - normative Ansprüche und kommunikative Praktiken*, 91–123. Bern: hep.
- Mundwiler, Vera, Judith Kreuz, Daniel Müller-Feldmeth, Martin Luginbühl & Stefan Hauser. 2019. Quantitative und qualitative Zugänge in der Gesprächsforschung: Methodologische Betrachtungen am Beispiel einer Studie zu argumentativen Gruppendiskussionen. *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion* 20. 323–383.

- Ningelgen, Jana & Peter Auer. 2017. Is there a multimodal construction based on non-deictic so in German? *Linguistics Vanguard* 3(1). 10.1515/lingvan-2016-0051.
- Ohlhus, Sören. 2014. *Erzählen als Prozess: Interaktive Organisation und narrative Verfahren in mündlichen Erzählungen von Grundschulkindern* (Stauffenburg Linguistik 79). Tübingen: Stauffenburg Verlag.
- Ohlhus, Sören. 2016. Narrative Verfahren und mediale Bedingungen. Überlegungen zur Analyse mündlicher und schriftlicher Erzählungen. In Ulrike Behrens & Olaf Gätje (Hrsg.), *Mündliches und schriftliches Handeln im Deutschunterricht: Wie Themen entfaltet werden* (Positionen der Deutschdidaktik 3), 39–65. Berlin: Peter Lang.
- Ohlhus, Sören & Juliane Stude. 2009. Erzählen im Unterricht der Grundschule. In Michael Becker-Mrotzek (Hrsg.), *Mündliche Kommunikation und Gesprächsdidaktik*, 471–486. Baltmannsweiler: Schneider-Verlag Hohengehren.
- Oloff, Florence. 2013. Embodied withdrawal after overlap resolution. *Journal of Pragmatics* 46(1). 10.1016/j.pragma.2012.07.005.
- Oloff, Florence. 2017. Genau als redebeitragsinterne, responsive, sequenzschließende oder sequenzstrukturierende Bestätigungspartikel im Gespräch. In Hardarik Blühdorn, Arnulf Deppermann, Henrike Helmer & Thomas Spranz-Fogasy (Hrsg.), *Diskursmarker im Deutschen: Reflexionen und Analysen*, 207–232. Göttingen: Verlag für Gesprächsforschung.
- Payrató, Lluís & Sedinha Teßendorf. 2014. 116. Pragmatic gestures. In Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill & Jana Bressemer (Hrsg.), *Body - Language - Communication: An international handbook on multimodality in human interaction* (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 38/2), 1531–1539. Berlin & Boston: De Gruyter Mouton.
- Pekarek Doehler, Simona. 2016. More than an epistemic hedge: French je sais pas ‘I don’t know’ as a resource for the sequential organization of turns and actions. *Journal of Pragmatics* 106. 10.1016/j.pragma.2016.06.014.
- Pekarek Doehler, Simona. 2019. At the Interface of Grammar and the Body: Chais pas (“dunno”) as a Resource for Dealing with Lack of Recipient Response. *Research on Language and Social Interaction* 52(4). 10.1080/08351813.2019.1657276.
- Pekarek Doehler, Simona, Leelo Keevallik & Xiaoting Li. 2022. Editorial: The Grammar-Body Interface in Social Interaction. *Frontiers in psychology* 13. 10.3389/fpsyg.2022.875696.
- Pekarek Doehler, Simona, Hilla Polak-Yitzhaki, Xiaoting Li, Ioana M. Stoenica, Martin Havlík & Leelo Keevallik. 2021. Multimodal Assemblies for Prefacing a Dispreferred Response: A Cross-Linguistic Analysis. *Frontiers in psychology* 12. 10.3389/fpsyg.2021.689275.
- Perelman, Chaim & Lucie Olbrechts-Tyteca. 1958/2004. *Die Neue Rhetorik: Eine Abhandlung über das Argumentieren*. Stuttgart: Frommann-holzboog.
- Pfänder, Stefan. 2016. Fragmented adverbial clauses as resources for negotiating alignment: concessive repair patterns in French talk-in-interaction. *Language sciences: a world journal of the sciences of languages* 58. 95–110.
- Pomerantz, Anita. 1984. Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shaped. In j. M. Atkinson & John Heritage (Hrsg.), *Structures of Social Action*, 57–101. Cambridge: Cambridge University Press.
- Pomerantz, Anita. 1986. Extreme case formulations: A way of legitimizing claims. *Human Studies* 9(2-3). 219–229.
- Pontecorvo, Clotilde & Alessandra Fasulo. 1997. Learning to Argue in Family Shared Discourse: The Reconstruction of Past Events. In Lauren B. Resnick (Hrsg.), *Discourse, Tools and Reasoning: Essays on Situated Cognition* (Nato asi Subseries F 160), 406–442. Berlin & Heidelberg: Springer Berlin / Heidelberg.
- Prediger, Susanne, Kirstin Erath, Uta Quasthoff, Vivien Heller & Anna Vogler. 2016. Befähigung zur Teilhabe an Unterrichtsdiskursen: Die Rolle von Diskurskompetenz. In Jürgen Menthe, Dietmar

- Höttecke, Thomas Zabka, Marcus Hammann & Martin Rothgangel (Hrsg.), *Befähigung zu gesellschaftlicher Teilhabe: Beiträge der fachdidaktischen Forschung* (Waxmann-E-Books Schule und Unterricht 10), 285–300. Münster & New York: Waxmann.
- Quasthoff, Uta. 2009. Entwicklung der mündlichen Kommunikationskompetenz. In Michael Becker-Mrotzek (Hrsg.), *Mündliche Kommunikation und Gesprächsdidaktik*, 84–100. Baltmannsweiler: Schneider-Verlag Hohengehren.
- Quasthoff, Uta. 2015. Ko-Konstruktion in Erwachsenen-Kind Interaktionen: membership und der Erwerb von sprachlicher Kompetenz. In Ulrich Dausendschön-Gay, Elisabeth Gülich & Ulrich Krafft (Hrsg.), *Ko-Konstruktion in der Interaktion: Die gemeinsame Arbeit an Äußerungen und anderen sozialen Ereignissen*, 287–312. Bielefeld: transcript.
- Quasthoff, Uta. 2021. Methodische Überlegungen zur Datenbasis in der Interaktionalen Diskursanalyse: Grundlegendes zur Erhebung von Stichproben und Korpora. In Uta Quasthoff, Vivien Heller & Miriam Morek (Hrsg.), *Diskurserwerb in Familie, Peergroup und Unterricht* (Reihe Germanistische Linguistik 324), 43–76. Berlin & Boston: De Gruyter.
- Quasthoff, Uta, Vivien Heller & Miriam Morek. 2017. On the sequential organization and genre-orientation of discourse units in interaction: An analytic framework. *Discourse Studies* 19(1). 10.1177/1461445616683596.
- Quasthoff, Uta, Vivien Heller & Miriam Morek. 2021a. Diskurskompetenz und diskursive Partizipation als Schlüssel zur Teilhabe an Bildungsprozessen: Grundlegende Konzepte und Untersuchungslinien. In Uta Quasthoff, Vivien Heller & Miriam Morek (Hrsg.), *Diskurserwerb in Familie, Peergroup und Unterricht* (Reihe Germanistische Linguistik 324), 13–34. Berlin & Boston: De Gruyter.
- Quasthoff, Uta, Vivien Heller & Miriam Morek. 2021b. Glossar - Diskurskompetenz und diskursive Partizipation. In Uta Quasthoff, Vivien Heller & Miriam Morek (Hrsg.), *Diskurserwerb in Familie, Peergroup und Unterricht* (Reihe Germanistische Linguistik 324), 35–41. Berlin & Boston: De Gruyter.
- Quasthoff, Uta, Vivien Heller, Susanne Prediger & Kirstin Erath. 2022. Learning in and through classroom interaction: On the convergence of language and content learning opportunities in subject-matter learning. *European Journal of Applied Linguistics* 10(1). 10.1515/eujal-2020-0015.
- Quasthoff, Uta & Friederike Kern. 2007. Familiäre Interaktionsmuster und Diskursfähigkeit: Mögliche Auswirkungen interaktiver Stile auf diskursive Praktiken und Kompetenzen bei Schulkindern. In Heiko Hausendorf (Hrsg.), *Gespräch als Prozess: Linguistische Aspekte der Zeitlichkeit verbaler Interaktion*, 277–305. Tübingen: Gunter Narr Verlag.
- Quasthoff, Uta & Christian Kluger. 2021. Familiäre Interaktionsmuster als Erwerbsressource im längsschnittlichen Verlauf: Wie nutzen Kinder interaktive Ressourcen zum Ausbau mündlicher Argumentationskompetenz? In Uta Quasthoff, Vivien Heller & Miriam Morek (Hrsg.), *Diskurserwerb in Familie, Peergroup und Unterricht* (Reihe Germanistische Linguistik 324), 107–156. Berlin & Boston: De Gruyter.
- Quasthoff, Uta & Antje Krah. 2012. Familiäre Kommunikation als Spracherwerbsressource: Das Beispiel argumentativer Kompetenzen. In Eva Neuland (Hrsg.), *Sprache der Generationen* (Thema Deutsch 12), 115–132. Mannheim: Dudenverlag.
- Quasthoff, Uta, Elke Wild, Madeleine Domenech, Jelena Hollmann, Christian Kluger, Antje Krah & Nantje Otterpohl. 2021c. Familiäre Ressourcen für den Erwerb von Argumentationskompetenz: Quantitative Befunde zur sprachlich bedingten Bildungsbenachteiligung. In Uta Quasthoff, Vivien Heller & Miriam Morek (Hrsg.), *Diskurserwerb in Familie, Peergroup und Unterricht* (Reihe Germanistische Linguistik 324), 79–105. Berlin & Boston: De Gruyter.
- Quasthoff, Uta, Elke Wild, Madeleine Domenech, Jelena Hollmann, Christian Kluger, Antje Krah, Nantje Otterpohl & Johannes Wibbing. 2015. Mündliches und schriftliches Argumentieren in der Sekundarstufe I: Können Eltern ihre Kinder beim Erwerb dieser Schlüsselkompetenz unterstützen? In Angelika Redder, Johannes Naumann & Rosemarie Tracy (Hrsg.), *Forschungsinitiative Sprachdiagnostik und Sprachförderung - Ergebnisse*, 207–234. Münster: Waxmann.
- Quasthoff, Uta M. 1980. *Erzählen in Gesprächen: linguistische Untersuchungen zu Strukturen und Funktionen am Beispiel einer Kommunikationsform des Alltags*. Tübingen: Narr.

- Quasthoff, Uta M., Friederike Kern, Sören Ohlhus & Juliane Stude. 2019. *Diskurse und Texte von Kindern: Praktiken - Fähigkeiten - Ressourcen: Erwerb*. Dortmund: Universitätsbibliothek Dortmund.
- Raymond, Geoffrey. 2010. Grammar and Social Relations. In Alice F. Freed & Susan Ehrlich (Hrsg.), *"Why do you ask?": The function of questions in institutional discourse*, 87–107. New York: Oxford University Press.
- Reig Alamillo, Asela, Jean-Marc Colletta & Michèle Guidetti. 2012. Gesture and language in narratives and explanations: the effects of age and communicative activity on late multimodal discourse development. *Journal of child language* 40(3). 10.1017/S0305000912000062.
- Reineke, Silke. 2016. *Wissenszuschreibung in der Interaktion: Eine gesprächsanalytische Untersuchung impliziter und expliziter Formen der Zuschreibung von Wissen* (OraLingua 12). Heidelberg: Universitätsverlag Winter.
- Reineke, Silke. 2018. Interaktionale Analysen kognitiver Phänomene. In Konstanze Marx & Simon Meier (Hrsg.), *Sprachliches Handeln und Kognition: Theoretische Grundlagen und empirische Analysen* (Linguistik - Impulse & Tendenzen 75), 183–204. Berlin & Boston: De Gruyter.
- Rezat, Sara. 2007. *Die Konzession als strategisches Sprachspiel* (Sprache - Literatur und Geschichte. Studien zur Linguistik /Germanistik 30). Heidelberg: Universitätsverlag Winter.
- Rohlfing, Katharina J. 2008. Language Acquisition: A Multimodal Avenue. In Stephan Gramley & Vivian Gramley (Hrsg.), *Bielefeld introduction to applied linguistics: A course book* (Bielefeld English and American studies 1). Bielefeld: Aisthesis-Verlag.
- Rönqvist, Sara & Jan Lindström. 2021. Turn Continuations and Gesture: “And Then”-Prefacing in Multi-Party Conversations. *Frontiers in Communication* 6. 10.3389/fcomm.2021.670173.
- Rossano, Federico. 2012. *Gaze behavior in face-to-face interaction*. Nijmegen: Radboud University Nijmegen.
- Ruusuvuori, Johanna. 2013. Emotion, Affect and Conversation. In Jack Sidnell & Tanya Stivers (Hrsg.), *The Handbook of Conversation Analysis*, 330–349. Chichester: Wiley-Blackwell.
- Ruusuvuori, Johanna & Anssi Peräkylä. 2009. Facial and Verbal Expressions in Assessing Stories and Topics. *Research on Language & Social Interaction* 42(4). 10.1080/08351810903296499.
- Sacks, Harvey. 1987. On the Preferences for Agreement and Contiguity in Sequences in Conversation. In Graham Button & John R. E. Lee (Hrsg.), *Talk and Social Organisation*, 54–69. Philadelphia: Multilingual Matters.
- Sacks, Harvey, Emanuel A. Schegloff & Gail Jefferson. 1974. A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language* 50(4). 10.2307/412243.
- Scarvaglieri, Claudio. 2017. “I a hundred per cent agree” – Konsensuelles Argumentieren in der interkulturellen Wissenschaftskommunikation. In Kristin Bührig & Stephan Schlickau (Hrsg.), *Argumentieren und Diskutieren* (Hildesheimer Schriften zur Interkulturellen Kommunikation 8), 145–176. Frankfurt am Main & New York: Peter Lang.
- Schefflen, Albert E. 1964. The significance of posture in communication systems. *Psychiatry: interpersonal and biological processes* 27(4). 316–331.
- Schegloff, Emanuel A. 1968. Sequencing in Conversational Openings. *American Anthropologist* 70(6). 1075–1095. 10.1525/aa.1968.70.6.02a00030.
- Schegloff, Emanuel A. 1984. On some gestures' relation to talk. In J. M. Atkinson & John Heritage (Hrsg.), *Structures of Social Action: Studies in Conversation Analysis* (Studies in emotion and social interaction), 266–296. Cambridge: Cambridge University Press.
- Schegloff, Emanuel A. 1996. Confirming Allusions: Toward an Empirical Account of Action. *American Journal of Sociology* 102(1). 10.1086/230911.
- Schegloff, Emanuel A. 1997. Practices and actions: Boundary cases of other-initiated repair. *Discourse Processes* 23(3). 10.1080/01638539709545001.
- Schegloff, Emanuel A. 1998. Body Torque. *Social Research* 65(3). 535–596.
- Schegloff, Emanuel A. 2007. *Sequence organization in interaction: A primer in conversation analysis I* (1). Cambridge: Cambridge University Press.

- Schegloff, Emanuel A. & Harvey Sacks. 1973. Opening up Closings. *Semiotica* 8(4). 10.1515/semi.1973.8.4.289.
- Schmitt, Reinhold. 2005. Zur multimodalen Struktur von turn-taking. *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion* 6. 17–61.
- Schmitt, Reinhold (Hrsg.). 2007. *Koordination: Analysen zur multimodalen Interaktion* (Studien zur Deutschen Sprache 38). Tübingen: Gunter Narr Verlag.
- Schmitt, Reinhold & Arnulf Deppermann. 2007. Monitoring und Koordination als Voraussetzungen der multimodalen Konstitution von Interaktionsräumen. In Reinhold Schmitt (Hrsg.), *Koordination: Analysen zur multimodalen Interaktion* (Studien zur Deutschen Sprache 38), 95–128. Tübingen: Gunter Narr Verlag.
- Schmitt, Reinhold & Heiko Hausendorf. 2016. Sprache und Raum: Eine neue Forschungsperspektive, ihre Ursprünge und ihr aktueller Entwicklungsstand. In Heiko Hausendorf, Reinhold Schmitt & Wolfgang Kesselheim (Hrsg.), *Interaktionsarchitektur, Sozialtopographie und Interaktionsraum* (Studien zur Deutschen Sprache 72), 9–23. Tübingen: Narr Francke Attempto Verlag.
- Schmitt, Reinhold & Ralf Knöbl. 2013. Recipient design aus multimodaler Sicht. *Deutsche Sprache* 41(3). 242–276.
- Schönfelder, Nora. 2024. Multimodal Repetitions in Children's Co-construction of Arguments. *European Journal of Psychology and Education*.
- Schönfelder, Nora & Vivien Heller. 2019. Embodied reciprocity in conversational argumentation. Soliciting and giving reasons with Palm Up Open Hand gestures. *Proceedings of the 6th Gesture and Speech in Interaction GESPIN*. 81–86.
- Schoonjans, Steven. 2018. *Modalpartikeln als multimodale Konstruktionen: Eine korpusbasierte Kookkurrenzanalyse von Modalpartikeln und Gestik im Deutschen* (Empirische Linguistik / Empirical Linguistics 8). Berlin: De Gruyter.
- Selting, Margret. 2004. Listen: Sequenzielle und prosodische Struktur einer kommunikativen Praktik – eine Untersuchung im Rahmen der Interaktionalen Linguistik. *Zeitschrift für Sprachwissenschaft* 23(1). 10.1515/zfsw.23.1.1.
- Selting, Margret. 2012. Complaint stories and subsequent complaint stories with affect displays. *Journal of Pragmatics* 44(4). 10.1016/j.pragma.2012.01.005.
- Selting, Margret, Peter Auer, Birgit Barden, Jörg Bergmann, Elizabeth Couper-Kuhlen, Susanne Günthner, Uta Quasthoff, Christoph Meier, Peter Schlobinski & Susanne Uhmman. 1998. Gesprächsanalytisches Transkriptionssystem (GAT). *Linguistische Berichte* 173. 91–122.
- Selting, Margret, Peter Auer, Dagmar Barth-Weingarten, Jörg Bergmann, Pia Bergmann, Karin Birkner, Elizabeth Couper-Kuhlen, Arnulf Deppermann, Peter Gilles, Susanne Günthner, Martin Hartung, Friederike Kern, Christine Mertzluft, Christian Meyer, Miriam Morek, Frank Oberzaucher, Jörg Peters, Uta Quasthoff, Wilfried Schütte, Anja Stukenbrock & Susanne Uhmman. 2009. Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion* 10. 292–341.
- Sidnell, Jack. 2011. The epistemics of make-believe. In Jakob Steensig, Lorenza Mondada & Tanya Stivers (Hrsg.), *The Morality of Knowledge in Conversation*, 131–155. Cambridge: Cambridge University Press.
- Sikveland, Rein O. & Richard Ogden. 2012. Holding gestures across turns. *Gesture* 12(2). 10.1075/gest.12.2.03sik.
- Spiegel, Carmen. 1999. Argumentation von Jugendlichen im Deutschunterricht. Zwei Argumentationsformen. *Zeitschrift für Angewandte Linguistik*(30). 17–40.
- Spiegel, Carmen. 2003. „zum beispiel es gibt ja leute...“. Das Beispiel in der Argumentation Jugendlicher. In Arnulf Deppermann & Martin Hartung (Hrsg.), *Argumentieren in Gesprächen: Gesprächsanalytische Studien* (Stauffenburg Linguistik), 111–129. Tübingen: Stauffenburg Verlag.
- Spiegel, Carmen. 2006a. Argumentieren lernen im Unterricht - ein funktional-didaktischer Ansatz. In Elke Grundler & Rüdiger Vogt (Hrsg.), *Argumentieren in Schule und Hochschule: Interdisziplinäre Studien*, 63–76. Tübingen: Stauffenburg Verlag.

- Spiegel, Carmen. 2006b. *Unterricht als Interaktion: Gesprächsanalytische Studien zum kommunikativen Spannungsfeld zwischen Lehrern, Schülern und Institution*. Radolfzell: Verlag für Gesprächsforschung.
- Spranz-Fogasy, Thomas. 2003. Alles Argumentieren, oder was? Zur Konstitution von Argumentation in Gesprächen. In Arnulf Deppermann & Martin Hartung (Hrsg.), *Argumentieren in Gesprächen: Gesprächsanalytische Studien* (Stauffenburg Linguistik), 27–39. Tübingen: Stauffenburg Verlag.
- Spranz-Fogasy, Thomas. 2005a. Argumentation als Kommunikationsideologie. *Deutsche Sprache* 2(05). 141–156.
- Spranz-Fogasy, Thomas. 2005b. 'widersprechen'. *Zu Form und Funktion eines Aktivitätstyps in Schlichtungsgesprächen. Eine gesprächsanalytische Untersuchung*. Radolfzell: Verlag für Gesprächsforschung.
- Stefani, Elwys de & Lorenza Mondada. 2021. Chapter 10. A resource for action transition. In Emma Betz, Arnulf Deppermann, Lorenza Mondada & Marja-Leena Sorjonen (Hrsg.), *OKAY across languages: Toward a comparative approach to its use in talk-in-interaction* (Studies in Language and Social Interaction 34), 302–336. Amsterdam & Philadelphia: John Benjamins Publishing Company.
- Stevanovic, Melisa. 2012. Establishing joint decisions in a dyad. *Discourse Studies* 14(6). 10.1177/1461445612456654.
- Stevanovic, Melisa. 2013. Constructing a proposal as a thought: A way to manage problems in the initiation of joint decision-making in finnish workplace interaction. *Pragmatics (Quarterly Publication of the International Pragmatics Association (IPrA))* 23(3). 10.1075/prag.23.3.07ste.
- Stevanovic, Melisa & Anssi Peräkylä. 2012. Deontic Authority in Interaction: The Right to Announce, Propose, and Decide. *Research on Language & Social Interaction* 45(3). 10.1080/08351813.2012.699260.
- Stivers, Tanya. 2008. Stance, Alignment, and Affiliation During Storytelling: When Nodding Is a Token of Affiliation. *Research on Language & Social Interaction* 41(1). 10.1080/08351810701691123.
- Stivers, Tanya. 2015. Coding Social Interaction: A Heretical Approach in Conversation Analysis? *Research on Language & Social Interaction* 48(1). 10.1080/08351813.2015.993837.
- Stivers, Tanya, Lorenza Mondada & Jakob Steensig. 2011. Knowledge, morality and affiliation in social interaction. In Tanya Stivers, Lorenza Mondada & Jakob Steensig (Hrsg.), *The morality of knowledge in conversation* (Studies in interactional sociolinguistics 29), 3–24. Cambridge & New York: Cambridge University Press.
- Stivers, Tanya & Federico Rossano. 2012. Mobilising response in interaction: a compositional view of questions. In Jan-Peter de Ruiter (Hrsg.), *Questions: Formal, functional and interactional perspectives* (Language culture and cognition 12), 58–80. Cambridge: Cambridge University Press.
- Stivers, Tanya & Jack Sidnell. 2005. Introduction: Multimodal interaction. *Semiotica* 1(156). <https://doi.org/10.1515/semi.2005.2005.156.1>.
- Stivers, Tanya & Jack Sidnell. 2016. Proposals for Activity Collaboration. *Research on Language and Social Interaction* 49(2). 10.1080/08351813.2016.1164409.
- Streeck, Jürgen. 1993. Gesture as communication I: Its coordination with gaze and speech. *Communication Monographs* 60(4). 10.1080/03637759309376314.
- Streeck, Jürgen. 2006. Gestures: Pragmatic Aspects. In Keith Brown (Hrsg.), *Encyclopedia of Language and Linguistics*, 71–76. Burlington: Elsevier.
- Streeck, Jürgen. 2007. Geste und verstreichende Zeit: Innehalten und Bedeutungswandel der "bietenden Hand". In Heiko Hausendorf (Hrsg.), *Gespräch als Prozess: Linguistische Aspekte der Zeitlichkeit verbaler Interaktion*, 157–180. Tübingen: Gunter Narr Verlag.
- Streeck, Jürgen. 2009a. Forward-Gesturing. *Discourse Processes* 46(2-3). 10.1080/01638530902728793.
- Streeck, Jürgen. 2009b. *Gesturecraft: Manufacturing understanding* (Gesture studies 2). Amsterdam: J. Benjamins Pub. Co.
- Streeck, Jürgen. 2010. Nick Enfield, The Anatomy of Meaning: Speech, Gesture, and Composite Utterances. *Pragmatics & Cognition* 18(2). 457–464.

- Streeck, Jürgen. 2013. 43. Praxeology of gesture. In Ellen Fricke, Alan Cienki, David McNeill, Silva Ladewig & Sedinha Tessendorf (Hrsg.), *Body - Language - Communication: An international Handbook on Multimodality in Human Interaction* (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science [HSK] 38/1), 674–688. Berlin & Boston: De Gruyter Mouton.
- Streeck, Jürgen. 2015. Embodiment in Human Communication. *Annual Review of Anthropology* 44(1). 10.1146/annurev-anthro-102214-014045.
- Streeck, Jürgen. 2016. Gestische Praxis und sprachliche Form. In Arnulf Deppermann, Helmuth Feilke & Angelika Linke (Hrsg.), *Sprachliche und kommunikative Praktiken*, 57–79. Berlin & Boston: De Gruyter.
- Streeck, Jürgen, Charles Goodwin & Curtis LeBaron (Hrsg.). 2011. *Embodied Interaction: Language and body in the material world*. Cambridge [u. a]: Cambridge University Press.
- Streeck, Jürgen & Ulrike Hartge. 1992. Previews: Gestures at the transition place. In Peter Auer & Aldo Di Luzio (Hrsg.), *The contextualization of language* (Pragmatics and beyond 22), 135–157. Amsterdam & Philadelphia: John Benjamins Publishing Company.
- Stude, Juliane. 2003. *Mündliche und schriftliche Spieleerklärungen von Grundschulkindern: Erste Analysen des OLDER-Korpus*. Dortmund.
- Stukenbrock, Anja. 2009. Herausforderungen der multimodalen Transkription: Methodische und theoretische Überlegungen aus der wissenschaftlichen Praxis. In Karin Birkner & Anja Stukenbrock (Hrsg.), *Die Arbeit mit Transkripten in Fortbildung, Lehre und Forschung*, 144–170. Mannheim: Verlag für Gesprächsforschung.
- Stukenbrock, Anja. 2015. *Deixis in der face-to-face-Interaktion* (Linguae & litterae 47). Berlin & München: De Gruyter.
- Stukenbrock, Anja. 2018. Forward-looking: Where do we go with multimodal projections? In Arnulf Deppermann & Jürgen Streeck (Hrsg.), *Time in embodied interaction: Synchronicity and sequentiality of multimodal resources* (Pragmatics & Beyond New Series 293), 31–68. Amsterdam & Philadelphia: John Benjamins Publishing Company.
- Sweetser, Eve & Marisa Sizemore. 2008. Personal and interpersonal gesture spaces: Functional contrasts in language and gesture. In Andrea Tyler, Yiyoun Kim & Mari Takada (Hrsg.), *Language in the Context of Use: Discourse and Cognitive Approaches to Language*, 25–51. Berlin & Boston: De Gruyter Mouton.
- Teßendorf, Sedinha. 2014. 117. Pragmatic and metaphoric - combining functional with cognitive approaches in the analysis of the "brushing aside gesture". In Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill & Jana Bressemer (Hrsg.), *Body - Language - Communication: An international handbook on multimodality in human interaction* (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science [HSK] 38/2), 1540–1558. Berlin & Boston: De Gruyter Mouton.
- Thurmair, Maria. 1989. *Modalpartikeln und ihre Kombinationen* (Linguistische Arbeiten 223). Tübingen: Max Niemeyer Verlag.
- Tiittula, Liisa. 2007. Blickorganisation in der side-by-side-Positionierung am Beispiel eines Geschäftsgesprächs. In Reinhold Schmitt (Hrsg.), *Koordination: Analysen zur multimodalen Interaktion* (Studien zur Deutschen Sprache 38), 225–261. Tübingen: Gunter Narr Verlag.
- Toulmin, Stephen E. 1958. *The Uses of Argument*. Cambridge: Cambridge University Press.
- Tracy, Karen & Jessica Robles. 2009. Questions, questioning, and institutional practices: an introduction. *Discourse Studies* 11(2). 10.1177/1461445608100941.
- Trautmann, Caroline. 2004. *Argumentieren: Funktional-pragmatische Analysen praktischer und wissenschaftlicher Diskurse* (Arbeiten zur Sprachanalyse 43). Frankfurt am Main & Berlin: Peter Lang.
- van Eemeren, Frans H. & Rob Grootendorst. 2004. *A systematic theory of argumentation: The pragma-dialectical approach*. New York: Cambridge University Press.
- van Eemeren, Frans H., Rob Grootendorst & Tjark Kruiger. 1987. *Handbook of argumentation theory: A critical survey of classical backgrounds and modern studies* (Studies of argumentation in pragmatics and discourse analysis 7). Dordrecht: Foris Publications.

- Vogt, Rüdiger. 2002. *Im Deutschunterricht diskutieren: Zur Linguistik und Didaktik einer kommunikativen Praktik* (Reihe Germanistische Linguistik 228). Berlin: De Gruyter.
- Völzing, Paul-Ludwig. 1982. *Kinder argumentieren: die Ontogenese argumentativer Fähigkeiten*. Paderborn & München: Schoeningh.
- Vygotsky, Lev S. (Hrsg.). 1978. *Mind in society: The development of higher psychological processes*. Cambridge: Harvard University Press.
- Wald, Benji. 1978. Zur Einheitlichkeit und Einleitung von Diskurseinheiten. In Uta M. Quasthoff (Hrsg.), *Sprachstruktur - Sozialstruktur: Zur linguistischen Theorienbildung* (Monographien Linguistik und Kommunikationswissenschaft 30), 128–149. Königstein: Scriptor-Verlag.
- Weinrich, Harald. 2003. *Textgrammatik der deutschen Sprache*. Hildesheim & Zürich: Olms.
- Weiser-Zurmühlen, Kristin. 2021. *Vergemeinschaftung und Distinktion: Eine gesprächsanalytische Studie über Positionierungspraktiken in Diskussionen über TV-Serien*. Berlin & Boston: De Gruyter.
- Whitehead, Kevin A. 2011. Some uses of head nods in “third position” in talk-in-interaction. *Gesture* 11(2). 10.1075/gest.11.2.01whi.
- Wild, Elke, Uta Quasthoff, Jelena Hollmann, Nantje Otterpohl, Antje Krah & Sören Ohlhus. 2012. Die Rolle familialer Unterstützung beim Erwerb von Argumentationskompetenz in der Sekundarstufe I. *Diskurs Kindheits- und Jugendforschung* 7(1). 101–112.
- Williams, Robert G. 2022. Coordinating and Sharing Gesture Spaces in Collaborative Reasoning. *Gesture* 21(1). 115–149. 10.1075/gest.21005.wil.
- Zadunaisky Ehrlich, Sara & Shoshana Blum-Kulka. 2014. “Now I said that Danny becomes Danny again”: a multifaceted view of kindergarten children’s peer argumentative discourse. In Asta Cekaite, Shoshana Blum-Kulka, Vibeke G. Aukrust & Eva Teubal (Hrsg.), *Children’s peer talk: Learning from each other*, 23–41. Cambridge: Cambridge University Press.
- Zifonun, Gisela, Ludger Hoffmann & Bruno Strecker. 1997. *Grammatik der deutschen Sprache, 1* (Schriften des Instituts für Deutsche Sprache 7/1). Berlin: De Gruyter.
- Zima, Elisabeth. 2017. Multimodale Mittel der Rederechtsaushandlung im gemeinsamen Erzählen in der Face-to-Face Interaktion. *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion* 18. 241–273.

14 Anhang

14.1 Abbildungsverzeichnis

Abbildung 1: Gestenphasierung nach Kendon (2004: 114)	29
Abbildung 2: Übersicht über McNeills (1992, 2005) Gestenklassifikationssystem (eigene Darstellung)	37
Abbildung 3: Übersicht über Gesten mit kommunikativer Funktion nach Bavelas et al. (1992, 1995) und Bavelas (1994) (eigene Darstellung)	39
Abbildung 4: Übersicht über die Funktionen von Gesten nach Kendon (2004, 2017) (eigene Darstellung)	40
Abbildung 5: Übersicht über die sechs Gestenökologien nach Streeck (2009) (eigene Darstellung)	42
Abbildung 6: Aufnahmesetting der Eltern-Kind-Dyaden	91
Abbildung 7: Struktur der F-Formation (adaptiert nach Kendon 1990).....	92
Abbildung 8: Unterteilung des Gestenraums bei sitzenden Personen aus McNeill (1992: 89).....	93

14.2 Tabellenverzeichnis

Tabelle 1: Zusammenfassende Übersicht über die Funktionen von Gesten unter Berücksichtigung der unterschiedlichen Klassifikationssysteme	46
Tabelle 2: Übersicht über performative Funktionen gestischer Ressourcen	48
Tabelle 3: Übersicht über modalisierende Funktionen von gestischen Ressourcen	49
Tabelle 4: Übersicht über strukturierende Funktionen gestischer Ressourcen	50
Tabelle 5: Übersicht über interaktionsorganisatorisch Funktionen von gestischen Ressourcen	51
Tabelle 6: Übersicht zu gängigen Gestenfamilien	52
Tabelle 7: Übersicht über das Datenkorpus	90
Tabelle 8: Beschreibung der multimodalen Gestalt „Hervorheben von Optionen im Beziehen von Positionen“	104
Tabelle 9: Beschreibung der multimodalen Gestalt „Position beziehen durch Ausschließen einer Option“	109
Tabelle 10: Beschreibung der multimodalen Gestalt „Präsentieren von (komplexeren) Begründungen“	113
Tabelle 11: Beschreibung der multimodalen Gestalt „Kurzzeitiges Präsentieren und Hervorheben relevanter Bedeutungselemente der Begründung“	118
Tabelle 12: Beschreibung der multimodalen Gestalt „Etablieren und Präsenthalten eines verhandelten Gesprächsgegenstandes“	122
Tabelle 13: Beschreibung der multimodalen Gestalt „Etablieren und Präsenthalten eines verhandelten Gesprächsgegenstandes bei konkurrierender Turnübernahme“	129
Tabelle 14: Beschreibung der multimodalen Gestalt „Referenzieren und Präsenthalten eines verhandelten Gesprächsgegenstandes als Verständendokumentation (rezipient*innenseitig)“	133
Tabelle 15: Beschreibung der multimodalen Gestalt „Multimodales Einführen einer Turnfortsetzung“	135
Tabelle 16: Beschreibung der multimodalen Gestalt „Multimodale Kohäsion und Kohärenz“	139
Tabelle 17: Beschreibung der multimodalen Gestalt „Verwerfen in Begründungen“	144
Tabelle 18: Beschreibung der multimodalen Gestalt „Nachdrückliches Zustimmung“	148
Tabelle 19: Beschreibung der multimodalen Gestalt „Schwaches Zustimmung in Konzessionen“	151
Tabelle 20: Beschreibung der multimodalen Gestalt „Schwaches Zustimmung zu evidenten Begründungen & Positionen“	153
Tabelle 21: Beschreibung der multimodalen Gestalt „Präsentieren während des Einwendens“	160
Tabelle 22: Beschreibung der multimodalen Gestalt „Sichtbares Explizieren von Negationen in Einwänden“	164
Tabelle 23: Beschreibung der multimodalen Gestalt „Anhalten einer vorangegangenen Opposition in Konzessionen“	167
Tabelle 24: Beschreibung der multimodalen Gestalt „Mahnendes Einwenden“	169
Tabelle 25: Beschreibung der multimodalen Gestalt „Einfordern von Rückmeldungen im Listenformat“	173
Tabelle 26: Beschreibung der multimodalen Gestalt „Präsenthalten des Gesprächsgegenstandes in eingeforderten Rückmeldungen“	177
Tabelle 27: Beschreibung der multimodalen Gestalt „Markieren von Evidenz durch Fallenlassen der Hand“	183
Tabelle 28: Beschreibung der multimodalen Gestalt „Markieren von Evidenz durch Palm Up Geste mit lateraler Bewegung“	184
Tabelle 29: Beschreibung der multimodalen Gestalt „Markieren von Evidenz durch Hand Shrug“	187
Tabelle 30: Beschreibung der multimodalen Gestalt „Markieren des ablehnenden Fallenlassens eines Gesprächsgegenstandes“	189
Tabelle 31: Beschreibung der multimodalen Gestalt „Anzeige von epistemischer Unsicherheit“	192
Tabelle 32: Beschreibung der multimodalen Gestalt „Markieren von Unbestimmtheit“	196
Tabelle 33: Übersicht über rekonstruierte multimodale Gestalten für das mündliche Argumentieren (inkl. Funktion und Beschreibung)	202
Tabelle 34: Übersicht über diskursstrukturelle Profile der Fokuskinder	208
Tabelle 35: Anzahl der Begründungen (Fokuskind I)	234
Tabelle 36: Funktionen und Anzahl multimodaler Gestalten zur Kontextualisierung (Fokuskind I)	235
Tabelle 37: Funktionen und Anzahl multimodaler Gestalten zur Vertextung (Fokuskind I)	240
Tabelle 38: Funktionen und Anzahl multimodaler Gestalten zur epistemischen Modalisierung (Fokuskind I)	249
Tabelle 39: Übersicht über Funktionen der eingesetzten multimodalen Gestalten für das Begründen (Fokuskind I)	257
Tabelle 40: Anzahl der Begründungen (Fokuskind II)	258

Tabelle 41: Funktionen und Anzahl multimodaler Gestalten zur Kontextualisierung (Fokuskind II)	258
Tabelle 42: Funktionen und Anzahl multimodaler Gestalten für das Vertexten (Fokuskind II)	263
Tabelle 43: Funktionen und Anzahl multimodaler Gestalten zur epistemischen Modalisierung (Fokuskind II)	270
Tabelle 44: Übersicht über Funktionen der eingesetzten multimodalen Gestalten für das Begründen (Fokuskind II)	275
Tabelle 45: Alters- und kindvergleichende Übersicht zu Funktionen der eingesetzten multimodalen Gestalten für das Begründen	276
Tabelle 46: Anzahl der Begründungen und multimodalen Gestalten im Vergleich (Fokuskind I & II)	280

14.3 Transkriptionskonventionen

Die videographierten Gesprächsdaten wurden nach dem Gesprächsanalytischen Transkriptionssystem 2 (GAT 2) nach Selting et al. (2009) transkribiert sowie um Aspekte der Multimodalität (vgl. Mondada 2018) und ihrer zeitlichen Entfaltung (vgl. Kendon 2004) erweitert, die mit Heller (2018a) unter dem Verbaltranskript angegeben werden. Die nachfolgende Übersicht über die verwendete Transkriptionskonvention wurde aus Selting et al. (2009: 330f.) entnommen und entsprechend um o. g. Aspekte ergänzt.

Sequenzielle Struktur/Verlaufsstruktur

[]	Überlappungen und Simultansprechen
[]	
=	schneller, unmittelbarer Anschluss neuer Sprecherbeiträge oder Segmente (latching)

Ein- und Ausatmen

°h / h°	Ein- bzw. Ausatmen von ca. 0.2-0.5 Sek. Dauer
°hh / hh°	Ein- bzw. Ausatmen von ca. 0.5-0.8 Sek. Dauer
°hhh / hhh°	Ein- bzw. Ausatmen von ca. 0.8-1.0 Sek. Dauer

Pausen

(.)	Mikropause, geschätzt, bis ca. 0.2 Sek. Dauer
(-)	kurze geschätzte Pause von ca. 0.2-0.5 Sek. Dauer
(--)	mittlere geschätzte Pause v. ca. 0.5-0.8 Sek. Dauer
(---)	längere geschätzte Pause von ca. 0.8-1.0 Sek. Dauer
(0.5)	gemessene Pausen von ca. 0.5 bzw. 2.0 Sek. Dauer
(2.0)	(Angabe mit einer Stelle hinter dem Punkt)

Sonstige segmentale Konventionen

und_äh	Verschleifungen innerhalb von Einheiten
äh öh äm	Verzögerungssignale, sog. "gefüllte Pausen"

Lachen und Weinen

haha hehe hihi	silbisches Lachen
((lacht)) ((weint))	Beschreibung des Lachens
<<lachend> >	Lachpartikeln in der Rede, mit Reichweite
<<:-)> soo>	"smile voice"

Rezeptionssignale

hm ja nein nee	einsilbige Signale
hm_hm ja_a	zweisilbige Signale
nei_ein nee_e	
?hm?hm,	mit Glottalverschlüssen, meistens verneinend

Sonstige Konventionen

((hustet)) para- und außersprachliche Handlungen u.
sprachbegleitende para- und außersprachliche

Ereignisse

<<hustend> > Handlungen und Ereignisse mit Reichweite
() unverständliche Passage ohne weitere Angaben
(xxx), (xxx xxx) ein bzw. zwei unverständliche Silben
(solche) vermuteter Wortlaut
(also/alo)

Sonstige segmentale Konventionen

: Dehnung, Längung, um ca. 0.2-0.5 Sek.
:: Dehnung, Längung, um ca. 0.5-0.8 Sek.
::: Dehnung, Längung, um ca. 0.8-1.0 Sek.
? Abbruch durch Glottalverschluss

Tonhöhenbewegung am Ende von Intonationsphrasen

? hoch steigend
, mittel steigend
– gleichbleibend
; mittel fallend
. tief fallend

Sonstige Konvention

<<erstaunt> > interpretierende Kommentare mit Reichweite

Akzentuierung

akZENT Fokusakzent
ak!ZENT! extra starker Akzent
akzEnt Nebenakzent

Auffällige Tonhöhen sprünge

↑ kleinere Tonhöhen sprünge nach oben
↓ kleinere Tonhöhen sprünge nach unten

Lautstärke- und Sprechgeschwindigkeitsveränderungen, mit Extension

<<f> > forte, laut
<<ff> > fortissimo, sehr laut
<<p> > piano, leise
<<pp> > pianissimo, sehr leise
<<all> > allegro, schnell
<<len> > lento, langsam
<<cresc> > crescendo, lauter werdend
<<dim> > diminuendo, leiser werdend
<<acc> > accelerando, schneller werdend

<<rall> > rallentando, langsamer werdend

Veränderung der Stimmqualität und Artikulationsweise

<<creaky> > glottalisiert, „Knarrstimme“

<<flüsternd> > Beispiel für Veränderung der Stimmqualität, wie angegeben

Körperlich-sichtbare Ressourcen

VA |ja | mit Schrägstrichen unter dem Verbaltranskript;

va-gs |öffnet rH Palm Up| Ressource wird mit Abkürzung angegeben
(gs = gesture; gz = gaze; bo = body etc.)

--->| Ressource wird bis Pfeilende gehalten

>Z. (XXX) Ressource wird bis Z. XXX gehalten

>> Ressourceneinsatz beginnt vor der fokussierten Sequenz

->> Ressourceneinsatz hält über die fokussierte Sequenz an

... Vorbereitung der Ressource bzw. des Ressourceneinsatzes

,, , Rückkehr der Ressource (in Ruheposition)

#1 Nummer des eingefügten Standbildes