

Blickpraktiken von SprecherInnen und AdressatInnen bei der Lokaldeixis: Mobile Eye Tracking-Analysen zur Herstellung von *joint attention*

Anja Stukenbrock

Abstract

Der folgende Beitrag untersucht unterschiedliche Blickpraktiken von SprecherInnen und AdressatInnen, die systematisch auftreten, wenn diese mittels deiktischer Referenzierungshandlungen einen gemeinsamen Aufmerksamkeitsfokus auf ein sichtbares Phänomen im Wahrnehmungsraum herstellen. Während eine wachsende Anzahl von Untersuchungen in der multimodalen Konversationsanalyse die zentrale Rolle gestischer Praktiken in deiktischen Referenzierungssequenzen nachgewiesen hat, ist das Blickverhalten bei der Deixis bislang kaum konversationsanalytisch untersucht worden. Die wenigen existierenden Studien basieren auf Videoaufnahmen, die keine präzise Analyse des Blickverhaltens gestatten. Demgegenüber gründen die Ergebnisse der vorliegenden Studie auf mobilen Eye Tracking-Aufnahmen von Beteiligten in natürlichen Alltagsinteraktionen. Der Beitrag erweitert und differenziert bisherige Ergebnisse zu Deixis, Gestik und Blick. Es wird argumentiert, dass deiktische Ausdrücke, Gesten und Blickpraktiken in einer spezifischen Weise sequenziell zusammenspielen: Der Blick konstituiert dabei ein Mittel interpersoneller Koordinierung, das über die regulatorische Funktion z.B. beim Turn-Taking hinausgeht. Innerhalb deiktischer Referenzierungssequenzen nehmen Blickpraktiken spezifische sequenzielle Positionen ein und tragen systematisch zur Intersubjektivierung eines gemeinsam wahrzunehmenden Phänomens bei.

Keywords: gemeinsame Aufmerksamkeitsorientierung, Blickpraktiken, Lokaldeixis, mobiles Eye Tracking, Konversationsanalyse.

English Abstract

The following paper explores different gaze practices that systematically occur when participants direct their interlocutor's attention to visible entities in the surroundings by means of deictic pointing. While a growing body of research in multimodal conversation analysis acknowledges the central role of gestures in acts of deictic reference, gaze has remained almost unexplored in studies on verbal and manual pointing. The few existing studies rely on video-recordings that do not allow a precise analysis of eye gaze. The results of the present study are based on mobile eye-tracking recordings undertaken with two pairs of eye tracking glasses worn by participants in non-laboratory, everyday interactions. The paper expands and refines previous observations on deixis, gestures and gaze. It claims that deictic expressions, gesture and eye gaze combine in a very specific way: Gaze constitutes a means of interpersonal coordination that goes beyond the regulatory functions at work, e.g. in turn taking. Within deictic referencing acts, eye gaze occupies specific slots or sequential positions that systematically contribute to, create and display a new status of intersubjectivity among the participants vis-à-vis the phenomenon of joint attention.

Keywords: joint attention, gaze practices, spatial deixis, mobile eye tracking, conversation analysis

1. Einleitung
2. Datenkorpus, methodologische Überlegungen und Herausforderungen
3. Deiktische Praktiken
 - 3.1. Was ist Deixis?
 - 3.2. Verbaldeixis erfordert Zeigegesten
 - 3.3. Verbaldeixis und Zeigegesten fordern den Adressatenblick ein
4. Analysen zum Blickverhalten in deiktischen Praktiken
 - 4.1. Der Gesten-Blick-Nexus in der *demonstratio ad oculos*
 - 4.2. Der Ziel-Blick-Nexus in der *demonstratio ad oculos*
 - 4.3. Der Blick-zu-Blick-Nexus in der *demonstratio ad oculos*
5. Diskussion: Was mobiles Eye Tracking für die Analyse von Blickpraktiken bei der *demonstratio ad oculos* in natürlicher Interaktion leistet
6. Literatur
7. Konventionen für die Transkription des Blickverhaltens

1. Einleitung

Seit einigen Jahrzehnten erbringen videoanalytische Untersuchungen in der Konversationsanalyse¹ wichtige Erkenntnisse über das Zusammenspiel verbaler und körperlich-visueller Ressourcen und verfeinern durch empirische Detailstudien die Konzeptualisierung verbaler Interaktion als leiblichem Gesamtgeschehen. Dabei wächst die Einsicht in die Schlüsselrolle, die das Blickverhalten der Beteiligten bei der on-line-Gestaltung sprachlicher Interaktion spielt.² Allerdings sind Videoaufzeichnungen häufig nicht präzise genug, um die Blickpraktiken der Interaktionsbeteiligten im Detail zu analysieren. Mobile Eye Tracking-Verfahren bieten seit einigen Jahren die Möglichkeit, außerhalb des Labors Untersuchungen zum Blickverhalten durchzuführen. Doch bis auf wenige Ausnahmen³ gibt es bislang kaum konversationsanalytisch fundierte Untersuchungen, die mit mobilem Eye Tracking arbeiten, um das Präzisionsdefizit in Videodaten zu überwinden.

Eye Tracking-Verfahren werden noch immer fast ausschließlich in psychologischen, psycholinguistischen und kognitionswissenschaftlichen Forschungsparadigmen verwendet, wo sie in stark kontrollierten Versuchsanordnungen in Laborsituationen zum Einsatz kommen, in denen sich die Beteiligten kaum bewegen. Dies gilt auch für Eye Tracking-Studien zur Referenz.⁴ Selbst bei denjenigen Studien, die mittels mobiler Eye Tracker das Blickverhalten bei der Referenz unter verschiedenen Bedingungen miteinander vergleichen, sind große Einschränkungen im Hinblick auf die ökologische Validität der Ergebnisse zu konstatieren. So wurden in Elizitierungsexperimenten Interaktionsdyaden unter Sitz- und Stehbedingungen aufgefordert, LEGO-Objekte als mögliche Replika moderner Kunstwerke zu diskutieren (Clark/Gergle 2010; Gergle/Clark 2011). Die Ergebnisse unterstreichen die zentrale Rolle des Blicks bei der Koordinierung visueller Aufmerksamkeit in Referenzierungshandlungen und weisen quantitative Unterschiede

¹ U.a. De Stefani 2010, 2013; De Stefani/Gazin 2014; Deppermann 2013, 2018; Goodwin 2000, 2003; Hausendorf/Mondada/Schmitt 2012; Heath 1986; Hindmarsh/Heath 2000; Mondada 2007, 2012a; Schmitt 2007; Streeck 2009; Streeck/Goodwin/LeBaron 2011; Stukenbrock 2009, 2014a, b, 2015.

² Argyle/Cook 1976; Goodwin 1980, 1981; Kendon 1990; Rossano 2012; Streeck 1993, 2002, 2014; Stukenbrock 2010, 2014a, 2015.

³ Auer im Druck; Holler/Kendrick 2015; Stukenbrock im Druck a, b; Stukenbrock/Dao im Druck; Weiss/Auer 2016.

⁴ Clark/Gergle 2010; Gergle/Clark 2011; Hanna/Brennan 2007; Land 2006.

zwischen sitzenden und stehenden (mobilen) Teilnehmern nach. Demnach verwenden stehende Dyaden einen höheren Anteil lokaler Deiktika zur Referenz als sitzende Teilnehmer und weisen zugleich ein geringeres Maß an Blicküberlappungen (*gaze overlap*) auf (Gergle/Clark 2011:442). Offen bleibt jedoch, welche unterschiedlichen Blickpraktiken SprecherInnen und AdressatInnen einsetzen, wie diese im lokalen Interaktionskontext temporal gestaltet und interpersonell koordiniert (Deppermann/Schmidt 2007) werden. Auch sind die Gesten der Beteiligten in der Untersuchung unberücksichtigt geblieben, obgleich sie eine unabdingbare Ressource multimodaler Referenz sind und das Blickverhalten entscheidend prägen (Stukenbrock 2009, 2015).

Im Unterschied zu artifiziellen Arrangements in Experimenten sind Referenzierungspraktiken im Alltag in Aktivitäten eingebettet, die in stationären, semi-stationären sowie in sich ständig verändernden, mobilen Situationen vollzogen werden. Dabei bewegen sich nicht nur die Beteiligten (vgl. Stukenbrock/Dao angenommen), sondern es können auch bewegliche und sich bewegende Objekte involviert sein, die wie die interpersonelle Koordinierung von verbalen und kinesischen Praktiken der Beteiligten in die Analyse des interaktiven Gesamtgeschehens, das sie zugleich reflektieren und konstituieren, einzubeziehen sind.⁵

Bisherige, auf experimentellen Daten basierende Untersuchungen konzentrieren sich auf vorab definierte, quantifizierbare Variablen und berücksichtigen nicht, dass multimodale Referenzierungsakte sich von Augenblick-zu-Augenblick interaktiv in der Zeit entfalten und dabei sowohl lokale, kontextspezifische als auch kontextübergreifende Ordnungsmuster erzeugen. Durch die eingeschränkte Untersuchungsperspektive und den Verzicht auf natürliche Daten werden die konstitutiven Bestandteile menschlicher Interaktion, allen voran ihre endogene, gemeinschaftliche Organisation, die Tatsache, dass sie von den Beteiligten lokal situiert und inkrementell in sequenzieller und zeitlicher Entfaltung hervorgebracht wird (Schegloff 1996:167; Mondada 2012b:33), außer Acht gelassen.

Im Gegensatz dazu stellt der vorliegende Beitrag ökologisch valide qualitative Eye Tracking-Analysen zum Blickverhalten von Zeigenden und ihren Adressaten in natürlichen Interaktionen vor. In den Analysen werden Präzisionsdefizite traditioneller Videoaufnahmen durch robuste Beobachtungen zu exakten Blickfixierungen und Blickbewegungsverläufen vermieden. Ausgangspunkt bildet eine empirisch fundierte multimodale Deixistheorie (MDT), die auf der Grundlage umfangreicher Videodaten entwickelt wurde (Stukenbrock 2009, 2015). Einige der dort präsentierten Teilergebnisse zum Blickverhalten können nun auf der Basis mobiler Eye Tracking-Daten verfeinert und ausdifferenziert werden.

Die Untersuchung postuliert, dass Deixis in der face-to-face-Interaktion eine multimodale Praktik der Referenzherstellung ist, die gestisches Zeigen und visuelle Wahrnehmung involviert. Sprache, Gestik und Blick sind in einer spezifischen Weise miteinander verwoben: Blick wird zu einer Ressource interpersoneller Koordinierung, die über regulatorische Funktionen wie z.B. im Turn-Taking hinausgeht. Mein Argument lautet, dass die Blickpraktiken der Beteiligten in deiktischen Referenzierungsakten bestimmte (mikro-)sequenzielle Positionen besetzen und dadurch systematisch dazu beitragen, in einer triadischen Konstellation einen neuen Status von Intersubjektivität zwischen den Beteiligten herzustellen und zu do-

⁵ De Stefani 2010, 2013; De Stefani/Gazin 2014; De Stefani/Mondada 2014; Haddington/Mondada/Neville 2013; Stukenbrock/Birkner 2010.

kumentieren. Triadisch wird die Konstellation bezeichnet, da sich die Beteiligten visuell nicht (nur) aufeinander, sondern auf ein drittes Phänomen, den Gegenstand gemeinsamer Aufmerksamkeitsfokussierung, orientieren.

2. Datenkorpus, methodologische Überlegungen und Herausforderungen

Die folgenden Analysen basieren auf Eye Tracking-Daten, die mit zwei mobilen Eye Tracking-Brillen (SMI, sampling rate 30 Hz) aufgenommen wurden und zusammen mit einer eingebauten Szenenkamera die sichtbare Umgebung der Beteiligten und deren Blickverhalten in Alltagssituationen aufzeichnen. Das im Aufbau befindliche Korpus umfasst Aufzeichnungen von Interaktionsdyaden, die 1. ein informelles Gespräch miteinander führen, 2. auf einem Markt zusammen einkaufen, 3. in einer Bibliothek ein Buch recherchieren, 4. gemeinsam ein Möbelstück aufbauen, 5. miteinander musizieren, 6. ein Museum besuchen.

Die Aufnahme jeder Interaktionsdyade setzt sich aus den Videos der Eye Tracking-Brillen zusammen, die die Beteiligten in der Interaktion tragen. Diese Videos bestehen aus den Aufzeichnungen der jeweiligen Szenenkamera, die in die Eye Tracking-Brille eingebaut ist. Zusätzlich wird binokular, d.h. für jedes Auge, mit zwei in das Brillengestell integrierten Augenkameras das Blickverhalten der beteiligten Personen registriert. Die Daten der Augenkameras werden in das Video der Szenenkamera integriert und erscheinen dort als farbig markierter Blickcursor (Abb 1: blauer Kreis im linken und grüner Kreis im rechten Bild). Der Blickcursor indiziert die Blickfixationen und -bewegungen (i.e. die foveale Augenausrichtung, cf. Duchowski 2003; Land 2006) der BrillenträgerInnen:



Für die Transkription und Analyse wurden die Eye Tracking-Videos der Beteiligten mit Adobe Premiere Pro CS4 synchronisiert und als integriertes audiovisuelles Split-Screen-Video (Abb. 1) exportiert. Für die Rekonstruktion der Mikro-Koordinierung von Sprache, Blick und körperlichem Verhalten der Beteiligten wurden die Videos in das multimodale Annotationstool ELAN⁶ importiert und manuell nach den untersuchungsrelevanten Parametern annotiert.

Der von der Kamera aufgezeichnete Ausschnitt entspricht nicht vollumfänglich der menschlichen Wahrnehmung, die weder ein rechteckiges Bild 'ausgibt', noch

⁶ <https://tla.mpi.nl/tools/tla-tools/elan/>

für das gesamte Gesichtsfeld dieselbe Sehschärfe aufweist, sondern sowohl foveales (d.h. zentrales, hochauflösendes) als auch peripheres Sehen umfasst (vgl. Rossano 2012:24). Gerade die periphere Wahrnehmung, die für bewegliche Phänomene wie z.B. Gesten eine wichtige Rolle spielt, wird von den Eye Trackern nicht bzw. nur sehr eingeschränkt erfasst. In Abb. 1 ist zu sehen, dass im Bildausschnitt von Person 1 (linkes Bild mit blauem Cursor) Ellenbogen und Beine von Person 2 'abgeschnitten' sind. Umgekehrt ist im Ausschnitt von Person 2 (rechtes Bild mit grünem Cursor) der Kopf eines vor der Käsetheke stehenden Kunden 'abgeschnitten', während die Interaktionspartnerin, Person 1, überhaupt nicht sichtbar ist. Diese Ausschnitte entsprechen nicht der insgesamt etwa 200 Grad umfassenden menschlichen Wahrnehmung. In stationären und semi-mobilen Settings wird daher eine dritte Kamera verwendet, die eine Beobachterperspektive auf die Aktivitäten der Beteiligten aus der Totalen bzw. Halbtotalen ermöglicht. Für die Analyse von ausschließlich mit Eye Tracking-Brillen aufgezeichneten Interaktionen stellen diese Aufnahmen besondere Herausforderungen dar und verlangen neue Lesepraktiken wie z.B. die *gaze angle*-Analyse (Schwarzkopf/Büchner/Hölscher/Konieczny 2017).

Bislang werden (mobile) Eye Tracking-Verfahren fast ausschließlich in experimentellen Forschungsdesigns auf den Gebieten der Kognitionswissenschaften, der Psychologie, von Markt- und Usability-Forschung genutzt (vgl. den Überblick bei Horsely et al. 2014); doch sie können auch für hochauflösende, multimodale Mikroanalysen verbaler Interaktion im Forschungsparadigma von Konversationsanalyse und Interaktionaler Linguistik verwendet werden. Allerdings sind eine Reihe von Herausforderungen zu bewältigen, die durch den Einsatz von Eye Tracking-Verfahren zur *in situ*-Untersuchung von natürlicher Interaktion auftreten.

Da in solchen Untersuchungen das Ziel darin besteht, soziale Interaktion und die dazu mobilisierten Ressourcen in ihrer endogenen Ordnung und Emergenz zu untersuchen, können die Daten nicht automatisch kodiert werden, wie dies in experimentellen Eye Tracking-Studien mit entsprechenden Programmen üblich ist. Die Arbeitsphasen der verbalen Transkription und multimodalen Annotierung nach konversationsanalytischen Standards (Hepburn/Bolden 2012; Selting et al. 2009) sind extrem zeitaufwändig. Zudem bedeutet das Aufnehmen von Daten 'in freier Wildbahn' ein erhöhtes Risiko, größere Datenmengen zu verlieren. Abgesehen von Problemen sich verschlechternder Kalibrierung, die aus spontanen Bewegungen der Beteiligten entstehen können (vgl. auch Clark/Gergle 2011), beeinflussen ständig wechselnde Lichtverhältnisse und andere externe Faktoren gegebenenfalls die Tracking-Qualität der Daten.

Eine weitere Herausforderung stellt die Mobilität der Beteiligten dar. Während in stationären und semi-stationären Settings der Einsatz einer externen Kamera zur Dokumentation der Interaktion aus der Totalen keinerlei Schwierigkeiten bereitet, ist dies in mobilen Settings auf Märkten o.ä. oft nicht möglich, ohne dass der natürliche Interaktionsfluss gestört wird. Zugleich liefern gerade in mobilen Settings Aufnahmen aus der Beobachterperspektive wichtige Informationen zu den Körperbewegungen der Beteiligten, die nicht in deren jeweiligen Szenekameras sichtbar sind.

Zuletzt schließlich erzeugt die Mobilität der Beteiligten auf mehrfache Weise Probleme der Nicht-Ikonizität in den Daten sowie in verschiedenen Formen der Daten(re)präsentation. Nicht-Ikonizität als unerwünschter Effekt der Synchroni-

sierung der Videos ist unvermeidbar, da den individuellen Videoaufzeichnungen der Beteiligten im Split Screen-Video fest definierte Positionen auf der rechten und linken Bildschirmseite zugeteilt werden müssen. Im Split Screen-Video sind Entscheidungen über die räumliche Anordnung des synchronisierten Mastervideos ein für allemal festgelegt. Doch in natürlichen Interaktionen verändern Beteiligte fortwährend ihre Positionen, sie begeben sich mal auf die rechte und im nächsten Moment auf die linke Seite ihres Partners, laufen vor- und hintereinander her, drehen sich plötzlich um und bewegen sich spontan in alle möglichen Richtungen. Doch sobald die Einzelvideos synchronisiert und in ein Split-Screen-Video fusioniert wurden, haben sie fixe Positionen in der zweidimensionalen Ansicht des Computerbildschirms. Folglich spiegelt die räumliche Anordnung der Beteiligten (rechts/links) im Split Screen-Video und in den Standbildern nicht immer deren aktuelle Position im Interaktionsraum wider, sondern kann in einem invertierten oder anderweitig nicht-ikonischem Verhältnis dazu stehen.

Nicht-Ikonizität beeinflusst nicht nur die visuelle Repräsentation in Standbildern für multimodale Transkripte, Vortragsfolien und Publikationen, sondern auch die Analysierbarkeit der Daten. Für eine sequenzielle Analyse, die sich nicht auf Daten aus der Beobacherperspektive einer dritten Kamera stützen möchte oder kann, sondern allein auf der Grundlage der Eye Tracking-Perspektiven unternommen wird, stellt die Nicht-Ikonizität eine perzeptorische und kognitive Herausforderung dar. Diese birgt allerdings zugleich das methodologische Potenzial, bisherige Seh- und Analysegewohnheiten zu hinterfragen und metareflexiv zugänglich zu machen sowie neue Wege der Analyse beweglicher Bilder zu beschreiten (Schwarzkopf/Büchner/Hölscher/Konieczny 2017).

Das erste Beispiel (Bib03) illustriert einen solchen Fall (siehe Abschnitt 4.1). Darin suchen zwei Teilnehmerinnen nach einem Buch in der Bibliothek. Im Verlauf ihrer Aktivitäten wechseln sie mehrfach ihre Positionen. In der zu analysierenden Sequenz befinden sie sich in einer *side-by-side*-Orientierung und gehen den Flur der Bibliothek entlang. Doch entspricht die Anordnung im Split Screen-Video, in welchem P's Perspektive links und A's Perspektive rechts zu sehen ist, nicht deren realem Nebeneinander in dieser Interaktionsphase: Die Teilnehmerin, deren Perspektive rechts zu sehen ist, läuft in diesem Moment in Wahrheit auf der linken Seite, umgekehrt befindet sich ihre Interaktionspartnerin A, die im Split Screen-Video links platziert wurde, auf der rechten Seite. In solchen Phasen liefert eine dritte Kamera wichtige Zusatzinformationen für die Analyse körperliche Bewegungen und Konfigurationen der Beteiligten im Raum. Das Problem der Nicht-Ikonizität hingegen wird dadurch nicht beseitigt, und zudem entsteht als neues Problem die Intrusion und Störung der natürlichen Interaktion.

Als vorläufige Bilanz lässt sich festhalten, dass eine ganze Reihe von Überlegungen die Entscheidung darüber mitbeeinflussen, ob mobile Eye Tracker verwendet werden oder nicht, ob die Aufnahmen durch zusätzliche externe Kameras begleitet werden sollen und in welcher Weise. Die Bemerkung, mit der Schegloff vor 20 Jahren das Verhältnis neuer Technologien und ihres analytischen Nutzens beschrieben hat, haben ihre Gültigkeit angesichts verfeinerter Datenaufnahmetechnologien nicht verloren, sondern an Aktualität noch gewonnen: "we need to ask what new analytic possibilities are made available by this new technology, and these may follow on the new observational possibilities" (Schegloff 1996:166).

3. Deiktische Praktiken

3.1. Was ist Deixis?

Der Terminus *deixis* leitet sich aus dem altgriechischen Wort δειξίς her, das 'Zeigen' bedeutet. Menschen können auf unterschiedliche Arten zeigen. Neben und im Verbund mit sprachlichen Praktiken werden eine Vielfalt körperlicher Zeigepraktiken verwendet (vgl. Kendon 2004; Kita 2003; Stukenbrock 2015, 2016), die von manuellen Zeigegesten über den Gebrauch anderer Körperteile wie des Kopfs, des Kinns oder der Lippen bis hin zu Instrumenten und Gegenständen reichen. Letztere können auf das Zeigen spezialisiert sein wie Zeigestöcke oder zufällig zur Hand sein (Stukenbrock 2015). In den Sprachen sind bestimmte Formen grammatikalisiert worden, die auf das verbale Zeigen spezialisiert sind und eine eigene Klasse linguistischer Elemente bilden: Deiktika.

3.1.1. Verbaldeixis erfordert Zeigegesten

Von zentraler Bedeutung für den Gebrauch verbaler Deiktika (insbesondere Lokaldeiktika) in der Interaktion von Angesicht zu Angesicht ist ihre Kombination mit Zeigegesten. Ein deiktischer Ausdruck und eine Geste werden zu einem multimodalen Zeigmittel verbunden, durch das eine Sprecherin die visuelle Aufmerksamkeit ihres Adressaten auf eine sichtbare Entität im gemeinsamen Wahrnehmungsraum lenkt.

Der erste Deixistheoretiker, der die zentrale Rolle gestischen Zeigens in deiktischen Referenzierungsakten hervorhebt und das körperliche Zeigen in seine Theorie integriert, ist der deutsche Sprachpsychologe Karl Bühler (1965[1934]). In seinen Überlegungen zur Deixis, die er in der *Sprachtheorie* (1934) formuliert hat, geht Bühler der Frage nach, wie eine spezifische Klasse linguistischer Ausdrücke dazu dienen kann, Äußerungen der Beteiligten in Relation zu deren Origo (i.e. dem Ausgangspunkt ihrer subjektiven Orientierung) bzw. genauer: im Hinblick auf die Dimensionen von Raum, Zeit und Person zu verankern. Die Personalpronomen *ich* und *du* beispielsweise enkodieren die Rolle der Teilnehmer am Gesprächsereignis; Lokaladverbien wie *hier* und *da* referieren auf proximale oder distale Räume in Relation zum Aufenthaltsort der Gesprächsbeteiligten zum Äußerungszeitpunkt; temporale Ausdrücke wie *jetzt*, *bald*, *nachher* und die Tempora des Verbs enkodieren zeitliche Referenz in Abhängigkeit zum Äußerungsmoment (vgl. Fillmore 1997; Levinson 1983).

Neben den drei Dimensionen Raum, Zeit und Person, die in anderen Theorien um weitere Dimensionen erweitert wurden (vgl. Levinson 2000), unterscheidet Bühler drei Modi des Zeigens: 1. *demonstratio ad oculos et ad aures*, d.h. Zeigen auf sichtbare Phänomene im unmittelbaren Wahrnehmungsraum, 2. *Anaphora*, d.h. Zeigen auf sprachliche Elemente im sprachlichen Kontext, und 3. *Deixis am Phantasma*, d.h. Zeigen auf abwesende Phänomene, die nicht im Wahrnehmungs-, sondern im Vorstellungs- bzw. Phantasieraum situiert sind und daher imaginiert werden müssen. Deixis am Phantasma erfordert die Versetzung des Ausgangspunkts unserer Orientierung, der *Origo* (Bühler 1965[1934]) bzw. des indexikalischen Grunds (Hanks 1990; Levinson 2004), an einen imaginierten Standort im Vorstellungsraum. Vorgestellte Phänomene oder Entitäten werden auf andere

raumzeitliche Koordinaten als den unmittelbaren Wahrnehmungsraum, in dem wir normalerweise verankert sind, bezogen. Sie werden nicht von unserem leiblichen, sondern einem zu imaginierenden Standort aus betrachtet (vgl. ausführlich Stukenbrock 2012, 2014b, 2015, 2016 sowie aus Sicht der *Mental Space Theory* Ehmer 2011). Der vorliegende Beitrag konzentriert sich auf die Analyse deiktischer Praktiken in der *demonstratio ad oculos et ad aures*. Deixis am Phantasma und anaphorisches Zeigen bleiben demgegenüber unberücksichtigt.

Bühlers Einsicht in die zentrale Rolle, die dem körperlichen Zeigen und der visuellen Wahrnehmungssteuerung beim Gebrauch von Deiktika in der face-to-face-Interaktion zukommt, wird durch eine wachsende Zahl von Untersuchungen in der Multimodalitätsforschung empirisch belegt.⁷ Multimodale Studien zur Deixis zeigen, dass der Gebrauch und das Verstehen deiktischer Ausdrücke leibgebunden ist. Die Rekonstruktion deiktischer Praktiken in der Interaktion ist folglich eine analytische Aufgabe, die Perspektivität und Intersubjektivität nicht nur als kognitives Problem, sondern als leibgebundenen Prozess begreifen muss. Dieser verlangt von den Beteiligten eine wechselseitige, von Augenblick-zu-Augenblick zu leistende Kalibrierung von Rede, Gestik, Blick und Bewegungen in der emergierenden Interaktion.

Im Zuge der Erkenntnis, dass im grammatischen System der Sprachen enkodiertes Zeigen (Deixis) ebenso wie gestisches Zeigen als kommunikatives Mittel zur Herstellung gemeinsamer Aufmerksamkeit konzeptualisiert werden kann (*joint attention*)⁸, hat sich das Augenmerk empirischer Untersuchungen zunächst auf Gesten konzentriert. Der Blick wurde bislang vernachlässigt. Wenn jedoch erstens die Herstellung eines gemeinsamen Aufmerksamkeitsfokus eine der zentralen Funktionen deiktischer Ausdrücke und Zeigegesten ist und wenn zweitens spezifische Blickpraktiken der Beteiligten den interaktiven Stellenwert eben jener Praktiken für die Beteiligten selbst sichtbar dokumentieren, ist es unabdingbar, auch das Blickverhalten systematisch zu untersuchen.

In deiktischen Praktiken ist der Blick eine zentrale Ressource für die Beteiligten, sie setzen spezifische Blickmuster nicht nur zum Turn-Taking u.ä., sondern auch zur Kontextualisierung der Relevanz visueller Wahrnehmung selbst ein: zum Zeigen mit dem Blick, zum *flagging* (Streeck 2002) relevanter Gesten im Augenblick ihrer Emergenz, zur sichtbaren, interaktiven Dokumentation visueller Aufmerksamkeitsorientierung, zum Monitoring der visuellen Aufmerksamkeit Anderer, zur Dokumentation von Wahrnehmung und zur Wahrnehmungswahrnehmung.

3.1.2. Verbaldeixis und Zeigegesten fordern den Adressatenblick ein

Im Folgenden knüpfe ich an Bühlers Behauptung an, dass Verbaldeixis auf Zeigegesten oder äquivalente visuelle Zeigehilfen angewiesen ist (vgl. Bühler 1965 [1934]:80 sowie Abschnitt 3.2), und argumentiere darüber hinaus, dass verbaldeiktische Formen in ihrer primordialen Verwendung in mitgegenwärtiger Interaktion nicht nur Zeigegesten benötigen, sondern überdies den Adressatenblick

⁷ Eriksson 2009; Fricke 2007; Goodwin 2003; Hausendorf 2003; Hindmarsh/Heath 2000; Keavallik 2013; Mondada 2007, 2012a; Streeck 2002; Stukenbrock 2008, 2009, 2010, 2015.

⁸ Vgl. Diessel 2006; Eriksson 2009; Goodwin 2003; Tomasello 2008; Stukenbrock 2009, 2015.

einfordern. Indem Deiktika eine multimodale Gestalt mit Zeigegesten bilden, die gesehen werden müssen, fordern sie den Blick des Adressaten ein (*request for gaze*, Stukenbrock im Druck b). Diese Blickeinforderung wird hier als erster Paarteil einer multimodal zu begreifenden Paarsequenz konzeptualisiert, die auch und zu meist äußerungsintern innerhalb anderer (übergeordneter) Sequenzstrukturen operiert. Ein deiktischer Ausdruck ist damit nicht allein eine grammatische Form, die bestimmte Merkmale der Origo (bzw. des indexikalischen Grunds), des Referenten und der Beziehung zwischen Origo und Referent (bzw. zwischen Figur und Grund) enkodiert, sondern interaktional formuliert fungiert er als Fokussierungsaufforderung (*summons*), der eine Fokussierungsbestätigung (*answer*), in diesem Fall die visuelle Aufmerksamkeit des Adressaten, seine Blickzuwendung, konditionell relevant setzt.

Die These, dass sowohl turn-initiale als auch turn-interne Merkmale in der emergierenden Rede eine Fokussierungsaufforderung implementieren können, die als Antwort den Blick des Adressaten einfordern, wurde bereits von Goodwin (1981) in seiner einflussreichen Studie zu *Restarts, pauses, and the achievement of mutual gaze in face-to-face interaction* formuliert und nachgewiesen. Darin stellt Goodwin eine systematische Beziehung zwischen spezifischen vokalen Praktiken des Sprechers wie Abbrüchen, Neuanfängen, Pausen und Verzögerungen und dem Blickverhalten des Hörers fest. Sein Argument lautet, dass diese Praktiken als "a request for the gaze of a hearer" operieren, so dass "the actions of speaker and hearer together [...] a particular type of summons-answer sequence" konstituieren (Goodwin 1981:280). Er stellt die Hypothese auf, "that summons-answer sequences might function not only to provide coordinated entry into a conversation as a whole (Schegloff (1968:1089) but also to establish the availability of participants toward each other within the turn itself" (Goodwin 1981:282f., Verweis auf Schegloff im Original).

Unabhängig davon, ob eine Blickeinforderung durch einen Abbruch, einen Neustart, eine Pause ("summons for gaze" nach Goodwin 1980, 1981) oder durch einen deiktischen Ausdruck (Stukenbrock im Druck a, b) ergeht, lässt sich festhalten, dass ein solcher "summons for gaze" nicht nur auf der Handlungsebene, sondern ganz offenbar auch unterhalb der Ebene genuiner Handlungen (*actions* im Sinne der Konversationsanalyse, vgl. Levinson 2012) geschehen kann. Der gestische Gebrauch eines deiktischen Ausdrucks weist temporale und sequenzielle Besonderheiten auf, da er den Adressaten auffordert, seinen Blick in einer spezifischen Weise zu gebrauchen und zuerst zum Sprecher zu schauen, um zusätzliche visuelle Informationen aus dessen verkörperten Ausdrucksressourcen (z.B. einer manuellen Zeigegeste) zu erlangen, und dann seinen Blick zum Objekt gemeinsamer Aufmerksamkeit zu reorientieren.

Bühler hat in kongenialer Weise der irreduziblen Rolle der menschlichen Sinnesorgane und insbesondere der Augen in deiktischen Referenzierungshandlungen Ausdruck verliehen, indem er den ersten Modus der Deixis mit dem lateinischen Syntagma *demonstratio ad oculos et ad aures* benannt hat. Der Terminus *demonstratio ad oculos* (*demonstratio* = 'Zeigen' und *ad oculos* = 'zu den/für die Augen') verbindet die Zeigegeste des Sprechers (die *demonstratio*) bzw. dessen zeigende Hand mit den Augen des Adressaten (*ad oculos*). Der hier empirisch zu untersuchende interpersonelle Gesten-Augen-Nexus ist in der Bühlerschen Terminologie bereits theoretisch kondensiert.

In den folgenden Analysen werde ich zeigen, dass der *Gesten-Augen-* bzw. *Gesten-Blick-Nexus* (vgl. Abschnitt 4.1) nur eine der systematisch auftretenden Relationen ist, die durch das Blickverhalten der Beteiligten in deiktischen Referenzierungshandlungen implementiert werden. Die unterschiedlichen Blickmuster treten an spezifischen sequenziellen Positionen auf und sind kontextsensibel, zeitlich flexible Komponenten einer extrem feinteiligen interpersonellen Koordination, auf die deiktische Praktiken angewiesen sind und die sie zugleich systematisch herstellen.

4. Analysen zum Blickverhalten in deiktischen Praktiken

Gegenstand der empirischen Analysen ist der Zusammenhang zwischen verbalen und körperlichen Zeigepraktiken einerseits und Blickpraktiken der Beteiligten andererseits. In den Analysen wird P durchgängig zur Bezeichnung der zeigenden Person und A für die Adressatin verwendet.

4.1. Der Gesten-Blick-Nexus in der *demonstratio ad oculos*

Referiert ein Sprecher deiktisch auf ein Phänomen im gemeinsamen Wahrnehmungsraum und verwendet eine Zeigegeste, um den Blick des Adressaten dorthin zu lenken, muss das körperliche Zeigen des Sprechers von seinem Adressaten visuell wahrgenommen werden. Blickt der Adressat nicht bereits zum Sprecher, so hat der Sprecher unterschiedliche verbale und körperliche Praktiken zur Verfügung, sich den Blick des Adressaten zu sichern wie z.B. Vokative, explizite Fokussierungsaufforderungen (*summons*), Körperbewegungen, Berührungen etc.

Das im Folgenden ausgeführte Argument lautet, dass deiktische Ausdrücke eine Fokussierungsaufforderung konstituieren und intrinsisch die Blickzuwendung des Adressaten einfordern. Die Fokussierungsaufforderung, die durch Deiktika ergeht, ist insofern spezifisch, als sie ein komplexes Blickorientierungsmuster initiiert – eines, das vom Adressaten verlangt, sich visuell auf den Körper und die Geste des Sprechers zu orientieren, um von dort vektorielle Informationen darüber zu erhalten, wohin er seinen Blick im nächsten Schritt richten muss, um das Zeigziel auffinden und Referenz herstellen zu können (zur Unterscheidung von Zeigziel und Referenten Quine 1960; Stukenbrock 2009, 2015).

Im ersten Beispiel belegt die Blickfixation in den Eye Tracking-Daten, dass die Adressatin A direkt auf P's Geste schaut, bevor sie ihren Blick auf den indizierten Suchraum und das Zeigziel richtet. Das Beispiel stammt aus dem Bibliothekskorpus. Die Beteiligten, zwei Studentinnen, haben soeben von der Bibliotheksangestellten erfahren, dass die Signatur des Buchs, das sie suchen, mit dem Buchstaben A beginnt. Der Ausschnitt setzt damit ein, dass die Studentinnen auf eine Glastür zulaufen, die die Bibliothekssektionen abtrennt. Im Transkript sind die aus dem Split Screen-Video extrahierten Standbilder zeitlich exakt mit den fett markierten Sprechsilben aligniert. P's Perspektive befindet sich auf der linken und A's Perspektive auf der rechten Seite.⁹

⁹ In diesem Fall ist die Rechts-Links-Anordnung der Beteiligten im Split-Screen-Videos nicht-ikonisch (vgl. Abschnitt 2).

Beispiel 1: "DA ist A"



Abb. 1.1 | Abb. 1.2
01 (3.0 |)



Abb. 1.3 | Abb. 1.4

02 P: °hh DA ist A.



Abb. 1.5 | Abb. 1.6
03 ((die Beteiligten gehen nacheinander durch die Tür))

Zu Beginn des Auszugs laufen die Beteiligten schweigend nebeneinander auf die Glastür zu. Sie befinden sich in einem "open state of talk" (Goffman 1981:134) bzw. einem "'continuing state of incipient talk'" (Schegloff/Sacks 1973:325). Abb. 1.1 hält die Perspektiven von A und P fest, während sie sich der Glastür nähern. Beide blicken geradeaus (Abb. 1.1/rechts und links), wobei P ihrer Freundin leicht vorausläuft, wie A's Szenenkamera belegt (Abb. 1.1/rechts).

Wie in Abb. 1.2 zu sehen ist, öffnet P im nächsten Moment den rechten Flügel der Glastür und richtet ihren Blick auf die Regale am hinteren Ende des Raums (Abb. 1.2/links: blauer Tracking-Kreis). In deixistheoretischer Terminologie (Stukenbrock 2009, 2015) etabliert sie visuell den künftigen Verweisraum einer deixischen Zeigehandlung (P-> VR).

Während sie einatmet und damit den Beginn einer neuen Äußerung projiziert (Z. 2: °hh), hebt P ihren linken Arm (Abb. 1.3/links) und zeigt mit ausgestrecktem Zeigefinger (Abb. 1.4/links) auf den Raumausschnitt, auf den sie im nächsten Moment auch deixistisch mit dem Lokaladverb *DA/there* (Z. 1: *DA ist A/there is A.*) referiert. Zunächst taucht die Geste in ihrer eigenen Szenenkamera auf (Abb. 1.3/links und Abb. 1.4/links), bevor sie auch in der Aufnahme ihrer Adressatin sichtbar wird (Abb. 1.4/rechts).

A beginnt sich leicht zu P zu drehen. A's Szenenkamera zeigt, dass sie nicht mehr direkt nach vorne durch die Glastür in die Bibliothek blickt, sondern zur Seite (Abb. 1.3/rechts). Genauer gesagt bestehen A's Eye Tracking-Daten aus Sakka-

den und offenbaren, dass sie ihren Blick in Richtung der Sprecherin P zu orientieren beginnt. Diese Blickbewegung kommt kurz zur Ruhe, als A die Hand mit dem ausgestreckten Zeigefinger von P erblickt (Abb. 1.4/rechts). In den Eye Tracking-Daten gibt es eine kurze Fixation auf P's Zeigefinger (Abb. 1.4/rechts), bevor A ihren Blick in die Richtung wendet, die P gestisch indiziert hat (Abb. 1.5/rechts).

Interessanterweise wird A nun in P's Szenenkamera sichtbar. Das bedeutet, dass das *side-by-side*-Arrangement, in dem P kurzzeitig leicht voranlief, eine neue Dynamik entwickelt. Nun hat A etwas Vorsprung gegenüber P, die nicht nur zeigt, sondern A auch die Tür aufhält. Diese Doppelgeste aus Zeigen und Türaufhalten kann von A als multimodale Aufforderung verstanden werden, der Zeigegeste nicht nur mit dem Blick zu folgen, sondern sie als direktiven Akt zu interpretieren, in die angegebene Richtung zu laufen.

A kommt der Aufforderung nach und tritt als erste durch die Glastür, was in P's Szenenkamera dokumentiert wird (Abb. 1.5-1.6/links). Als sie durch die Glastür geht, fokussiert A ihren Blick nicht auf einen spezifischen Ort, sondern scannt die Regale auf beiden Seiten der Bibliothek sowie auf den oberen und unteren Ebenen des Raums (Abb. 1.5-1.6/rechts). Inzwischen hat sie ihre Interaktionspartnerin P überholt und läuft vor ihr (Abb. 1.6/rechts). P ihrerseits retrahiert ihre Zeigegeste und richtet ihren Blick von den Bücherregalen (Abb. 1.4/links) auf ihre Adressatin A, als diese in ihr Gesichtsfeld tritt. Aufgrund der hohen Präzision der Eye Tracking-Daten lässt sich noch genauer feststellen, dass P ihren Blick auf die Region von A's Augen orientiert (Abb. 1.5/links). Auch wenn die Eye Tracking-Brillen in diesem Moment eine direkte Wahrnehmung von A's Augenausrichtung erschweren, kann P nichtsdestotrotz aus der vektoriellen Orientierung von A's Brillengläsern inferieren, dass sie zum hinteren Ende des Raums blickt.

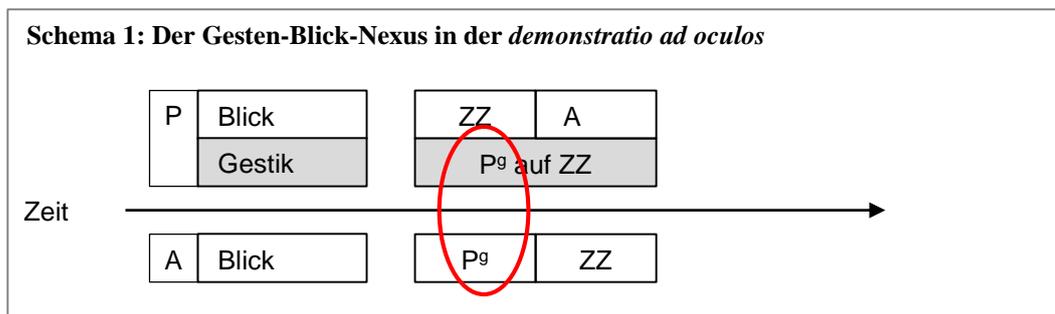
Als A durch die Glastür geht, die P für sie aufhält, löst sich das *side-by-side*-Arrangement weiter auf und wird von den Beteiligten in eine *face-to-back*-Orientierung transformiert. P, deren Sicht in den Raum hinein nun teilweise durch A blockiert wird, die vor P läuft (Abb. 1.6/links), beginnt wie A die Regale in dem neuen Raum zu scannen (Abb. 1.6/links: blauer Tracking-Kreis auf den Bücherregalen).

Zusammenfassend: Die Analyse hat verdeutlicht, dass die zeigende Person P auf das Zeigziel blickt, den lokaldeiktischen Ausdruck *DA* in Verbindung mit einer Zeigegeste verwendet, während die Adressatin A, als sie das deiktische Adverb *DA* hört, ihren Blick reorientiert und auf die zeigende Sprecherin, genauer gesagt auf deren Zeigegeste, richtet. Dieser Moment konstituiert den hier als *Gesten-Blick-Nexus* konzeptualisierten Zusammenhang zwischen körperlichen Zeigepraktiken und Blickpraktiken in der *demonstratio ad oculos*.

Dieses Phänomen ist nicht immer in den Eye Tracking-Daten in Form von Blickfixationen auf dem gestikulierenden Körperteil der Sprecherin oder des Sprechers zu beobachten. Experimentelle Studien zu Blickfixationen des Adressaten auf Gesten von Sprechern und der Verarbeitung gestisch vermittelter Informationen weisen darauf hin, dass es keine direkte Beziehung zwischen Adressatenfixationen und der Verarbeitung direktonaler Gesteninformationen gibt, dass beides aber zunimmt, wenn die Sprecher auch selbst auf ihre Gesten blicken (Gullberg/Kita 2009). Wie bereits Streeck (1993, 2002) in konversationsanalytischen Untersuchungen festgestellt hat, lenken Sprecher durch den Blick die visuelle Aufmerksamkeit der Adressaten auf ihre Gesten. Doch im Gegensatz zu den von

Streeck untersuchten ikonischen bzw. metaphorischen Gesten richten Sprecher bei deiktischen Zeigegesten ihren Blick nicht auf die eigene Geste, da diese nicht das Ziel der Adressatenaufmerksamkeit darstellt. Stattdessen blicken sie auf das Zeigeziel und liefern dem Adressaten dadurch zusätzliche vektorielle Informationen zum Auffinden des Zeigeziels. Daher kann ausgeschlossen werden, dass eine erhöhte Anzahl von Adressatenfixationen auf Sprechergesten dadurch zustande kommt, dass Sprecher auf ihre Zeigegesten blicken und diese dadurch für Adressaten in der Relevanz hochstufen.

Abstrakt lässt sich das Ergebnis der Analyse in folgendem Schema zusammenfassen. Darin repräsentiert "P" die zeigende Person, hier mit den relevanten Parametern Blick und Gestik, und "A" die Adressatin, für die vorerst nur der Blick relevant ist. "ZZ" steht für Zeigeziel:



Die relevante Phase interpersoneller Koordinierung, in der der Gesten-Blick-Nexus als relevante Komponente von *demonstratio ad oculos*-Praktiken interaktiv hergestellt wird, kann folgendermaßen beschrieben werden: Demnach blickt und zeigt P auf das Zeigeziel. Zugleich richtet A ihren Blick auf P, in diesem Fall speziell auf P's Geste (P^g). Dieser im Schema rot markierte Moment konstituiert den interpersonellen Gesten-Blick-Nexus zwischen Sprecherhand und Adressatenaug. Wie später zu zeigen sein wird (siehe Abschnitt 4.3, Schema 4), richtet P anschließend den Blick auf A, um zu überprüfen, ob A das Zeigeziel sieht.

4.2. Der Ziel-Blick-Nexus in der *demonstratio ad oculos*

Einen gemeinsamen Aufmerksamkeitsfokus auf ein sichtbares Objekt in Form exakt verfolgbare Blickmuster zwischen der zeigenden Person P und dem Adressaten A in den Eye Tracking-Daten nachzuweisen, bedeutet, dass diese zu einem bestimmten Moment in der emergierenden Interaktion Blickfixationen sowohl von P als auch von A auf demselben Phänomen aufweisen müssten, dem Gegenstand gemeinsamer Aufmerksamkeitsfokussierung. Dies kann gleichzeitig geschehen, so dass die synchronisierten Eye Tracking-Aufnahmen von P und A eine Fixation auf demselben Phänomen im selben Moment aufweisen; oder die Fixationen von P und A treten nacheinander auf. In letzterem Fall blickt die zeigende Person P in der Regel zum Adressaten A, um dessen visuelle Aufmerksamkeitsorientierung zu kontrollieren, oder sie schaut woanders hin.

Typischerweise ist das Blickmuster von P so organisiert, dass sie zunächst auf das Zeigeziel blickt und zeigt, dann die Geste einfriert und den Blick vom Zeigeziel auf den Adressaten A richtet, um dessen Aufmerksamkeitsorientierung zu

überprüfen (Stukenbrock 2009, 2015:314-330). Wenn P aus der Blickausrichtung des Adressaten A inferieren kann, dass A die visuelle Aufmerksamkeit ebenfalls auf das indizierte Objekt gerichtet hat, wird P den Blick wieder auf das Objekt gemeinsamer Aufmerksamkeit richten. Dieses bei deiktischen Zeigehandlungen standardmäßig auftretende Blickmuster schafft auf mikrointeraktiv geordnete Weise Momente gemeinsamer visueller Aufmerksamkeit zwischen den Beteiligten.

Das nächste Beispiel aus dem Marktkorpus repräsentiert dieses klassische Muster. Die zwei Beteiligten, P und A, haben soeben den Marktplatz betreten. In den folgenden Standbildern ist P's Perspektive auf der linken und A's Perspektive auf der rechten Seite repräsentiert. In diesem Beispiel ist die Position der Teilnehmer im Split-Screen-Videos ikonisch.

Beispiel 2: "Kürbis"

01 A: () einmal GUCKen,
 02 (3.1)
 03 A: soll_ma einmal kurz ähm SCHRÄG laufen?=oder: ähm-
 04 (2.8)



Abb. 2.1

05 P: **jah:;**
 06 (-)
 07 A: nee?
 08 (0.9)
 09 P: ha_ja_der ist ja gar nicht so GROß,



Abb. 2.2

10 A: **nee;**
 11 das (.) das geht glaub ich SCHNELL;



Abb. 2.3

12 (0.9)



Abb. 2.4

13 P: **HIER**;
14 =GUCK mal;



Abb. 2.5

15 (1.2)



Abb. 2.6

16 P: da gibt_s **KÜR**bisse;



Abb. 2.7

17 A: oKAY;
18 ja die sind mir zu GROSS aber- (.)

Nachdem sie gemeinsam den Marktplatz betreten haben, handeln P und A zunächst aus, wie sie vorgehen wollen (Z. 1-11). Schließlich einigen sie sich auf den Vorschlag von A (Z. 3), erst über den Markt zu laufen und zu schauen, was es gibt. Wie in Abb. 2.1 zu sehen ist, blicken beide während des Aushandlungsprozesses geradeaus auf den Marktplatz (Abb. 2.1/rechts und links). Dabei lässt A gemäß mit ihrer Projektion, den Markt zunächst in Augenschein zu nehmen, ihren Blick über die Marktstände gleiten und scannt die Stände in mittlerer Distanz (Abb.2.2-2.3/rechts). Demgegenüber fokussiert P, der dem Vorschlag nur zögernd zugestimmt an, einen Typ von Waren am ersten Marktstand: Kürbisse (Abb. 2.2-2.3/links).

Laut multimodaler Deixistheorie (Stukenbrock 2009, 2015) blickt P damit auf ein Phänomen, das zum Zeigeziel (ZZ) innerhalb eines Suchraums (SR) wird, wenn und sobald P die visuelle Aufmerksamkeit seiner Adressatin A darauf lenkt. In der Pause (Z. 12) bleibt sein Blick auf die Kürbisse fixiert, dann äußert er das Lokdaldeiktikon *HIER* (Z. 13), auf das in der nächsten TCU ein visuelles Wahrnehmungsverb in der Imperativform folgt: *GUCK mal* (Z. 14). P's Äußerung konstituiert den ersten Paarteil (EPT) einer multimodalen Paarsequenz (vgl. De Stefani/Gazin 2014; Stukenbrock 2014a, im Druck b), deren zweiter Paarteil leiblich realisiert wird. Als erster Paarteil macht eine Fokussierungsaufforderung eine Fokussierungsbestätigung konditionell relevant; in diesem Fall heißt das konkret, dass A ihrem Interaktionspartner erkennbar ihren Blick zuwendet und dann dort hin orientiert, wohin er zeigt.

Dass P den Blick seiner Adressatin genau in diesem Moment in der emergierenden Interaktion einfordert und auf eine bestimmte Ware lenkt, bedeutet, dass er den Status des ersten Marktstandes von einem beiläufig zu passierenden Stand unter vielen anderen in einen transformiert, der ein potenzielles Kaufobjekt darbietet (Dao/Stukenbrock im Druck). Damit suspendiert und verändert er den interaktiv hergestellten Projektionsbogen, der einen gemeinsamen Orientierungsgang über den Markt vorsah.

In der anschließenden Pause (Z. 15) richten beide Teilnehmer ihren Blick neu aus. A folgt der Fokussierungsaufforderung und blickt zu dem von P indizierten Ort (Abb. 2.5/rechts). Im Gegensatz dazu ist P's Blick nicht mehr auf die Kürbisse, sondern auf das Gesicht und die Augen seiner Adressatin gerichtet (Abb. 2.5/links), um ihre visuelle Aufmerksamkeit zu prüfen.¹⁰ Die Eye Tracking-Daten offenbaren, dass A's Blick genau in dem Moment, in dem P zu A blickt und seine Eye Tracking-Aufnahme eine Blickfixation auf A's Augenregion aufweist (Abb. 2.5/links: grüner Tracking-Kreis auf A's Gesicht und Augen), bereits auf den Kürbis gerichtet ist (Abb. 2.5/rechts: blauer Tracking-Kreis auf Kürbis).

Dies ist ein interaktiver Schlüsselmoment im temporalen Ablaufformat deiktischer Zeigehandlungen: Die Blickpraktiken der Beteiligten und deren temporale Koordinierung stellen systematisch einen Moment der Meta-Wahrnehmung her: Der Augenblick, in dem der zeigende Teilnehmer P seine visuelle Wahrnehmung auf die Wahrnehmung seiner Adressatin ausrichtet, konstituiert einen Moment der Wahrnehmungswahrnehmung. Sichtbar und inferierbar wird dieser Moment der Wahrnehmungswahrnehmung in den Eye Tracking-Daten durch Blickfixationen von P auf A's Augen als visuell wahrnehmbaren Indikatoren ihrer Blickorientierung und zu interpretierenden Quellen ihrer Wahrnehmung (siehe Abschnitt 4.3). Im vorliegenden Beispiel schaut P zu A und richtet seinen Blick auf ihre Augen (Abb. 2.5/links), um aus der vektoriellen Orientierung ihres Blicks Schlüsse darüber zu ziehen, ob sie die Kürbisse sieht.

Als A sich umdreht, äußert P eine Präsentativkonstruktion, die die Referenz seines deiktischen Akts (Z. 13: *HIER*) präzisiert: *da gibt_s KÜRbisse*; (Z. 16). Diese Feststellung macht eine verbale Reaktion konditionell relevant. Sie enthält ein Lokaladverb (*da*), das sich gleichermaßen deiktisch auf ein sichtbares Phäno-

¹⁰ Diese Blickpraktik, die andernorts als "Kontrollblick zur Wahrnehmungswahrnehmung" (Stukenbrock 2009; 2015:314ff.) konzeptualisiert wurde, tritt in deiktischen Praktiken regelmäßig auf und besetzt im multimodalen Gesamtformat von Zeigehandlungen einen systematischen Slot (vgl. Abschnitt 4.3.).

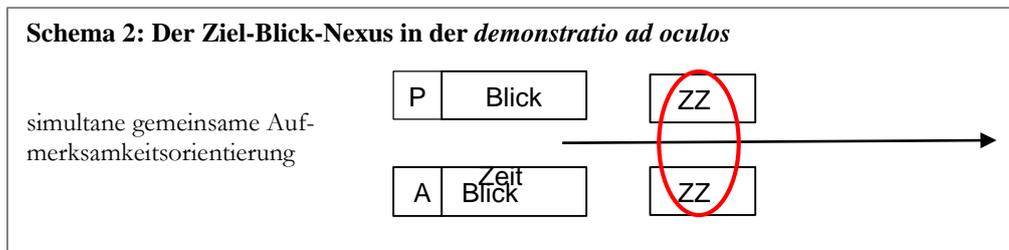
men im Wahrnehmungsraum und anaphorisch auf den zuvor deiktisch (Z. 13: *HIER*) hergestellten Raum bezieht und diesen elaboriert. Die Äußerung wird durch eine Zeigegeste begleitet, die P auf Hüfthöhe (Abb. 6/rechts) relativ niedrig, aber auf der Blickachse von A ausführt (Abb. 2.6/rechts: blauer Tracking-Kreis für A's Blick entspricht der Höhe von P's Unterarm in A's Szenenvideo). Währenddessen schaut P zu seiner Adressatin, die bereits auf die Kürbisse blickt.

Am Äußerungsende blickt P wieder auf die Kürbisse. Die Eye Tracking-Daten dokumentieren, dass A's Blick ebenfalls auf die Kürbisse gerichtet ist. Dies konstituiert einen genuinen Moment simultaner Aufmerksamkeitsorientierung auf ein gemeinsames Anschauungsobjekt. Beide, der grüne und der blaue Tracking-Kreis dokumentieren zur selben Zeit eine Blickfixation der Beteiligten P und A auf dem Kürbis (Abb. 2.7/links und rechts).

Dabei ist festzuhalten, dass *joint attention* sowohl simultan als auch zeitlich zerdehnt hergestellt werden kann, weshalb eine automatisierte Errechnung von sogenannten "gaze cross recurrence"-Instanzen (Jermann/Nüssli 2012) keine validen Ergebnisse zur erfolgreichen Herstellung von *joint attention* im Sinne von gemeinsamer und als gemeinsam gewußter Wahrnehmung eines Phänomens erbringt (vgl. Stukenbrock im Druck b). Dazu gehören weitere interaktive Schritte wie die Wahrnehmungswahrnehmung (Abschnitt 4.3) und die Verstehensdokumentation (vgl. auch Stukenbrock 2015:85-93, 313-343), mit denen sich die Beteiligten gegenseitig Wahrnehmen und Verstehen als gemeinsam vollzogene Prozesse anzeigen.

In unserem Beispiel wenden sich die Beteiligten nach dem Moment gemeinsamer Aufmerksamkeitsorientierung von den Kürbissen ab und beginnen erneut, die Marktstände zu scannen. A äußert ein Rezipientensignal (Z. 17: *oKAY*), das ihr Verstehen dokumentiert, aber zugleich ihre Ambivalenz in Bezug auf die Kürbisentdeckung kontextualisiert und eine dispräferierte Reaktion projiziert. In der Tat folgt eine negative Bewertung (Z. 18: *ja die sind mir zu GROSS aber-*), die die Möglichkeit, den Kürbis als Kaufobjekt in Betracht zu ziehen, ausschließt. Das angeschlossene *aber* allerdings projiziert bereits eine mögliche Neubewertung. Der damit verbundene Fortgang der Interaktion wird hier nicht weiter verfolgt (siehe aber Dao/Stukenbrock im Druck).

Zusammenfassend ist festzuhalten, dass es Momente gemeinsamer visueller Aufmerksamkeitsorientierung gibt, die sich, wenn sie simultan auftreten, in den Eye Tracking-Daten in dem *frame*-genau synchronisierten Split Screen-Video als gleichzeitige Blickfixationen in P's und A's jeweilige Tracking-Aufnahme auf dem Zeigeziel darstellen. Dies lässt sich schematisch wie folgt veranschaulichen:



Der Begriff *Ziel-Blick-Nexus* wird hier in doppelter Weise für Zeigende und für deren Adressaten verwendet, die, der Geste des Zeigenden folgend, mit ihrem Blick das Zeigeziel aufsuchen. Dabei stellt sich die Beziehung zwischen Blick und Zeigeziel bzw. zwischen Sehen und zu Sehendem sowohl zeitlich als auch interaktiv in je spezifischer Weise für Zeigende und für Adressaten her. Während es für den ersteren darum geht, durch das Wahrnehmen des Zeigeziels die eigene Zeigegeste zu monitoren und möglichst präzise und sichtbar für den Adressaten zu zeigen, steht letzterer vor der Aufgabe, das Zeigeziel anhand der imaginär auf das Zeigeziel hin zu extrapolierenden Zeigegeste des Zeigenden zu finden. Wie im Folgenden argumentiert wird, ist dies aber keine mathematisch-vektorielle, optische Verlängerungsübung, sondern eine interaktive und sozio-kognitive Herstellungsleistung.

Zeigende blicken, bevor sie die Aufmerksamkeit ihrer Adressatinnen durch einen Zeigeakt auf ein Zeigeziel lenken, zunächst selbst dorthin (Stukenbrock 2009, 2015:56-59, 241-252) und koordinieren die leiblich-räumliche Selbstverankerung triadisch mit der des Adressaten und des Zeigeziels, bevor sie eine zeitlich und räumlich auf die Aufmerksamkeitsorientierung und leibliche Position des Adressaten zugeschnittene körperliche Zeigegeste ausführen. Findet der Adressat das Zeigeziel nicht, so dass es zu einer Reparatursequenz kommt, erfordert dies vom Zeigenden eine Rekalibrierung seiner Zeigegesten in räumlicher Feinkoordination der triadischen Beziehung zwischen ihm, dem Zeigeziel und dem visuellen Zugang des Adressaten zu beiden.

Um die Beziehung zwischen Zeigeziel und Adressatenblick auszudifferenzieren und dem Faktum Rechnung zu tragen, dass Adressaten das Zeigeziel nicht immer erfolgreich auffinden, sondern in einer auf den indizierten Raum gerichteten Suchbewegung verbleiben, ist in Anlehnung an Goodwins Konzept der *domain of scrutiny* (Goodwin 2003:211) der Begriff des *Suchraums* (SR) für den Raum eingeführt worden, auf den sich Adressatinnen und Adressaten orientieren, wenn Zeigende ihre Aufmerksamkeit auf ein sichtbares Zeigeziel im umgebenden Raum lenken (Stukenbrock 2009, 2015:71-72, 272-281). Als Komplementärbegriff für den Raum, auf den Zeigende sich orientieren, bevor sie die visuelle Aufmerksamkeit ihrer Adressaten auf etwas Drittes orientieren können, fungiert der Begriff *Verweisraum* (VR) (Stukenbrock 2009, 2015:56-59, 241-252). In Beispiel 2 ("Kürbis") war zu sehen, wie P vorauslaufend auf die Kiste mit Kürbissen blickt (Beispiel 2: Abb. 2.2/links, Abb. 2.3/links), bevor er die Aufmerksamkeit seiner Adressatin darauf lenkt.

Praktiken vorauslaufender intra-personeller Selbstkalibrierung des Zeigenden (Sich-zum-Zeigen-Anschickenden) wie räumliche Repositionierung, Blickwechsel zwischen Verweisraum und Adressaten sind Ausdruck und Mittel der zwischenleiblichen Sensibilität des Zeigenden gegenüber der triadischen, deiktisch zwischen ihm, dem Adressaten und einem sichtbaren Phänomen herzustellenden Beziehung. Diese systematisch auftretenden Praktiken der Selbstorientierung können, wenn sie vom Adressaten wahrgenommen werden, eine interpersonell wirksam werdende Projektionskraft entfalten. Dabei sind die Übergänge zwischen mehr oder weniger ostentativen perzeptorischen Selbstorientierungen fließend, diese können unsichtbar, sichtbar und schließlich auch als Display und adressatenseitige "pre-invitation" vollzogen werden.

Das folgende Beispiel 3 aus einer stationären Gesprächssituation ist ein weiterer Beleg für die beschriebenen Blickpraktiken. Zwei Studentinnen unterhalten sich im Gemeinschaftsraum des Instituts über Dekorationen für das bevorstehende Weihnachtsfest. Die Sequenz setzt ein, nachdem A sich bei der Gesprächspartnerin über ihre Mutter beklagt hat, die einen künstlichen Weihnachtsbaum aufstellen möchte. Die Zeigende P, deren Perspektive sich im Splitscreen-Video auf der linken Seite befindet, zeigt daraufhin auf einen Plastikbaum im Gemeinschaftsraum, um in spaßhafter Modalität eine visuelle Gegenevidenz gegen die Scheußlichkeit von Plastikbäumen zu liefern. A's Perspektive ist auf der rechten Seite zu sehen.

Beispiel 3: "Plastikbaum"

01 A: <<acc>da ha_ich_gsa_höma_was machen wir mit_m PLASTikbaum;



Abb. 3.1

02 =ins **WOHN**zimmer stellen,
 03 darauf hab ich !GAR! kein bock;
 04 dann [(...)]

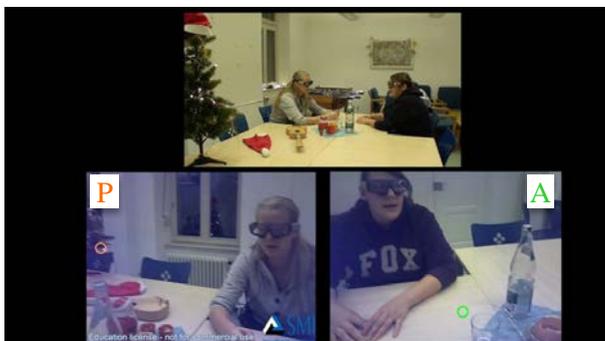


Abb. 3.2

05 P: [°h <<acc>ja aber GUCK_mal;



Abb. 3.3

06 **SCHAU** mal;



Abb. 3.4

07 DESCH **AUCH**_n plastikbaum;

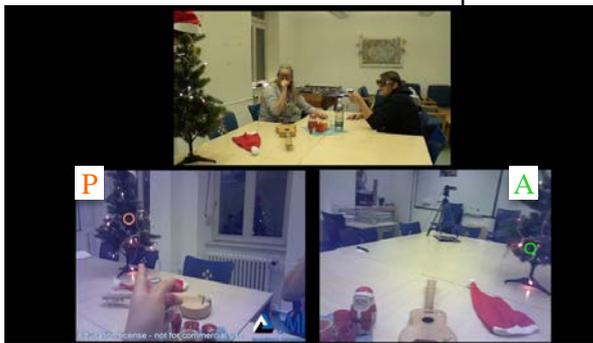


Abb. 3.5

Der Ausschnitt beginnt mit einem Selbstzitat von A, in welchem diese ihre Reaktion auf die mütterliche Idee eines künstlichen Weihnachtsbaums wiedergibt. Im Verlauf der Redewiedergabe (Z. 1-2) blickt sie zu ihrer Gesprächspartnerin P (Abb. 3.1/links), die ihrerseits A anblickt, während sie zuhört (Abb. 3.1/rechts), so dass Blickkontakt herrscht (Goodwin 1981; Kendon 1967, 1990).

Im Anschluss an das Selbstzitat formuliert A eine negative Bewertung (Z. 03). Als erster Paarteil einer Paarsequenz macht eine erste Bewertung eine – präferenzziell gleichlaufende – zweite Bewertung relevant (Pomerantz 1984), was im vorliegenden Fall bedeutet, dass die Gesprächspartnerin eine mit A übereinstimmende negative Bewertung formuliert. P hingegen projiziert eine abweichende Antwort, indem sie an die hier nicht als Affirmationssignal, sondern als Diskursmarker (Auer/Günthner 2005) verwendete Partikel *ja* die Disjunktion *aber* anschließt und damit eine Nicht-Übereinstimmung einleitet. Anstatt ihrer Gesprächspartnerin jedoch explizit zu widersprechen, lenkt P deren Aufmerksamkeit durch einen deiktischen Zeigeakt auf einen Plastikbaum, der Teil der Weihnachtsdekoration des Gemeinschaftsraums ist. Bevor sie die Zeigegeste im Verbund mit dem gestisch verwendeten (Fillmore 1997) Demonstrativpronomen *DESCH* (Z. 07) ausführt, reorientiert sie ihren Blick vorauslaufend, d.h. bereits beim Einatmen (Z. 05), auf den Verweisraum (Abb. 3.2/links). Gesprächspartnerin A wendet ihrerseits den Blick von P ab und schaut auf den Tisch (Abb. 3.2/rechts).

P's visuelle Reorientierung ist integraler Bestandteil des multimodalen Ablaufformats deiktischer Zeigehandlungen. Der Moment, in welchem P's Blick sich auf den Christbaum richtet, konstituiert den *Ziel-Blick-Nexus* innerhalb des Verweisraums. Die Blickpraktik der vorauslaufenden Verweisraum- und Zeigzielherstellung des Zeigenden dient vor allem der intrapersonellen Koordinierung (Depper-

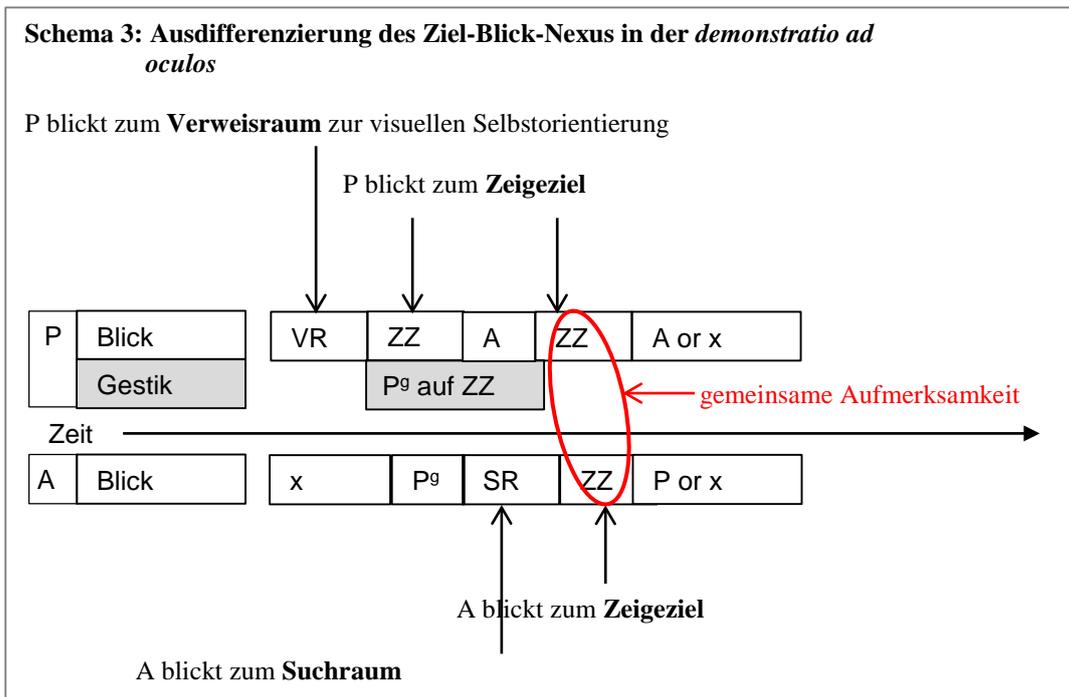
mann/Schmitt 2007). Sie kann von beiden Beteiligten aber auch interaktiv zur interpersonellen Koordination genutzt werden: Beispielsweise können Zeigende ihren Blick explizit und sichtbar für ihre Interaktionspartner auf ein Phänomen im umgebenden Raum richten, bevor sie gestisch darauf zeigen und zum *gaze following* einladen. Umgekehrt folgen Adressaten gelegentlich dem Blick ihrer Gesprächspartner, bevor diese auf das künftige Zeigeziel zeigen und dazu den Blick der Adressaten deiktisch einfordern. Und schließlich nutzen Zeigende die menschliche Fähigkeit zum *gaze following* auch, um ausschließlich mit dem Blick zu zeigen. Die Praktik des Zeigens mit dem Blick tritt insbesondere in Aktivitätskontexten auf, in denen die Beteiligten mit manuellen Aktivitäten (z.B. Kochen) beschäftigt sind, oder wenn das Zeigen verdeckt erfolgen soll (Stukenbrock 2015:177-192).

Wie zu sehen war, erfolgt P's visuelle Orientierung auf den Verweisraum und das künftige Zeigeziel, bevor sie die Zeigegeste ausführt (sichtbar in der Szenenkamera ihrer Adressatin: Abb. 3.3/rechts). P's Hand verlässt die Ruheposition erst, als sie die zweite der beiden Aufmerksamkeitsfokussierungen (*summons*) äußert (Z. 06: *SCHAU mal*). Diese bestehen aus einem visuellen Wahrnehmungsverb im Imperativ (Z. 05: *GUCK mal*; Z. 06: *SCHAU mal*), das explizit den Adressatenblick einfordert, bevor die multimodale Gestalt aus deiktischer Referenz und Zeigegeste realisiert wird. Während P's Blick auf den Plastikbaum gerichtet bleibt (Abb. 3.3/links), reorientiert A ihren Blick auf P's Gesicht und Augenregion (Abb. 3.3/rechts). In diesem Augenblick kann sie P's Blickrichtung wahrnehmen und dieser folgen, um herauszufinden, was es zu sehen gibt. P's Hand hat zu diesem Zeitpunkt für die Adressatin sichtbar die Ruheposition verlassen, wodurch eine Zeigegeste projiziert wird, die zusammen mit der Blickorientierung der Zeigenden von der Adressatin zur Lokalisierung des Zeigeziels genutzt werden kann.

Wie in Abb. 3.4 deutlich wird, verbindet P den deiktischen Gebrauch des Demonstrativpronomens *DESCH* (Z. 07) mit einer Zeigegeste. Während P's Blick weiterhin auf das Zeigeziel fixiert ist (Abb. 3.4/links), wendet A ihren Blick ab und orientiert sich auf den durch P's Geste und Blick doppelt indizierten Suchraum und das Zeigeziel (Abb. 3.4/rechts). Mit dem Äußerungsende erreicht A's Blick das Zeigeziel (Abb. 3.5/rechts). Da auch P's Blick noch immer auf den Baum gerichtet ist, entsteht ein Moment gemeinsamer Aufmerksamkeit, der in den Eye Tracking-Daten in den simultanen Blickfixationen von P und A auf demselben Objekt dokumentiert ist.

Oben wurde das Konzept des *Verweisraums* von dem des *Suchraums* abgegrenzt, um den Raum theoretisch und empirisch zu erfassen, den Zeigende systematisch durch eine bestimmte Blickpraktik herstellen. Das *Verweisraum*-Konzept trägt der Tatsache Rechnung, dass gemeinsame Aufmerksamkeitsräume und -foci Schritt für Schritt in der Zeit emergieren, so dass aus subjektiv wahrgenommenen und konzeptualisierten Phänomenen Objekte gemeinsamer Aufmerksamkeit und interaktiv hergestellter Intersubjektivität werden.

Zusammenfassend: Der in diesem Abschnitt diskutierte *Ziel-Blick-Nexus* konzeptualisiert zum einen a) die Beziehung zwischen einer Sprecherin P und einem visuell wahrnehmbaren Objekt, das diese mit der Adressatin teilen möchte, und zum anderen b) die Beziehung zwischen der Adressatin A und dem Objekt, das P wahrgenommen hat und durch eine deiktische Zeigehandlung in den visuellen Aufmerksamkeitsfokus von A rückt. Es wurde deutlich, dass eine Ausdifferenzierung in *Verweisraum* für den Raum, in dem P den Raum mit dem künftigen Zeigeeziel ins Auge fasst, und *Suchraum* für den Raum, in dem A das Zeigeeziel mit dem Blick auffindet, erforderlich ist. Schematisch lassen sich die Konzepte und die zeitliche Emergenz, in der sie im Zeigeakt relevant werden, folgendermaßen veranschaulichen:



4.3. Der Blick-zu-Blick-Nexus in der *demonstratio ad oculos*

Augenkontakt zwischen den Beteiligten tritt regelmäßig in der Kommunikation von Angesicht zu Angesicht auf und erfüllt unterschiedliche Funktionen, insbesondere Monitoring, Sprecherwechselregulierung, Ausdruck von Affiliation oder Aggression sowie Anerkennung.¹¹ In seinen grundlegenden Darlegungen zur fokussierten Interaktion (*focused interaction, face engagement*) hebt Goffman hervor, dass "an eye-to-eye ecological huddle tends to be carefully maintained, maximizing the opportunity for the participants to monitor one another's perceivings" (1963:95). Dass Interaktionsbeteiligte die Möglichkeit zu fortwährender wechselseitiger Wahrnehmung und wechselseitiger Wahrnehmung der Wahrnehmungen des Anderen ("one another's perceivings") aktiv suchen und aufrecht erhalten, bedeutet im Kontext deiktischer Praktiken etwas ganz Spezifisches. Gemäß der These dieses Beitrags stellt der Gebrauch verbaler Deiktika in der face-to-face-Inter-

¹¹ Kendon 1990; Goodwin 1981; Rossano 2012; Streeck 2014.

aktion eine Aufforderung der Blickzuteilung (*request for gaze*, vgl. Stukenbrock in press b) an die Adressatin dar. Folglich sind typische Blickmuster erwartbar, die die sequenzielle Implementierung dieser Aufforderung (*request*) und ihrer Folgehandlungen leisten. Insofern bedeutet der *Blick-zu-Blick-Nexus in deiktischen Praktiken* etwas anderes als Augenkontakt im Sinne von *mutual gaze* und ist nicht damit zu verwechseln. Das heißt nicht, dass es keine Überschneidungen zwischen den als *mutual gaze* beschriebenen Funktionen des Augenkontakts beim Sprecherwechsel (Kendon 1967) bzw. bei der Handlungskonstitution (*action formation*, Rossano 2012; Streeck 2014) und dem deiktischen *Blick-zu-Blick-Nexus* gibt. Doch dessen Spezifik besteht in der triadischen Relation zwischen den Interaktionsbeteiligten und einem von ihnen gemeinsam wahrzunehmenden Dritten sowie dem Prozess der Intersubjektivierung dieses gemeinsamen Wahrnehmungsprozesses als eines gemeinsam gewussten, der als solcher durch eine reziproke, auf die Wahrnehmung des Anderen bezogene Wahrnehmung, d.h. durch und als Wahrnehmungswahrnehmung,¹² intersubjektiviert wird. Dieser spezifische Moment der Wahrnehmungswahrnehmung, in welchem P wahrnimmt, dass A das durch den Zeigeakt angezeigte Wahrzunehmende wahrnimmt, wurde bereits in der "Kürbis"-Sequenz erkennbar: Dort drehte A sich um, um auf den Kürbis zu blicken, während P seinen Blick auf A richtete, um deren visuelle Aufmerksamkeit im Hinblick auf dieses Dritte, i.e. den Kürbis, zu überprüfen (siehe in Abschnitt 4.2, Beispiel 2, Abb. 2.5/links).

Die Analyse des folgenden Beispiels ("fett DRAN") konzentriert sich auf den Augenblick, in welchem die zeigende Teilnehmerin P ihren Blick auf die Augen ihrer Interaktionspartnerin A richtet und sieht, dass diese auf das Ziel von P's Zeigegeste blickt. In der Sequenz suchen zwei Studentinnen nach einem Buch in der Bibliothek. Nach anfänglichen Schwierigkeiten beim Auffinden des Buchs haben sie eine Bibliotheksangestellte kontaktiert. Der Ausschnitt setzt ein, als die beiden sich von der Bibliothekarin abwenden, um ihre Suche fortzusetzen. Die Perspektive der Zeigenden P ist rechts, die ihrer Interaktionspartnerin A ist links abgebildet.

Beispiel 4: "fett DRAN" (Bib03_00:46:49)

```

01   A:   also::
02   L:   (also [HIER HINTen) die TREppe hoch,
03   P:   [also HIER;
04       (--)
05   A:   geNAU;
06       oKEE;
07   L:   [(.....)

```

¹² Luhmann 1984:560; Hausendorf 2003; Stukenbrock 2009, 2015:85-90, 313-330.

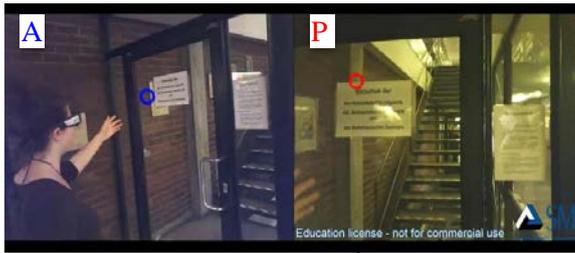


Abb. 4.1
08 P: [GUCK_ma **hier** stehst auch fett **DRAN**;

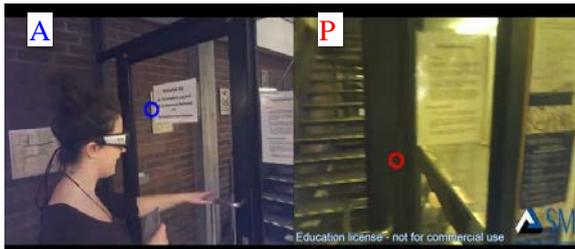


Abb. 4.2



Abb. 4.3

09 (1.3 _____)



Abb. 4.4

09 (1.3 _____)

10 A: ha;

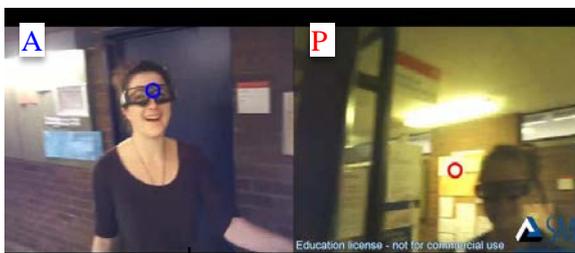


Abb. 4.5

11 P: **ha**;

12 A: ((kichert))

13 (-)

14 A: DANkeschö::n;



Abb. 4.6

15 ((laufen die Treppe hoch))

Nachdem die zwei Studentinnen sich von der Bibliothekarin (L) abgewandt haben, äußert A den Diskursmarker *also::* (Z. 01) und re-initiiert die Suche. In Überlappung mit einer nachgeschobenen Instruktion der Bibliothekarin (Z. 02: *also HIER HINten die TREPpe hoch*) beginnt auch P eine Äußerung mit *also* in der Funktion als Reformulierungindikator (Gülich/Kotschi 1987), auf den das deiktische Adverb *HIER* (Z. 03) folgt.

Wir überspringen den Klärungsprozess mit der Bibliothekarin und setzen die Analyse damit fort, dass P ein Schild an der Glastür entdeckt, das diejenige Bibliotheksabteilung ("Germanistische Linguistik") anzeigt, die sie gesucht haben. P äußert eine Fokussierungsaufforderung (*GUCK_MA*), auf die das deiktische Adverb *hier* folgt, das von einer mit offener Hand ausgeführten Zeigegeste auf das Türschild begleitet wird (vgl. zu verschiedenen Gestenformen Kendon 2004; Stukenbrock 2015:97-230). Währenddessen blickt auch P auf das Schild (Abb. 4.1/rechts). Wie auf der Komplementäraufnahme zu sehen ist (Abb. 4.1/links), befindet sich P im Kameraframe ihrer Adressatin A, die ebenfalls auf das Schild blickt (Abb. 4.1/links). Da P jedoch nicht zu A blickt, kann sie nicht wissen, dass A das Schild bereits entdeckt hat. P senkt ihren Arm, greift nach dem Knauf der Glastür, wie im Szenenvideo ihrer Interaktionspartnerin zu sehen ist (Abb. 4.2/links), und koordiniert ihren Blick mit ihrer Handbewegung (Abb. 4.2/rechts). Dadurch projiziert sie als Folgehandlung, dass sie die Tür öffnen und den gemeinsamen Weg in die entsprechende Richtung fortsetzen wird. Am Ende ihrer Äußerung und mit der Hand bereits am Türknauf wendet P sich zu A und richtet ihren Blick auf deren Augen (Abb. 4.3/rechts). Sie kann sehen und aus der Blickausrichtung ihrer Interaktionspartnerin inferieren, dass diese in diesem Augenblick auf das Türschild schaut. A's Eye Tracking-Daten dokumentieren, dass ihr Blick in der Tat auf das Türschild fixiert ist (Abb. 4.3/links).

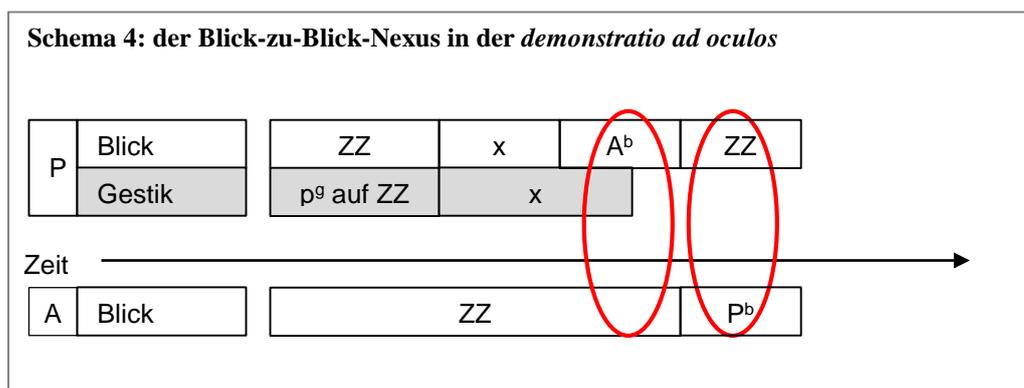
Der Augenblick, in welchem eine zeigende Person ihren Blick auf die Augen der Adressatin richtet, um deren visuelle Aufmerksamkeitsorientierung zu überprüfen, wird im vorliegenden Beitrag als *Blick-zu-Blick-Nexus* in der *demonstratio ad oculos* konzeptualisiert. Entscheidend gegenüber den andernorts funktional und sequenziell beschriebenen Momenten interaktiven Augenkontakts in der Gesprächsdyade ist dies gerade kein dyadischer Augenblick, sondern ein *triadischer Moment*, in welchem Zeigende wahrnehmen, dass ihre AdressatInnen eine dritte Größe wahrnehmen, i.e. das Zeigziel der deiktischen Zeigehandlung.¹³

¹³ Genauer müsste dieser Moment daher eigentlich *Blick-zu-Blick-Ziel-Nexus* heißen.

Nachdem P wahrgenommen hat, dass A das Türschild wahrgenommen hat, i.e. nach diesem Moment der Wahrnehmungswahrnehmung, richtet P ihren Blick wieder auf das Türschild (Abb. 4.4/rechts), während A ihren Blick nun umgekehrt vom Schild ab- und auf die Augenregion ihrer Interaktionspartnerin P wendet (Abb. 4.4/links). Es entsteht ein spiegelbildlicher Augenblick zum vorherigen *Blick-zu-Blick-Moment*, nur sind die Rollen nun vertauscht. Diejenige (A), die von ihrer Partnerin (P) wahrgenommen wurde, wird nun zur Wahrnehmenden (A), ihr Gegenüber wird zur Wahrgenommenen (P) bzw. noch präziser: zu derjenigen, deren Wahrnehmung von der Interaktionspartnerin wahrgenommen wird.

Im Folgenden spiegeln die beiden Beteiligten sich auch verbal. Zunächst äußert A einen *response cry* (Goffman 1981), der einen veränderten epistemischen Zustand (*epistemic stance*, cf. Heritage 2012) hinsichtlich des Wegproblems dokumentiert (Z. 10: *ha*), woraufhin P den *response cry* wiederholt (Z. 11: *ha*) und damit hörbar die kognitive und affektive Haltung teilt, die in A's initialem *response cry* inkarniert war. Darüber hinaus tauschen die beiden wechselseitige Blicke und Lächeln aus (Abb. 4.5/links und rechts), woraufhin sie eine *side-by-side*-Orientierung herstellen und gemeinsam die Treppe hinaufgehen (Abb. 4.6/links und rechts).

Diese Analyse validiert und verfeinert videobasierte Ergebnisse zur Rolle der Wahrnehmungswahrnehmung in deiktischen Zeigehandlungen (Hausendorf 2003; Stukenbrock 2009, 2015) auf der Grundlage von Eye Tracking-Daten, die zeitlich hochauflösende Beobachtungen zur interaktiven Feinkoordinierung von Sprecher- und Adressatenblick bei der Intersubjektivierung von Wahrnehmung ermöglichen. So lässt sich der Moment, in dem der Adressatenblick auf das Ziel gerichtet und der Sprecherblick zwecks Adressatenmonitoring auf die Adressatenaugen orientiert ist, exakt erfassen. Schematisch werden die Analyse-Ergebnisse des *Blick-zu-Blick-Nexus* in deiktischen Praktiken wie folgt zusammengefasst:



In der Sequenz zeigt und blickt P zunächst auf das Zeigziel, dann richtet sie ihren Blick auf die Tür (im Schema mit dem Platzhalter "x" notiert) und greift mit der Hand nach dem Türknauf, bevor sie ihren Blick schließlich der Adressatin A zuwendet. In diesem Augenblick kann sie sehen, dass A ebenfalls auf das Türschild blickt, das zugleich das Zeigziel von P darstellt, das A allerdings bereits seit einer Weile anschaut. Sequenziell ist dies nicht der Standardfall, im Gegenteil blicken Adressatinnen in Fällen, in denen Zeigende einen neuen Aufmerksamkeitsfokus herstellen, zunächst zum Zeigenden und dessen Geste und/oder orientieren

sich an dessen Blickausrichtung, um mit Hilfe dieser körperlichen Zeigemittel den Suchraum herzustellen und das Zeigeziel darin aufzufinden (vgl. Abschnitt 4.1).

Der entscheidende Moment der im vorliegenden Abschnitt analysierten Sequenz wird dadurch konstituiert, dass P die Wahrnehmung der Adressatin A wahrnehmen kann, dadurch dass sie deren Blickorientierung überprüft (als A^b notiert) und visuell verifiziert, dass A sieht, was es zu sehen gibt und was P bereits gesehen hat. Bemerkenswerterweise tritt in der vorliegenden Sequenz eine spiegelbildliche Synchronisierung der beiden Interaktantinnen ein, als P ihren Blick von A wieder auf das Türschild und A umgekehrt ihren Blick vom Türschild weg und auf P's Augen richtet, um zu schauen, wo P hinschaut (P^b). Anhand von P's neuerlicher Blickausrichtung kann A inferieren, dass P das Türschild sieht. Dies markiert den Übergangspunkt von einer gemeinsamen Raumorientierungssequenz zu einer *stance taking*-Sequenz, in der beide humorvoll anerkennen, dass sie zuvor "den Wald vor lauter Bäumen" nicht gesehen haben.

5. Diskussion: Was mobiles Eye Tracking für die Analyse von Blickpraktiken bei der *demonstratio ad oculos* in natürlicher Interaktion leistet

Gegenstand dieses Beitrags waren spezifische Blickpraktiken von SprecherInnen und ihren AdressatInnen, die systematisch auftreten, wenn deiktisch auf Sichtbares im Wahrnehmungsraum referiert und ein gemeinsamer Aufmerksamkeitsfokus auf das entsprechende Phänomen hergestellt wird. Der analytische Fokus lag auf den Blickpraktiken der Zeigenden P und ihrer Adressaten A sowie auf der interpersonellen Feinkoordinierung wechselseitiger Wahrnehmung und Wahrnehmungswahrnehmung, die sich in rekurrenten Blickmustern manifestieren. Diese Aspekte wurden bislang weder in ihrem zeitlichen Verhältnis zur online-Emergenz verbaler Äußerungen noch im Zusammenspiel mit leiblichen Praktiken wie Zeigegesten, Bewegungen und Konfigurationen der Beteiligten im Raum untersucht.

Die Analysen basieren auf mobilen Eye Tracking-Daten, die außerhalb des Laboratoriums aufgenommen wurden, natürliche Interaktionen dokumentieren und dadurch ökologisch valide sind. Damit leistet diese Untersuchung auch einen methodologischen Beitrag dazu, wie innovative, bislang fast ausschließlich im Labor für experimentelle Untersuchungen eingesetzte Technologien im Rahmen konversationsanalytischer Forschung für *in situ*-Studien zur sozialen Interaktion genutzt werden können.

Festzuhalten bleibt, dass mobile Eye Tracking-Aufnahmen nicht in naturalistischer Weise die "reale" Perspektive der Beteiligten wiedergeben. Dies gilt sowohl für das von der Szenekamera aufgenommene Video als auch für das Eye Tracking durch die eingebauten Kameras. Das aus diesen beiden Aufnahmen fusionierte Video, in dem das Video der Szenekamera mit dem beweglichen Cursor der Eye Tracking-Aufnahmen überlegt ist, bietet einen künstlichen Bildrahmen, in dem das Blickfeld der Beteiligten auf ein rechteckiges Fenster reduziert und einen wesentlichen Anteil peripherer Wahrnehmung der Beteiligten buchstäblich abschneidet (vgl. Abschnitt 2). Doch auch wenn mobile Eye Tracking-Aufnahmen keine naturalistische Aufzeichnung der emischen Perspektive der Beteiligten darstellen, bieten sie nichtsdestotrotz eine neuartige, beteiligtenzentrierte visuelle

Perspektive auf das, was sich in Echtzeit in sozial und räumlich situierter, leiblicher Interaktion vollzieht. Durch diese Art der Aufnahmen entsteht ein vollkommen neues Datenset, das sich radikal von den bisher für multimodale Interaktionsanalysen verwendeten traditionellen Videoaufnahmen aus Beobachterperspektive unterscheidet, ungewohnte Seh- und Analysepraktiken erfordert und dadurch theoretische und methodologische Reflexionen über die besonderen Leistungen und Probleme beider Aufnahmetypen erzwingt.

Die Detailanalysen zu Blickpraktiken in deiktischen Referenzierungshandlungen (Abschnitt 4) haben gezeigt, dass dem Blick eine herausgehobene Rolle in der Koordinierung von Verbaläußerungen, Gestik, Körperorientierung und Bewegung im Raum zukommt. Die im Folgenden zusammengefassten Blickpraktiken bilden einen konstitutiven Bestandteil deiktischer Referenz in der Interaktion von Angesicht zu Angesicht und treten systematisch in spezifischen sequenziellen Positionen bei der Herstellung gemeinsamer Aufmerksamkeitsfoci auf. Sie wurden als *Blick-Gesten-Nexus*, *Blick-Ziel-Nexus* und *Blick-zu-Blick-Nexus* in der *demonstratio ad oculos* konzeptualisiert und werden als Ergebnis der empirischen Analysen im Folgenden noch einmal differenziert aufgeführt.

Zunächst bildet der *Blick-Ziel-Nexus* einen Kernbestandteil der Blickpraktiken sowohl von P als auch von A:

P► ZZ = P blickt zum Zeigeziel
 A► ZZ = A blickt zum Zeigeziel

Allerdings vereinfacht dieses Konzept, was sich Augenblick-für-Augenblick in der Interaktion zwischen A und P in deiktischen Zeigehandlungen vollzieht, indem wesentliche Zwischenschritte unter dem Begriff *Zeigeziel* subsumiert werden. Auf der Mikroebene einer zeitlich und räumlich hochaufgelösten Detailanalyse wurde nachgewiesen, dass ein Zeigender P sich zunächst selbst im Wahrnehmungsraum orientiert, einen potenziellen *Verweisraum* (VR) mit einem möglichen Zeigeziel (ZZ) in den Blick nimmt, bevor er diesen Raum bzw. das Zeigeziel für die Adressatin A deiktisch und gestisch relevant setzt. Die in den Eye Tracking-Daten klar nachzuweisende visuelle Prä-Orientierung kann zeitlich deutlich vorauslaufen, was darauf hindeutet, dass P den Zeigeakt zeitlich flexibel in das emergente Gesamtgeschehen einbettet:

P► VR = P blickt zum Verweisraum

Umgekehrt gilt, dass Adressaten, wenn sie ihren durch P's Zeigegeste angeleiteten Blick auf ein Zeigeziel im Wahrnehmungsraum richten, zunächst den Suchraum (SR) herstellen, bevor sie das eigentliche Zeigeziel (ZZ) auffinden, was sich unter Umständen als problematisch erweisen kann:

A► SR = A blickt zum Suchraum

Die Herstellung eines gemeinsamen visuellen Aufmerksamkeitsfokus bildet den Kernmoment deiktischer Zeigehandlungen in der face-to-face-Interaktion. Dieser kommt dadurch zustande, dass beide Beteiligte auf dasselbe Phänomen blicken und dies auch voneinander wissen. Eine gemeinsame Aufmerksamkeitsorientie-

ung kann simultan oder zeitlich versetzt auftreten, gemeinsame Aufmerksamkeit (*joint attention*) kann daher nicht mechanisch aus den Eye Tracking-Daten als Fixation beider Blick-Cursor auf demselben Phänomen abgelesen bzw. durch das Errechnen von *gaze cross recurrence*-Instanzen (Jerma/Nüssli 2012) automatisiert werden. Solche Fixationen können auch zufällig sowie als intersubjektiv nicht vergemeinschaftete Momente in den Daten auftreten. Diese Zufallsmomente sind von gemeinsamer und als gemeinsam gewusster Wahrnehmung zu unterscheiden. Letztere ist hier gemeint. Sie kann in der Interaktion in unterschiedlichen Zeitlichkeitsverhältnissen (simultan, teil-simultan, zeitlich versetzt) emergieren:

P► ZZ = P blickt zum Zeigeziel
 A► ZZ = A blickt zum Zeigeziel

Der zweite wichtige Komplex interagierender Blickpraktiken bei der *demonstratio ad oculos* ist der *Blick-Gesten-Nexus*. Er geht über die wechselseitige visuelle Orientierung der Beteiligten hinaus und betrifft spezifische Momente, in denen A den Blick auf P bzw. bestimmte Körperregionen von P (und dessen Gesten) richtet, um vektorielle Informationen zum Auffinden des Zeigeziels zu erhalten:

A► P^g = A blickt zu P's Geste

Dabei ist festzuhalten, dass videobasierte Aussagen über Blickorientierungen auf die Geste des Interaktionspartners anhand von Eye Tracking-Daten einer genaueren Prüfung unterzogen werden müssen, da Adressaten offenbar auch andere Quellen zur Vektorbildung wie z.B. den Blick des Zeigenden anstelle der Geste nutzen und darüber hinaus vektorielle Informationen auch ohne Blickfixationen auf die Geste des Zeigenden erhalten.

Konstitutiv für die systematische und interaktive Organisation des Blicks in deiktischen Zeigesequenzen ist drittens der *Blick-zu-Blick-Nexus*. Wie bei der Organisation des Sprecherwechsels oder bestimmter Handlungen und Aktivitäten blicken P und A regelmäßig in die Augen ihres Interaktionspartners. Wenn P und A gleichzeitig mit ihrem Blick die Augen ihres Gegenübers aufsuchen, entsteht Augenkontakt:

A^b ◄.....► P^b = A blickt zu P's Augen and P blickt zu A's Augen

Das Spezifische am Blickkontakt zwischen P und A bei der *demonstratio ad oculos*, wodurch dieser sich vom Blickkontakt beim Sprecherwechsel unterscheidet, liegt darin, dass deiktisch eingebettete Blickpraktiken intrinsisch darauf zugeschnitten sind, Informationen über die auf ein Drittes orientierte Wahrnehmung des Interaktionspartners zu gewinnen, i.e. das Ziel gemeinsamer und als gemeinsam gewusster Wahrnehmung (*joint attention*). Diese stellt eine triadische Relation dar, die durch eine gemeinsames *timing* und *spacing* verbaler und leiblicher Praktiken der interpersonellen Koordinierung von Handlung, Wahrnehmung und Kognition zustande kommt.

Als Fortführung und Differenzierung früherer Untersuchungen zu Deixis, Zeigegesten und Blick (Stukenbrock 2009, 2010, 2015) haben die hier vorgelegten

Eye Tracking-Analysen systematisch wiederkehrende Momente in der Interaktion identifiziert, die ein gesteigertes, für die Interaktion konstitutives perzeptorisches Interesse Egos an Alters Wahrnehmungen, i.e. einen auf Alters visuelle Wahrnehmung gerichteten meta-perzeptorischen Gebrauch der eigenen visuellen Wahrnehmung Egos dokumentieren. Dieses reziproke Wahrnehmungsinteresse der Beteiligten setzt sich in spezifischen Blickpraktiken interaktiv ins Werk, wird an deren auf den Blick des Anderen gerichteten Blick (öffentlich/für die Beteiligten selbst) sichtbar und dadurch der sequenziellen Analyse leiblicher Interaktionspraktiken zugänglich:

A► P^b = A blickt zu P's Augen
 P► A^b = P blickt zu A's Augen

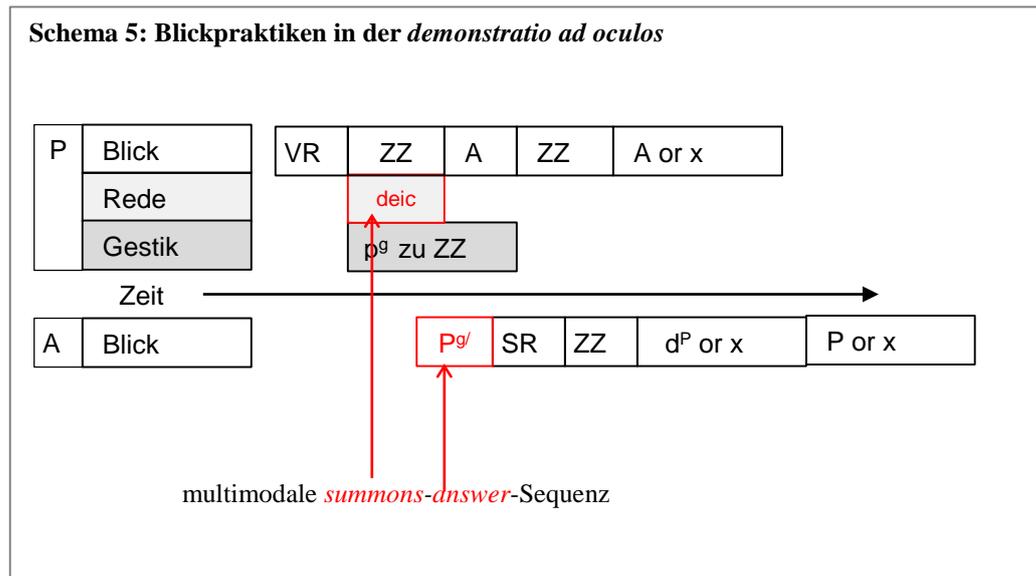
A's visuelle Orientierung auf P's Blick und/oder Geste in der *demonstratio ad oculos* tritt nach und als Antwort auf eine deiktische Fokussierungs- bzw. Blickaufforderung durch ein von P gestisch gebrauchtes Deiktikon auf. Dessen Gebrauch macht für A erforderlich, zusätzliche Informationen aus P's leiblichen Praktiken zu gewinnen und erzwingt daher eine Blickzuwendung. A wird folglich auf P's Geste oder alternativ auf dessen Blickorientierung achten.

Für P demgegenüber besteht die Systematik darin, zu A zu blicken und insbesondere auf ihren Blick zu achten, kurz nachdem er eine deiktische Blickaufforderung geäußert hat. A's Blick auf sich zu orientieren ist nicht wie bei Goodwins *restarts* (1981) das eigentliche Ziel von P's deiktischer Blickeinforderung, sondern diese dient dazu, die visuelle Aufmerksamkeit der Adressatin durch eine Zeigegeste auf ein Drittes zu lenken. P wird daher das leibliche Zeigen genau mit dem Aufmerksamkeitsverhalten seiner Adressatin koordinieren um sicherzustellen, dass A in dem Augenblick, indem sie sich P zuwendet, dessen Zeigehandlung sehen und befolgen kann. Der Kontrollblick, mit dem P das Blickverhalten von A überprüft, dient P dazu wahrzunehmen, ob auch A wahrnimmt, was P ihr zur Wahrnehmung gezeigt hat, so dass gemeinsame Aufmerksamkeit und wechselseitiges Verstehen – Intersubjektivität angesichts eines Dritten – interaktiv hergestellt werden.

Allerdings blickt P nicht immer zu A, um deren Blickorientierung vis-à-vis des Zeigziels zu kontrollieren. Sind Beteiligte in manuelle Aktivitäten involviert, die ihre ganze Aufmerksamkeit erfordern, und sind sie sich wechselseitig dieser Involviertheit bewusst, kann auf Blickpraktiken wechselseitiger Wahrnehmungskontrolle verzichtet werden. Meta-perzeptorische Blickpraktiken, die auf die visuelle Wahrnehmung des Anderen orientiert sind, treten nicht nur in spezifischen sequenziellen Positionen auf, sie sind darüber hinaus aktivitäts- und kontextsensitiv (Stukenbrock 2015).

Um die detaillierten Beobachtungen zur interaktiven Gestaltung der Blickpraktiken beim Zeigen in eine Zusammenschau zu integrieren können, komme ich abschließend auf die zentrale Funktion der Sprache, genauer, der deiktischen Formen, in multimodalen Referenzierungshandlungen zurück. Der vorliegende Beitrag hat zwei Hauptthesen formuliert: erstens, dass sprachliche Deiktika mit Zeigegesten zur visuellen Aufmerksamkeitsorientierung des Adressaten auf sichtbare Phänomene im gemeinsamen Wahrnehmungsraum verbunden werden, und zweitens, dass der gestische Gebrauch sprachlicher Deiktika notwendigerweise den

Blick des Adressaten einfordert. Innerhalb einer deiktischen Referenzierungshandlung vollzieht sich folglich eine multimodale *summons-answer*-Sequenz die eine systematische Position im sequenziellen Ablaufformat besetzt und rekurrente Blickmuster von P und A implementiert, die aktivitäts- und kontextsensitiv sind und in temporaler Feinabstimmung mit dem emergierenden verbalen und leiblichen Interaktionsgeschehen eingesetzt werden. Die verbaldeiktische Form kann wie folgt in das Gesamtschema integriert werden:



6. Literatur

- Argyle, Michael / Cook, Mark (1976): *Gaze and Mutual Gaze*. London: Cambridge University Press.
- Auer, Peter / Günthner, Susanne (2005): Die Entstehung von Diskursmarkern im Deutschen - ein Fall von Grammatikalisierung? In: Leuschner, Thorsten / Mortelmans, Tanja / De Groot, Sarah (Hg.), *Grammatikalisierung im Deutschen*. Berlin / New York: de Gruyter, 335-362.
- Brooks, Rechele / Meltzoff, Andrew N. (2014): Gaze following: A mechanism for building social connections between infants and adults. In: Mikulincer, Mario / Shaver, Phillip R. (Hg.), *Mechanisms of social connection: From brain to group*. Washington, D.C.: American Psychological Association, 167-183.
- Bühler, Karl (1965[1934]): *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Stuttgart: Gustav Fischer Verlag.
- Clark, Alan T. / Gergle, Darren (2011): *Mobile Dual Eye-Tracking Methods: Challenges and Opportunities*. Paper presented at DUET 2011: Dual Eye Tracking Workshop at ECSCW 2011 (Aarhus, Denmark, September 25, 2011).
- Clark, Alan T. / Gergle, Darren (2012): *Know what I'm talking about? Dual eye-tracking in multimodal reference resolution*. Paper presented at DUET 2012: Dual Eye Tracking workshop at CSCW 2012 (Bellevue, WA, February 12, 2012).

- Clark, Herbert H. / Schreuder, Robert / Buttrick, Samuel (1983): Common Ground and the Understanding of Demonstrative Reference. In: *Journal of Verbal Learning and Verbal Behavior* 22(2), 245-258.
- Couper-Kuhlen, Elisabeth / Selting, Margret (2001): *Studies in Interactional Linguistics*. Amsterdam: John Benjamins.
- Dao, Anh Nhi (in Vb.): Blickverhalten bei sprachlichen Lokalisierungen in der Interaktion. PhD thesis Freiburg/Br.
- Deppermann, Arnulf (2013): Multimodal interaction from a conversation analytic perspective. In: *Journal of Pragmatics* 46, 1-7.
- Deppermann, Arnulf (2018): Instruction practices in German driving lessons: Differential uses of declaratives and imperatives. In: *International Journal of Applied Linguistics* 28 (2), Special Issue on Driving School Interaction, 265-282.
- Deppermann, Arnulf / Schmitt, Reinhold (2007): Koordination. Zur Begründung eines neuen Forschungsgegenstandes. In: Schmitt, Reinhold (Hg.), *Koordination. Analysen zur multimodalen Interaktion*. Tübingen: Narr, 15-54.
- De Stefani, Elwys (2010): Reference as an interactively and multimodally accomplished practice. Organizing spatial reorientation in guided tours. In: Pettorino, Massimo et al. (Hg.), *Spoken Communication*. Newcastle upon Tyne: Cambridge Scholars Publishing, 137-170.
- De Stefani, Elwys (2013): The collaborative organisation of next actions in a semiotically rich environment. Shopping as a couple. In: Haddington, Pentti / Mondada, Lorenza / Nevile, Maurice (Hg.), *Interaction and Mobility. Language and the Body in Motion*. Berlin/Boston: De Gruyter, 123-151.
- De Stefani, Elwys / Gazin, Anne-Danièle (2014): Instructional sequences in driving lessons: Mobile participants and the temporal and sequential organization of actions. In: *Journal of Pragmatics* 65, 63-79.
- De Stefani, Elwys / Mondada, Lorenza (2014): Reorganizing mobile formations: When 'guided' participants initiate reorientation in guided tours. In: *Space and Culture* 17 (2), 157-175.
- Diessel, Holger (2006): Demonstratives, joint attention, and the emergence of grammar. In: *Cognitive Linguistics* 17 (4), 463-489.
- Duchowski, Andrew T. (2003): *Eye Tracking Methodology. Theory and Practice*. Berlin: Springer.
- Ehmer, Oliver (2011): *Imagination und Animation. Die Herstellung mentaler Räume durch animierte Rede*. Berlin, Boston: de Gruyter.
- Eriksson, Mats (2009): Referring as Interaction: On the Interplay Between Linguistic and Bodily Practices. In: *Journal of Pragmatics* 41, 240-262.
- Fricke, Ellen (2007): *Origo, Geste und Raum. Lokaldeixis im Deutschen*. Berlin: de Gruyter.
- Fillmore, Charles (1997): *Lectures on Deixis*. Stanford, CA: CSLI Publ.
- Garfinkel, Harold (1967): *Studies in Ethnomethodology*. Englewood Cliffs, Nj: Prentice-Hall.
- Gergle, Darren / Clark, Alan T. (2010): Effects of Shifting Spatial Context on Referential Form. In: *Proceedings of the 20th Annual Meeting of the Society for Text and Discourse*.
- Gergle, Darren / Clark, Alan T. (2011). See What I'm Saying? Using Dyadic Mobile Eye-Tracking to Study Collaborative Reference. In: *Proceedings of CSCW 2011*, 435-444.

- Goffman, Erving (1963): *Behavior in Public Places. Notes on the Social Organization of Gatherings*. New York: The Free Press.
- Goffman, Erving (1981): *Forms of Talk*. Philadelphia: University of Pennsylvania Press.
- Goodwin, Charles (1980): Restarts, Pauses, and the Achievement of a State of Mutual Gaze at Turn-Beginning. In: *Sociological Inquiry* 50 (3-4), 272-302.
- Goodwin, Charles (1981): *Conversational Organization: Interaction Between Speakers and Hearers*. New York: Academic Press.
- Goodwin, Charles (2000): Action and embodiment within situated human interaction. In: *Journal of Pragmatics* 32, 1489-1522.
- Goodwin, Charles (2003): Pointing as Situated Practice. In: Kita, Sotaro (Hg.), *Pointing. Where Language, Culture, and Cognition Meet*. NJ: Erlbaum, 217-241.
- Goodwin, Charles (2007): Interactive Footing. In: Holt, Elizabeth / Clift, Rebecca (Hg.), *Reporting talk. Reported speech in interaction*. Cambridge: Cambridge University Press, 16-46.
- Gülich, Elisabeth / Kotschi, Thomas (1987): Reformulierungshandlungen als Mittel der Textkonstitution. Untersuchungen zu französischen Texten aus mündlicher Kommunikation. In: Motsch, Wolfgang (Hg.), *Satz, Text, sprachliche Handlung*. Berlin: Akademie Verlag, 199-261.
- Gullberg, Marianne / Kita, Sotaro (2009): Attention to Speech-Accompanying Gestures: Eye Movements and Information Uptake. In: *Journal of Nonverbal Behavior* 33, 251-277.
- Haddington, Pentti / Mondada, Lorenza / Neville, Maurice (2013): *Interaction and Mobility. Language and the Body in Motion*. Berlin, Boston: de Gruyter.
- Hanks, William F. (1990): *Referential practice. Language and lived space among the Maya*. Chicago: University of Chicago Press.
- Hanna, Joy E. / Brennan, Susan (2007): Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation. In: *Journal of Memory and Language* 57, 596-651.
- Hausendorf, Heiko (2003): Deixis and speech situation revisited. The mechanism of perceived perception. In: Lenz, Friedrich (Hg.), *Deictic Conceptualisation of Space, Time and Person*. Amsterdam: John Benjamins, 249-269.
- Hausendorf, Heiko / Mondada, Lorenza / Schmitt, Reinhold (Hg.) (2012): *Raum als interaktive Ressource*. Tübingen: Narr.
- Heath, Christian (1986): *Body movement and speech in medical interaction*. Cambridge: Cambridge University Press.
- Hepburn, Alexa / Bolden, Galina (2012): The Conversation Analytic Approach to Transcription. In: Sidnell, Jack / Stivers, Tanya (Hg.): *The Handbook of Conversation Analysis*. Oxford: Wiley Blackwell, 57-76.
- Heritage, John (2012): Epistemics in Action: Action Formation and Territories of Knowledge. In: *Research on Language and Social Interaction* 45(1), 1-29.
- Hindmarsh, Jon / Heath, Christian (2000): Embodied reference: A study on deixis in workplace interaction. In: *Journal of Pragmatics* 32 (12), 1855-1878.
- Holler, Judith / Kendrick, Kobin (2015): Unaddressed participants' gaze in multi-person interaction: Optimizing reciprocity. In: *Frontiers in Psychology* 6(98), 1-14.

- Horsley, Mike / Eliot, Matt / Knight, Bruce Allen / Reilly, Ronan (Hg.) (2014): *Current Trends in Eye Tracking Research*. Heidelberg: Springer.
- Keevallik, Leelo (2013): Here in time and space: Decomposing movement in dance instruction. In: Haddington, Pentti / Mondada, Lorenza / Neville, Maurice (Hg.), *Interaction and Mobility. Language and the Body in Motion*. Berlin, Boston: de Gruyter, 345-370.
- Kendon, Adam (1967): Some function of gaze-direction in social interaction. In: *Acta Psychologica* 26, 22-63.
- Kendon, Adam (1990): *Conducting Interaction. Patterns of Behavior in Focused Encounters*. Cambridge / New York: Cambridge University Press.
- Kendon, Adam (2004) *Gesture. Visible Action as Utterance*. Cambridge: Cambridge University Press.
- Kita, Sotaro (Hg.) (2003): *Pointing. Where Language, Culture, and Cognition Meet*. Mahwah, NJ: Erlbaum.
- Land, Michael F. (2006): Eye movement and the control of actions in everyday life. In: *Progress in Retinal and Eye Research* 25, 296-324.
- Land, Michael F. / Tatler, Benjamin W. (2009): *Looking and Acting. Vision and eye movements in natural behaviour*. Oxford: Oxford University Press.
- Levinson, Stephen C. (2000): *Pragmatik*. 3. Aufl. Tübingen: Niemeyer.
- Levinson, Stephen C. (2004): Deixis. In: Horn, Laurence R. (Hg.), *The Handbook of Pragmatics*. Oxford: Blackwell, 97-121.
- Levinson, Stephen C. (2012): Action Ascription and Formation. In: Sidnell, Jack / Stivers, Tanya (Hg.), *The Handbook of Conversation Analysis*. Oxford: Wiley Blackwell, 103-130.
- Luhmann, Niklas (1984): *Soziale Systeme. Grundriß einer allgemeinen Theorie*. Frankfurt/M.: Suhrkamp.
- Mondada, Lorenza (2007): Multimodal Resources for Turn-Taking: Pointing and the Emergence of Possible Next Speakers. In: *Discourse Studies* 9(2), 195-226.
- Mondada, Lorenza (2012a): Deixis: an integrated interactional multimodal analysis. In: Bergmann, Pia / Brenning, Jana / Pfeiffer, Martin / Reber, Elisabeth (Hg.), *Prosody and Embodiment in Interactional Grammar*. Berlin / Boston: de Gruyter, 173-206.
- Mondada, Lorenza (2012b): The Conversation Analytic Approach to Data Collection. In: Sidnell, Jack / Stivers, Tanya (Hg.), *The Handbook of Conversation Analysis*. Oxford: Wiley Blackwell, 32-56.
- Pomerantz, Anita (1984): Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In: Atkinson, J. Maxwell / Heritage, John (Hg.), *Structures of social action: Studies in conversation analysis*. Cambridge: Cambridge University Press, 57-101.
- Quine, Willard Orman van (1960): *Word and Object*. Cambridge: Technology Press of the Massachusetts Institute of Technology.
- Rossano, Federico (2012): *Gaze Behavior in Face-to-Face Interaction*. Nijmegen: Max Planck Institute for Psycholinguistics Series.
- Schegloff, Emanuel A. (1968): Sequencing in Conversational Openings. In: *American Anthropologist* 70(6), 1075-1095.
- Schegloff, Emanuel A. (1996): Confirming Allusions: Toward an Empirical Account of Action. In: *American Journal of Sociology* 102(1), 161-216.

- Schegloff, Emanuel A. (2007): *Sequence Organization in Interaction*. Cambridge: Cambridge University Press.
- Schegloff, Emanuel A. / Sacks, Harvey (1973): *Opening Up Closings*. In: *Semiotica* 8, 289-327.
- Schmitt, Reinhold (Hg.) (2007): *Koordination. Analysen zur multimodalen Interaktion*. Tübingen: Narr.
- Schwarzkopf, Sarah / Büchner, Simon / Hölscher, Christoph / Konieczny, Lars (2017): *Perspective tracking in the real world. Gaze angle analysis in a collaborative wayfinding task*. In: *Spatial Cognition & Computation* 17:1-2, 143-162.
- Selting, Margret et al. (2009): *Gesprächsanalytisches Transkriptionssystem 2 (GAT 2)*. In: *Gesprächsforschung* 10, 353-402.
- Sidnell, Jack / Stivers, Tanya (Hg.) (2013): *The Handbook of Conversation Analysis*. Oxford: Wiley Blackwell.
- Streeck, Jürgen (1993): *Gesture as communication I: Its coordination with gaze and speech*. In: *Communication Monographs* 60, 275-299.
- Streeck, Jürgen (2002): *Grammars, Words, and Embodied Meanings: On the Uses and Evolution of So and Like*. In: *Journal of Communication* 52 (3), 581-596.
- Streeck, Jürgen (2009): *Gesturecraft. The manufacture of meaning*. Amsterdam / Philadelphia: John Benjamins.
- Streeck, Jürgen (2014): *Mutual gaze and recognition: Revisiting Kendon's "Gaze direction in two-person conversation"*. In: Seyfeddinipur, Mandana / Gullberg, Marianne (Hg.), *From Gesture in Conversation to Visible Action as Utterance. Essays in honor of Adam Kendon*. Amsterdam / Philadelphia: Benjamins, 35-55.
- Streeck, Jürgen / Goodwin, Charles / LeBaron, Curtis (Hg.) (2011): *Embodied Interaction. Language and the Body in the Material World*. Cambridge: Cambridge University Press.
- Stukenbrock, Anja (2008): *"Wo ist der Hauptschmerz?" – Zeigen am eigenen Körper in der medizinischen Kommunikation*. In: *Gesprächsforschung* 9, 1-33.
- Stukenbrock, Anja (2009): *Referenz durch Zeigen: Zur Theorie der Deixis*. In: *Deutsche Sprache*. 37, 289-315.
- Stukenbrock, Anja (2010): *Überlegungen zu einem multimodalen Verständnis der gesprochenen Sprache am Beispiel deiktischer Verwendungsweisen des Ausdrucks "so"*. In: *InLiSt – Interaction and Linguistic Structures* 47.
- Stukenbrock, Anja (2012): *Imagined spaces as a resource in interaction*. In: *Bulletin Suisse de Linguistique Appliquée* 96, 141-161.
- Stukenbrock, Anja (2014a): *Take the words out of my mouth: Verbal instructions as embodied practices*. In: *Journal of Pragmatics* 65, 80-102.
- Stukenbrock, Anja (2014b): *Pointing to an 'empty' space: Deixis am Phantasma in face-to-face interaction*. In: *Journal of Pragmatics* 74, 70-93.
- Stukenbrock, Anja (2015): *Deixis in der face-to-face-Interaktion*. Berlin / Boston: de Gruyter.
- Stukenbrock, Anja (2016): *Deiktische Praktiken: Zwischen Interaktion und Grammatik*. In: Deppermann, Arnulf / Feilke, Helmut / Linke, Angelika (Hg.), *Sprachliche und kommunikative Praktiken. Jahrbuch 2015 des Instituts für Deutsche Sprache*. Berlin / Boston: de Gruyter, 81-126.
- Stukenbrock, Anja (im Druck a): *Mobile dual eye-tracking in face-to-face interaction: the case of deixis and joint attention*. In: Brône, Geert / Oben, Bent (Hg.),

Eye-tracking in interaction. Studies on the role of eye gaze in dialogue. Amsterdam: Benjamins.

Stukenbrock, Anja (im Druck b): Forward-looking: Where do we go with multi-modal projections? In: Deppermann, Arnulf / Streeck, Jürgen (Hg.), Modalities and Temporalities: Convergences and divergences of bodily resources in interaction.

Stukenbrock, Anja / Dao, Anh Nhi (im Druck): Joint attention in passing. What dual mobile eye-tracking reveals about gaze in coordinating embodied activities on a market. In:

Tomasello, Michael (2008): Origins of Human Communication. Cambridge / Massachusetts: MIT Press.

Weiß, Clarissa / Auer, Peter (2016): Das Blickverhalten des Rezipienten bei Sprecherhäsitationen. Eine explorative Studie. In: Gesprächsforschung 17, 132-167.

7. Konventionen für die Transkription des Blickverhaltens

In den Analysen wird P konsistent für die zeigenden TeilnehmerInnen und A für die AdressatInnen verwendet.

der Ziel-Blick-Nexus

Der Ziel-Blick-Nexus bedeutet, dass P oder A auf den relevanten Raumausschnitt bzw. das entsprechende Zeigeziel blicken. Folgende Differenzierungen sind erforderlich:

| | | |
|---------|---|------------------------------|
| P -> VR | = | P blickt auf den Verweisraum |
| A -> SR | = | A blickt auf den Suchraum |
| P -> ZZ | = | P blickt zum Zeigeziel |
| A -> ZZ | = | A blickt zum Zeigeziel |

der Körper-Blick-Nexus, genauer: der Gesten-Blick-Nexus

| | | |
|---------------------|---|----------------------------|
| A -> P ^k | = | A blickt zu P (P's Körper) |
| A -> P ^g | = | A blickt auf P's Geste |

der Blick-zu-Blick-Nexus

| | | |
|-----------------------------------|---|---|
| A -> P ^b | = | A fokussiert auf P's Augenregion |
| P -> A ^b | = | P fokussiert auf A's Augenregion |
| P ^b <=> A ^b | = | A fokussiert P's und P reziprok auf A's Augenregion |

Prof. Dr. Anja Stukenbrock
Universität Lausanne
Quartier UNIL-Chamberonne
Anthropole 4063
CH - 1015 Lausanne

anja.stukenbrock@unil.ch

Veröffentlicht am 31.7.2018

© Copyright by GESPRÄCHSFORSCHUNG. Alle Rechte vorbehalten.