

Frames und Diagramme im Dialog: Indexikalische und ikonische Zeichenmodi im verbal-gestischen Ko-Konstruieren von Reiserouten

Irene Mittelberg und Linn-Marlen Rekittke, RWTH Aachen University

Summary. This paper aims to show how semantic frames (Fillmore 1982) and diagrams (Peirce 1960) interact in the verbo-gestural co-construction of travelroutes in multimodal German dialogue data. Taking a cognitive-semiotic perspective (Mittelberg), we demonstrate how indexical and iconic signs jointly engender semiotic structure and meaning, in both an intramodal and intermodal (cross-modal) fashion. First, we focus on how linguistic and gestural means activate frames through instantiating frame elements, distinguishing between material or body-based frame elements and those that are less grounded in direct experience and pertain to more abstract knowledge structures. We further present distinct semiotic profiles of three phases underpinning the dynamic process of dialogic travel planning: a predominantly indexical, an indexical-iconic, and a diagrammatic-iconic phase. Overall, indexicality again proves to be a fundamental principle of multimodal interaction. Indexicality conditions and orchestrates not only the interlocutors' bodily anchorage in integrated physical and mental spaces (Fauconnier, Sweetser, Turner), but also their (inter-)subjective dialogue behavior and the multimodal fabric of emerging diagrams.

Keywords. mental spaces, semantic frames, diagrams, cross modal, gestures, meaning constitution, indexicality, iconic signs, multimodal interaction.

Zusammenfassung. Anhand multimodaler deutscher Dialogdaten zeigt dieser Beitrag, inwiefern semantische Frames (Fillmore 1982) und Diagramme (Peirce 1960) im verbal-gestischen Ko-Konstruieren von Reiserouten interagieren. Auf Basis eines kognitiv-semiotischen Theorierahmens (Mittelberg) wird herausgearbeitet, wie indexikalische und ikonische Zeichenmodi bedeutungs- und strukturstiftend ineinandergreifen: intramodal und intermodal (cross-modal). Zunächst gilt unser Interesse dem Instantiieren von materiellen, körperbasierten und im Gesprächskontext präsenten Frame-Elementen sowie dem Aktivieren komplexer, abstrakterer Wissensstrukturen. Anschließend präsentieren wir unterschiedliche semiotische Profile dreier Phasen der dialogischen Reiseplanung: eine prädominant indexikalische, eine indexikalisch-ikonische und eine

diagrammatisch-ikonische Phase. Dabei erweist sich Indexikalität auch hier als grundlegendes Prinzip multimodaler Interaktion. Indexikalität bedingt und orchestriert nicht nur die körperliche Verankerung der Sprechenden in integrierten physischen und mentalen Räumen (Fauconnier, Sweetser, Turner), sondern auch das (inter-)subjektive Dialogverhalten und cross-modale Entstehen von Diagrammen.

Schlüsselbegriffe. Mental Spaces, semantische Frames, Diagramme, cross-modal, Gesten, Bedeutungskonstitution, Indexikalität, ikonische Zeichen, multimodale Interaktion

1. Einleitung

Dieser Beitrag untersucht, wie DialogpartnerInnen mithilfe von verbalen und gestischen Zeichen alternative Reiserouten entwerfen und sich schließlich auf eine Strecke einigen. Sein Anliegen ist es aufzuzeigen, inwiefern die von den Gesprächsteilnehmenden diskutierten Entwürfe verschiedener Reisstrecken von multimodal evozierten semantischen Frames (vgl. Fillmore 1982) und Diagrammen (vgl. Peirce 1960) unterfüttert und vorangetrieben werden. Es soll gezeigt werden, inwiefern im Diskursverlauf gewisse Frame-Elemente und -Relationen, subjektiv wie intersubjektiv, eine kurzlebige semiotische Realität erlangen. Diese können im Gestenraum für einen Moment im geteilten Fokus des Dialoggeschehens stehen und so in die laufenden multimodalen Semioseprozesse sinn- und strukturstiftend einfließen.

Unser Interesse gilt einerseits insbesondere indexikalischen Ausdrucksformen, um ihre referentiellen und interaktiven Funktionen hinsichtlich der Reiseplanung und Verständigung zwischen den Planenden herauszuarbeiten. Schwerpunkte sind die indexikalischen Dimensionen interaktiver Praktiken des Sich-aufeinander-Beziehens und Sich-Einigens im Dialog (vgl. Bavelas u.a. 1995). Gleichzeitig geht es darum, wie indexikalische Elemente (z.B. Funktionswörter und deiktische Gesten) und ikonische Elemente (z.B. Inhaltswörter und gestisch evozierte Linien und Figurationen) ineinandergreifen und so nicht nur Frames aktivieren, sondern auch multimodale Diagramme im Dialograum entstehen lassen.

Mit Hinblick auf den *Reisen*-Frame¹ als übergreifende Framestruktur zeigt dieser Beitrag, wie die DialogpartnerInnen anvisierte Reiseziele gestisch in den Raum setzen, Richtungen andeuten und Reiserouten schematisch in die Luft zeichnen. Es kann angenommen werden, dass diesen ad hoc produzierten Diagrammen mehr oder minder ausdifferenzierte kognitive Karten (vgl. Downs und Stea 1982) der Gesprächsteilnehmenden zugrunde liegen. Gleichzeitig werden durch die multimodale Aktivierung der entsprechenden Frames enzyklopädische Wissensbestände von bestimmten Reiseformen, Orten und Ländern sowie persönliche Erinnerungen und Vorstellungen aufgerufen und Anknüpfungspunkte für weiterführende Überlegungen kreiert.

Die diesen Ausführungen zugrundeliegenden deutschen Konversationsdaten bestehen aus zwei mit Audio, Video und einem Bewegungserfassungssystem (Motion-Capture-System) aufgenommenen Dialogen, in denen jeweils zwei Gesprächsteilnehmende der ihnen gestellten Aufgabe nachgehen, gemeinsam eine fiktive Interrailtour durch Europa zu planen. Als ein erster Eindruck von unserem Untersuchungsgegenstand dient das folgende Transkript der verbalsprachlichen Äußerung einer multimodalen Sequenz. Der Routenvorschlag von Gesprächspartnerin GP_{R1} im Dialog 1 lässt sich anfangs noch ohne Berücksichtigung der simultan produzierten Gesten nachvollziehen ((1), Z01–07)², da Ortsnamen genannt werden, die man auf der eigenen kognitiven Karte von Europa verorten kann und von denen man eine gewisse Vorstellung aufgrund geographischen Wissens und eigener Reiseerfahrungen hat. Hinsichtlich der darauf formulierten ‚wieder nach Hause‘ führenden Route (Z08–14) ist dies nicht der Fall. Letztere besteht aus einer Aneinanderreihung von Funktionswörtern wie Präpositionen, lokalen und temporalen Adverbien:

(1) Verbale Routenplanung ohne Gesten-Annotation.

- | | | |
|----|--------------------|---|
| 01 | GP _{R1} : | Pass auf! |
| 02 | | wir fahren von Stockholm nach Prag |
| 03 | | und von Prag nach Kiew |
| 04 | | und von da nach unten irgendwo |
| 05 | GP _{L1} : | Ich weiß überhaupt nicht, wo Kiew liegt relativ zu Prag |
| 06 | GP _{R1} : | ich auch nicht %lacht |
| 07 | GP _{L1} : | aber gut vorgetäuscht |
| 08 | GP _{R1} : | ich denke |
| 09 | | nee ich denke wir fahren dann so weißte von |
| 10 | | da |
| 11 | | nach da |
| 12 | | rüber |
| 13 | | runter |
| 14 | | und dann nach da und dann wieder nach Hause |

Wie im Verlauf des Artikels detailliert dargelegt, ergeben sich lokal erzeugte Form-Bedeutung-Korrelationen gerade erst im Verbund mit den synchron produzierten Gesten, insbesondere Gestendiagramme, sowie dem im Gespräch aufgebauten multimodalen Diskurskontext.

Durch das Visualisieren solcher virtueller Referenzstrukturen mittels Motion-Capture Technologie wird illustriert, wie sie sich im Gestenraum schrittweise entfalten und dabei den Gesprächspartnern als kognitiv-semiotische Ressourcen dienen können (vgl. Schüller und Mittelberg 2016, 2017). Solche kollaborativen Diskursprozesse des Konstruierens und Aushandelns von Reiserouten werden in diesem Beitrag anhand deutscher multimodaler Dialogdaten untersucht und durch das Zusammenführen der im Folgenden aufgeführten Ansätze theoretisch untermauert.

2. Kognitiv-semiotischer Theorierahmen

2.1 *Peirce, Jakobson und kognitive Linguistik*

Von einem verkörperten und multimodalen Verständnis von Sprache, Kognition und Interaktion ausgehend, nähert sich dieser Beitrag crossmodalen Strategien der Bedeutungskonstitution mit einem kognitiv-semiotischen Ansatz (vgl. Mittelberg 2006, 2008, 2019a/b), der kognitiv-linguistische Ansätze (vgl. u.a. Dancygier und Sweetser 2014; Fillmore 1982; Gibbs 2006; Johnson 1987) und klassische semiotische Theorien von Charles S. Peirce (vgl. 1955, 1960) und Roman Jakobson (vgl. 1956, 1966, 1971) verknüpft. Diese Perspektiven sind insofern kompatibel, als sie bei der Analyse von bedeutungsstiftenden Prozessen besonders die Motivation, Funktion und Struktur von Sprachformen und Sprachgebrauch berücksichtigen. Beide Traditionen räumen verinnerlichte („embodied“; vgl. Gibbs 2006) wahrnehmungs- und handlungsbasierten Mustern von Erfahrung und Interpretation einen hohen Stellenwert ein (vgl. Mittelberg 2013). Eine weitere geteilte Prämisse ist, dass Bedeutung einem Zeichenträger, etwa einem Wort oder einer Geste, nicht innewohnt, sondern im Moment der Interpretation dynamisch im Bewusstsein entsteht und bestimmt durch pragmatische Kräfte, kognitiv-semiotische Prinzipien und Erfahrung geleitete Vorstellungen und Assoziationen hervorruft. Dies wird beispielsweise durch das Rekrutieren von mentalen Strukturen wie Frames (vgl. Coulson 2001; Ziem 2014) und interagierenden metonymischen und metaphorischen Prozessen (vgl. Dancygier und Sweetser 2014) geleitet.

Indem diese Ansätze in ihrer Zusammenschau einen weiteren semiotischen Geltungshorizont aufspannen als rein sprachbasierte Theorierahmen, sind sie im besonderen Maße geeignet, auf Gesten und andere Kommunikationsformen angewandt zu werden, welche Modalitäten mit einbeziehen, denen im Vergleich zur Laut- und Schriftsprache kein hochgradig konventionalisierter symbolischer Kode zugrunde liegt. Gesten sind zwar zu unterschiedlichen Graden konventionalisiert, sei es bedingt durch kulturelle Praktiken, Bewegungsmuster oder einverlebte konzeptuelle Strukturen und Operationen.³ Da in redebegleitenden Gesten jedoch das Verhältnis von semiotischen Modi, wie Ikonizität, Indexikalität und Konventionalität (oder Gewohnheit; vgl. Peirce 1960), eine größere Varianz aufweist und in jeder multimodalen Performanz unter Einfluss von pragmatischen Faktoren sich stets neu austariert, wird hier den oben genannten semiotischen Theorien ein zentraler Stellenwert eingeräumt. Das Ziel besteht dabei, und auch im Folgenden, darin, der spezifischen Medialität von Gesten und ihrer Verflechtung mit sprachlichen Äußerungen soweit wie möglich gerecht zu werden.

Unter Gesten werden hier Handkonfigurationen und Handbewegungen sowie Körperposen und -bewegungen verstanden, welche die visuell-aktionalen Komponenten einer multimodalen Äußerung ausmachen und verschiedene kommunikative Funktionen aufweisen (vgl. Kendon 2004;

Müller 1998). In Anlehnung an Peirce (1960) und Jakobson (1960) betonen wir die hierarchisierte Schichtung von unterschiedlichen, interagierenden Zeichenmodi und multiplen Funktionen. Danach bestimmt die prädominante Funktion die Bedeutung eines Zeichens letztendlich, wobei andere Funktionen gleichzeitig in unterschiedlichem Grade mitwirken (vgl. Mittelberg 2013 und Mittelberg und Waugh 2009; Müller 1998). Hier geht es insbesondere um das cross-modale Ineinandergreifen von ikonischen und indexikalischen Zeichenmodi.

2.2 Semantische Frames

Gemäß des Themenschwerpunktes dieses Sonderheftes bildet die Theorie der ‚Mental Spaces‘ (vgl. Fauconnier 1985/1994, 1997) und insbesondere solche Ansätze, die eine framebasierte Struktur von mentalen Räumen annehmen (vgl. Fauconnier 1985/1994; Coulson 2001; Dancygier und Sweetser 2014), den globalen konzeptuellen Rahmen für unsere Untersuchungen.⁴ Unsere Analyse stützt sich jedoch in erster Linie auf den Begriff der semantischen Frames aus Fillmores (1977, 1982, 1985) grundlegenden Schriften. Neuere FrameNet-Arbeiten⁵ werden hier dahingehend berücksichtigt, dass wir zeigen, wie in unseren Dialogdaten gestische und sprachliche Zeichen Frame-Elemente instantiieren und so einen Frame – oder mehrere Frames gleichzeitig – aktivieren. Syntaktische Realisationen und multimodale Konstruktionen können nicht thematisiert werden.⁶

Gemäß den hier zugrunde gelegten Arbeiten von Fillmore sind Frames semantische Systeme, welche die Bedeutung sprachlicher Ausdrücke bestimmen und sowohl Sprachproduktion als auch Diskursverstehen unter Einbeziehung enzyklopädischer Wissensbestände und verinnerlichter Erfahrungskontexte (vgl. ‚scenes‘, Fillmore 1977) leiten. Frames bestehen aus miteinander verbundenen Teilen, wobei ein Element ein korreliertes Element sowie das gesamte System von Konzepten aktivieren kann (vgl. Fillmore 1982: 111).

A ‘frame’, as the notion plays a role in the description of linguistic meanings, is a system of categories structured in accordance with some motivating context. [...] The motivating context is some body of understandings, some pattern of practices, or some history of social institutions, against which we find intelligible the creation of a particular category in the history of the language community (Fillmore 1982: 119).

Für unsere Reiseplanungs-Studie ist der *Reisen*-Frame (engl. *Travel*-Frame) besonders relevant; er vereint in sich Handlungsmuster, Wissensstrukturen und motivierende Kontexte unterschiedlich starker Grade von Abstraktion und Komplexität. Im FrameNet Projekt des Deutschen wird er wie folgt definiert:

In diesem Frame begibt sich ein REISENDER auf eine Reise, eine Aktivität, die im Allgemeinen im Voraus geplant ist, bei der sich der REISENDE von einem Startpunkt (der QUELLE) zu einem ZIEL entlang eines WEGES oder innerhalb eines BEREICHS bewegt. Die Reise kann von MITBETEILIGTEN und GEPÄCK begleitet sein. Die DAUER oder die ENTFERNUNG der Reise, beide normalerweise lang, können auch als FORTBEWEGUNGSART bezeichnet werden. Wörter in diesem Frame betonen den gesamten Prozess, von einem Ort zum anderen zu gelangen, anstatt nur den Anfang oder das Ende der Reise zu beschreiben (FrameNet des Deutschen: <https://gsw.phil.hhu.de/framenet/frame?id=974> [31.5.2021]).

Der *Reisen*-Frame hat folgende Kern-Frame-Elemente (in alphabetischer Reihenfolge): BEREICH, FORTBEWEGUNGSART, QUELLE, REISENDER, RICHTUNG, WEG, ZIEL. Zu den Nicht-Kern-Frame-Elementen gehören ART_UND_WEISE, DAUER, ENTFERNUNG, GESCHWINDIGKEIT, MITBETEILIGTE, ZEIT und ZWECK; der gängigen Konvention folgend sind sie hier in Kapitälchen gesetzt. Im Fall der vorgegebenen Dialogsituation sind die REISENDEN die zwei GesprächspartnerInnen. Der vorgegebene BEREICH ist Europa, welcher in Unterbereiche (z.B. Länder, Regionen, Landschaften) gegliedert ist und auf Teilstrecken zwischen vier zu wählenden Etappenzielen zu bereisen ist. Meist bildet dabei der jeweilige Wohnort den Ausgangs- und Endpunkt (QUELLE und ZIEL). Es handelt sich um Zugreisen (InterRail). Die ZEIT des Reisens beträgt zwei Wochen. Der ZWECK ergibt sich aus den Präferenzen der Teilnehmenden.

Wir gehen davon aus, dass während der auf die Aufgabe fokussierten Dialoge der *Reisen*-Frame im Hintergrund aktiviert bleibt, da die Aufgabe darin besteht, eine Reiseroute auszuhandeln. Dabei interagiert er, wie in Abschnitt 4 und 5 zu zeigen ist, mit anderen Frames, je nach Phase des Planungsprozesses und den Vorschlägen und Erzählungen der SprecherInnen.

2.3 Diagrammatische Ikonizität in der multimodalen Reiseplanung

Diagramme gelten allgemein als ikonische Zeichenprozesse, „icons of relations“ (vgl. Peirce 1955, 1960), die Beziehungen zwischen Elementen darstellen. In Peirces (1960: 157, 2.277) eigenen Worten: Ikone, „which represent the relations, mainly dyadic, (...) of the parts of one thing by analogous relations in their own parts, are diagrams“ (vgl. Sternfeld 2007). Straßen- und U-Bahnkarten sowie Familienstammbäume sind bekannte Beispiele.

In unserem Korpus machen GesprächspartnerInnen räumliche Relationen und Reisebewegungen in Form von, mal mehr und mal weniger, flüchtigen gestischen Diagrammen sichtbar. In die Luft gezeichnete virtuelle Linien, Konfigurationen von in den Raum gesetzten Punkten oder andere Varianten von sich allmählich aufbauenden Gestenkomplexen vermögen aufzuzeigen, wie sich für die jeweiligen SprecherInnen einzelne Orte,

Länder oder Aktivitäten räumlich und/oder zeitlich zueinander verhalten. Dabei geht es nicht um eine genaue (z.B. metrische) Abbildung der räumlichen Relationen und Distanzen; letztere werden im Gestenraum auf die dort möglichen Größenverhältnisse ad hoc angepasst und oft flexibel verkürzt dargestellt. Hinsichtlich der Gestendiagramme ist hier entscheidend, dass schematische Figurationen dieser Art oft durch den Einsatz unterschiedlich gestalteter indexikalischer Handformen und richtungsangeben-der Bewegungen emergieren. SprecherInnen setzen zum Beispiel mit Zeigegesten Punkte in den Gestenraum, um Personen, Dinge, Orte, Ideen oder Diskursinhalte, über die sie sprechen, zu vergegenwärtigen und diese zueinander in Beziehung zu setzen. Die Verbindungen zwischen den Punkten können gedacht oder durch gestische Bewegungen von Punkt zu Punkt im Gesprächsverlauf hergestellt werden (vgl. Enfield 2009; Fricke 2007, 2012; Mittelberg 2006, 2014). Dabei ist von zentraler Bedeutung, dass Diagramme nicht nur bereits bestehende Relationen abbilden, sondern auch neue, im Denkprozess entstehende Verbindungen schlagen können. Sie haben somit ein epistemisches Potential, das auch in unserem Kontext der Reiseentwürfe zum Tragen kommt (vgl. Pombo und Gerner 2010).

3. Indexikalität als Grundprinzip multimodaler Interaktion

Indexikalische Zeichenmodi erwiesen sich in unseren multimodalen Dialogdaten als prädominant, da sie bei dem Evozieren von Frames und Diagrammen verschiedene Funktionen erfüllen. Um diese erklären zu können, sei hier ein kurzer Einblick in die den Gesten inhärente Indexikalität gegeben.

Gesten sind genuin an den sie hervorbringenden kommunizierenden Körper gebunden und können nicht losgelöst von diesem als Zeichen fungieren (Haviland 2000; Mittelberg 2010a/b, 2013, 2019a/b; Sweetser 2012). Nach Peirce (vgl. 1960: 143) gründen indexikalische Zeichen bekanntlich in Kontiguitätsbeziehungen wie einer faktischen oder kausalen Verbindung, Angrenzung oder Nähe. Sie sind von ihrem Referenzobjekt konditionierte Zeichen. Zeigegesten, die visuell wahrnehmbar auf etwas anderes verweisen als sich selbst, sind indexikalische Zeichen par excellence. Das Referenzobjekt konditioniert sie in dem Sinne, dass die Sprechenden Arm und Hand (und in gewisser Weise den gesamten Körper) als Zeichenträger auf das ausrichten, auf das sie zeigen, und so eine Kontiguitätsbeziehung herstellen.

In der Literatur werden verschiedene Praktiken der sprachlich-gestischen Bezugnahme unterschieden: v.a. Zeigen auf gegenständliche oder abstrakte, auf im situativen Kontext existierende oder gedachte Referenzobjekte, sowie richtungsangebendes Zeigen (vgl. Fricke 2007; Hassemer 2015; Kita 2003; McNeill 1992; Wilkens 2003).⁷ Abgeschwächte Grade von Indexikalität wurden als Grundlage metonymischer Modi gestischer Zeichenkonstitution beschrieben, z.B. zwischen gestikulierenden Händen und

den imaginären Objekten, die sie zu halten scheinen (vgl. Mittelberg 2019b; Mittelberg und Waugh 2014). In diesem Beitrag fragen wir nach den indexikalischen Dimensionen interaktiver und diagrammatischer Gesten.

3.1 Indexikalische Verankerung materieller und mentaler Diskurselemente

Der multimodal kommunizierende Körper wird im vorliegenden Beitrag als indexikalischer Anker verstanden: Von ihm aus organisieren sich die raumzeitlichen Koordinaten multimodaler Diskursentfaltung im Sinne von Bühlers Konzept der ‚Origo‘ (vgl. Bühler 1934/1982; Fricke 2007; Müller 1998) stets von Neuem. Dabei entstehen materielle, imaginierte und interaktive Gesprächskontexte. SprecherInnen können sich zudem in andere Personen und Kontexte hineinversetzen und Szenarien aus verschiedenen Perspektiven (‚viewpoint‘; vgl. McNeill 1992) beschreiben. Ikonische Gesten sind, wie bildliche Vorstellungen und Beschreibungen, tendenziell indexikalisch konditioniert (vgl. Mittelberg 2010b sowie Mittelberg und Waugh 2014 hinsichtlich Ikonizität-Indexikalitäts-Kontinuen für Gesten).

In der multimodalen Interaktion können gegenständliche Strukturen und ihre Affordanzen (vgl. Gibson 1977) als ‚material anchors‘ (vgl. Hutchins 1995) fungieren. Als solche lassen sie sich in kognitive Operationen sowie diskursive Erklärungs- und Verstehensprozesse (‚conceptual integration‘ oder ‚blending‘; vgl. Fauconnier und Turner 2002) einbeziehen. Dazu gehört, dass SprecherInnen durch Gesten den Kontakt (d.h. Kontiguität) mit der materiellen Kultur und Umwelt aktiv herstellen, indem sie mit Oberflächen, Gegenständen und Instrumenten interagieren (siehe ‚environmentally-coppled gestures‘ nach Goodwin 2007; Streeck u.a. 2011). Sweetser sieht im menschlichen Körper selbst einen funktionsstarken und stets perspektivierten ‚material anchor‘:

Gesture, like language, shows deictic centers, displacement phenomena and blended deictic structure. Our bodies are the most flexible and powerful material anchors [...] for representing and expressing viewpoint. [...] there is no more powerful icon for a bodily viewpoint than an actual body with an actual inherent viewpoint. Gestures therefore express spatial indexicality via embodied spatial indexicality – for humans, a pointing gesture directs joint attention to a particular actual location relative to the actual pointing body. This in itself is a cognitive achievement [...] (Sweetser 2012: 13).

Zwecks Orientierung und Streckenplanung wird typischerweise mit Land- und Straßenkarten interagiert: durch das Zeigen auf bestimmte Orte und eingezeichnete Wege (vgl. Fricke 2007).

3.2 *‘Imaginary material anchors‘ für geteilte Intentionalität im Gestenraum*

In unseren Daten besteht die kognitiv-semiotische Leistung im kollaborativen Erstellen von fiktiven Reiserouten. Da die GesprächspartnerInnen die ihnen vorgegebenen kognitiven und kommunikativen Aufgaben gemeinsam bewältigen müssen, benötigen wir für die Analyse Konzepte, die dieser semiotischen Komplexität gerecht werden können. Es geht primär um die gemeinsame Ideenentwicklung in gedanklichen Welten, von denen Ausschnitte und Dimensionen gerade durch Gesten und multimodal erwirkte Diagramme semiotisch habhaft gemacht und im Gestenraum zum Fokus der geteilten Aufmerksamkeit und Intentionalität werden (*‘joint attention‘*, vgl. Tomasello 1995; *‘joint intentionality‘*, vgl. Tomasello und Carpenter 2007).

Hier können wir uns auf Hutchins’ (vgl. 2005: 1575) Konzept der *‘imaginary material anchors‘* stützen, d.h. auf solche Anhaltspunkte und Elemente, die nicht als gegenständliche Strukturen im Kontext vorhanden sind, sondern die man sich unter Rückgriff auf Wissen und verinnerlichte Erfahrung (mehr oder weniger genau) vorstellen kann. In unserem Falle geht es um imaginäre Kartenstrukturen und das Manipulieren besteht beispielsweise im Zeigen auf gedachte Orte sowie im virtuellen Skizzieren von Reisstrecken.

Die kognitive Semantik und Semiotik geben uns also Instrumente an die Hand, die es ermöglichen, Vorstellungswelten in die Analyse mit einzu beziehen und so aufzuzeigen, inwiefern DialogpartnerInnen imaginäre und irreale Welten, sprich verschiedene *‘mental spaces‘*, im Gesprächsverlauf erschaffen und sich zwischen dem Real Space (vgl. Liddell 2003), dem von den GesprächsteilnehmerInnen geteilten Raum der Gesprächssituation, und dem jeweiligen sprachlich-gestisch eröffneten und fortentwickelten Story Space hin und her bewegen (vgl. Sweetser und Stec 2016: 239). Ihre Körper und Gesten bilden dabei als indexikalische Anker ein sichtbares Bindeglied.

3.3 *Dialogische Indexikalität: Interaktive Gesten*

In unseren Dialogdaten interagieren jeweils zwei indexikalisch verankerte und räumlich unterschiedlich ausgerichtete, nämlich sich zugewandte, Personen. In diesen konversationellen Interaktionsformen spielen indexikalische Elemente und Praktiken wie Funktionswörter (z.B. Personal- und Demonstrativpronomen) und gestisches Aufeinander-Beziehen eine zentrale Rolle. Gemäß Bavelas (vgl. 1994: 213) gibt es vier Typen von interaktiven Gesten: Die *‘Delivery Gesture‘* bezieht sich auf von GesprächspartnerInnen ausgetauschte Informationen: Sie wird unterteilt in eine *‘General Delivery Gesture‘*, die neue Informationen sozusagen an den/die GesprächspartnerIn weiterreicht. Eine *‘Shared Information Gesture‘* bezieht sich auf Inhalte, die den GesprächspartnerInnen bereits bekannt sind (*‘common*

ground‘; vgl. Clark 1996). ‚Citing Gestures‘ sind anaphorische Verweise auf eine im Gespräch vorausgegangene Äußerung, wobei zwischen einer ein Zitat kennzeichnende ‚General Citing Gesture‘ und einer ‚Acknowledgement Gesture‘ unterschieden wird. Durch letztere signalisieren Zuhörende, dass sie die Äußerung ihrer GesprächspartnerInnen vernommen bzw. verstanden haben. Die ‚Seeking Gesture‘ erbittet eine Antwort, Hilfe oder Zustimmung. Schließlich regeln ‚Turn Gestures‘ den Sprecherwechsel: Mit einer ‚Giving Turn Gesture‘ gibt die/der Sprechende die Sprecherrolle an eine/einen GesprächspartnerIn, und mit einer ‚Taking Turn Gesture‘ wird die Sprecherrolle übernommen (vgl. Bavelas u.a. 1995).

Die Analysen in den folgenden Abschnitten bringen die diesen interaktiven Gesten inhärente Indexikalität zutage: Sie zeigen auf im Dialog verbal und/oder gestisch instantiierte Frame-Elemente und andere Äußerungskomponenten, welche sozusagen virtuell ‚im Raum stehen‘.

4. Tendenzen im verbal-gestischen Evozieren von Frames

Für unser Gestenverständnis ist hier grundlegend, dass Frames erfahrungsbedingt sind, d.h. „motivated by human experiences, social institutions, and cultural practices“ (Coulson 2001: 18). Außersprachliche, kontextuelle und ko-textuelle Faktoren zählen zu den Kräften, die unser Verständnis von Begriffen und ihre Vernetzung konditionieren (vgl. Fillmore 1982; vgl. Ziem 2014: 24, 188ff.). Wie wir zeigen möchten, motivieren solche Erfahrungsmuster und kulturellen Praktiken, und damit insbesondere ‚Scenes‘ in Fillmores (1977) Sinn, nicht nur das Entstehen von Frames, sondern auch Prozesse gestischer Zeichenkonstitution und -interpretation. Aufbauend auf Mittelbergs (2017, 2019b) Arbeiten zu „Embodied Frames und Scenes“ wird nun herausgearbeitet, inwiefern sich diese in unseren Reiseplanungsdaten manifestieren. Nach einer kurzen Beschreibung der Datengrundlage (Abschnitt 4.1) liegt der Fokus zunächst auf dem multimodalen Aktivieren semantischer Framestrukturen und damit vor allem auf inhaltlichen (Frame-)Elementen in der Reiseplanung (Abschnitt 4.2). Anschließend widmen wir uns der Verbindung von Frame-Elementen in Diagrammen (Abschnitt 4.2).

4.1 Studiendesign, Korpus, Annotation

Die der hier vorgestellten Studie zugrundeliegenden Diskursdaten stammen aus dem *MuSKA Korpus* (*Natural Media Lab, RWTH Aachen*; Brenner und Mittelberg 2015). Insgesamt wurden sechs Stunden Audio-, Video- und Motion-Capture Daten aufgenommen. Das Korpus umfasst zwei Gesprächsgenres (Dialog und Narration) in zwei Sprachen (Deutsch und Amerikanisches Englisch). Für die vorliegende Studie wurden die deutschen Dialogdaten von sechzehn Personen (acht Paaren) berücksichtigt.

Die StudienteilnehmerInnen waren zwischen 20 und 35 Jahre alt. Hier präsentieren wir Ergebnisse qualitativer Analysen ausgewählter Sequenzen aus zwei der Dialoge von sich zuvor bekannten GesprächspartnerInnen (jeweils RechtshänderInnen).

Der Versuchsaufbau war wie folgt: Jeweils zwei GesprächspartnerInnen sitzen sich im Labor auf erhöhten Stühlen gegenüber. Ihre kommunikativen Bewegungen wurden mit 14 Infrarotkameras sowie einer HD-Videokamera erfasst. Dafür wurden bei jeder Person 35 reflektierende Marker an den Händen, Armen, Schultern, Knien, Hüfte und Nacken befestigt. Vor dem Gespräch wurde ihnen erklärt, dass sie ein fiktives Interrail-Ticket gewonnen haben und sie innerhalb von zwei Wochen insgesamt fünf unterschiedliche Orte innerhalb Europas besuchen können. Gemeinsam sollen sie planen, wo sie hinfahren, in welcher Reihenfolge sie die Ziele ansteuern und was sie dort unternehmen wollen (z.B. Stadtbesichtigung, Kultur, Entspannung oder Abenteuer).

Die verbalen Äußerungssequenzen wurden anhand der Diskurs-Transkriptionskonvention nach Du Bois u.a. (1993) und Duncan (vgl. McNeill 2005: 275ff.) in Intonationseinheiten unterteilt und transkribiert. Die Gesten wurden mit Hilfe der Annotationssoftware Elan mit Fokus auf die folgenden drei Aspekte annotiert: Gestenform und Bewegungsgestalt (gemäß Bresssem 2013), die exakte Sprach-Gesten-Synchronie sowie semantisch-pragmatische Beziehung zur verbalen Äußerung. Jede Gesteneinheit wurde in Phasen segmentiert und im Transkript durch eckige Klammern um die parallel hierzu produzierten verbalen Äußerungselemente markiert. Jede Geste ist mit einem hochgestellten G sowie ihrer jeweiligen Nummer gekennzeichnet ([...]G¹). Die fett markierten Elemente in den Sprachtranskripten repräsentieren Segmente, auf die der bedeutungstragende ‚stroke‘ fällt; zwei vertikale Linien orten einen ‚stroke hold‘ (vgl. Kita u.a. 1998). Dabei können sowohl die Bewegungstrajektorie, die Handform, die Position zum Körper und im Gestenraum sowie die Orientierung der Handflächen ausschlaggebend sein (vgl. Bresssem 2013; McNeill 1992).

4.2 Verbal-gestisches Instantiieren körperlich erfahrbarer und im Kontext präsenter Frame-Elemente

Das Fundament für unseren Ansatz bilden lebensweltliche Erfahrungskontexte einschließlich die für sie habituellen Gegenstände, räumlichen Verhältnisse und Handlungen. Es handelt sich hier um körperliche Bewegungsmuster wie z.B. Gehen, Radfahren, oder jemanden umarmen, sowie um Gegenstände als solche und das routinierte Handhaben von Objekten und Instrumenten. Unser (Körper-)Wissen wurzelt in den entsprechenden einverleibten (‚embodied‘) Erfahrungsmustern, die in der eigenen Körperlichkeit und dem direkten Interagieren mit der materiellen Kultur und sozialen Umwelt tief verankert sind (vgl. Mittelberg 2017, 2019b; Bresssem und Müller 2017; Müller 2017).

Um zu sehen, wie sich solche basalen Erfahrungskontexte im multimodalen Evozieren von hier maßgeblichen Frames manifestieren, geht es im ersten Beispiel um den Begriff *Wetter*, der sich, obwohl wir Wetterverhältnisse körperlich erfahren, gestisch nicht gut unmittelbar darstellen lässt. Der *Wetter*-Frame ist wie folgt definiert:

Die Umgebungsbedingungen von Temperatur, Niederschlag, Wind und Sonne beziehen sich auf einen bestimmten ORT und eine bestimmte ZEIT. Eine weitere NÄHERE_BESTIMMUNG der betreffenden Bedingungen kann ebenfalls angegeben werden (FrameNet des Deutschen: <https://gsw.phil.hhu.de/framenet/frame?id=932> (31.5.2021)).

Kern-Frame-Elemente sind ORT und ZEIT. So interessiert uns vor allem, wie der *Wetter*-Frame in der folgenden Sequenz multimodal (teil-)evoziert und durch Verbindungen zu anderen Frames unterschiedlich spezifiziert wird.

In der folgenden Sequenz geht es um eine zurückliegende Reiseerfahrung von Sprecherin GP_{R2} in Budapest. Budapest initiiert hier das Kernelement ZIEL im *Reisen*-Frame. Da sich GP_{R2} nicht an das genaue Datum (Kernelement ZEIT im *Reisen*-Frame) erinnern kann, hilft ihr GP_{L2} auf die Sprünge, indem er fragt, ob Schnee lag ((2), Z01). Die lexikalische Einheit ‚Schnee‘ aktiviert hier den *Niederschlag*-Frame⁸, und zwar so, dass Winterwetter suggeriert wird und damit gleichzeitig der *Wetter*-Frame nicht nur aktiviert, sondern auch spezifiziert wird. Dabei produziert GP_{L2} eine beidhändige ‚Palm-Up Open Hand‘ Geste (PUOH, vgl. Müller 2004) mit pragmatischer Funktion (Abb. 1a). Er erbittet, sozusagen mit leeren Händen, eine Antwort, was einer interaktiven ‚Seeking Gesture‘ gleichkommt (vgl. Bavelas u.a. 1995, s. Abschnitt 3.3).

(2) Wetterverhältnisse in Budapest.

- 01 GP_{L2}: **[Lag Schnee?]**^{G1}
- 02 GP_{R2}: nein (-) es [war **schon warm**]^{G2}
- 03 also so **[T-Shirt-Wetter]**^{G3} und
- 04 **[Shorts Wetter]**^{G4} war richtig schön und so
- 05 **[Schwimm-Wetter]**^{G5}
- 06 GP_{L2}: %lacht **[okay]**^{G6}

GP_{R2} erinnert sich daraufhin an die sommerlichen Wetterverhältnisse. Zunächst eröffnet sie den *Umgebungstemperatur*-Frame⁹ verbal durch „war schon warm“ (Z02). Hiermit wird gleichzeitig der *Wetter*-Frame erneut teinstantiiert: das Frame-Element SPEC. ist nun mit einer Instanz („warm“) spezifiziert. GP_{R2} macht dabei ebenfalls eine PUOH-Geste (Abb. 1b) mit vorwiegend modalen Funktion (vgl. Kendon 2004; Müller 2004). Sie unterstreicht damit, dass es für sie zweifelsohne „warm“ war (und eben kein winterliches Wetter herrschte).



1a: „Seeking Gesture“ (G1) (Z 01).



1b: „General Delivery Gesture“ (G2) (Z02).



1c: Gestischer Bezug auf T-Shirt (G3).



1d: Illustration der Länge einer Shorts (G4).



1e: Ikonisch-aktionale Schwimm-Geste (G5).



1f: Nachahmung der Schwimm-Geste (G6).

Abb. 1a–f: Wetter-Frame Sequenz.

Dann führt GP_{R2} mehrere Indizien für warmes Wetter auf, indem sie Elemente aus dem *Kleidung*-Frame integriert:¹⁰ „T-Shirt-Wetter“ begleitet sie durch beidhändige zirkulierende Handbewegungen dicht an ihrem mit einem T-Shirt bekleideten Oberkörper (Abb. 1c), welche gleichzeitig Wärmeempfinden suggerieren kann. Es folgt ein gestischer Verweis auf ihre Oberarme bzw. die kurzen Ärmel ihres T-Shirts. Zeitgleich mit „Shorts-Wetter“ markiert sie mit den Handkanten auf ihren Oberschenkel, wo Hosenbeine von Shorts enden (Abb. 1d). Dann ahmt sie auf „Schwimm-Wetter“ (Z05) die für das Brustschwimmen typische Arm- und Handhaltung sowie Bewegungen schematisch nach (Abb. 1e). Bemerkenswert ist hier, wie GP_{L2} die Wetterbeschreibung seiner Reisepartnerin multimodal rahmt: eingangs durch seine Frage und PUOH ‚seeking‘ Geste, die sofort von GP_{R2} gespiegelt wird (allerdings als ‚General Delivery Gesture‘; vgl. Bavelas 1994), und abschließend durch eine sein Gegenüber nachahmende Schwimm-Geste (Abb. 2f). Diese Mimikry-Gesten vermitteln ein starkes Alignment der GesprächspartnerInnen (vgl. Holler und Wilkin 2011).

In dieser Sequenz fällt außerdem auf, dass nur die beiden Schwimm-Gesten (1e und 1f) ikonischer Natur sind: Sie sind von einem leicht erkennbaren Bewegungsmuster motiviert, welches metonymisch ohne Hinzunehmen der Beine exemplifiziert wird (vgl. ‚action image icon‘ gemäß Mittelberg und Waugh 2014). Die zweite Schwimmgeste ist noch reduzierter als die ursprüngliche. Indexikalität ist in den beiden Gesten prädominant, mit denen GP_{R2} T-Shirt und Shorts am eigenen Körper verortet und damit dort, wo man sie typischerweise trägt (vgl. ‚body part index‘ gemäß Mittelberg und Waugh 2014). Hier initiieren die verbal geäußerten lexikalischen Einheiten (‚T-Shirt‘) und (‚Shorts‘) mit den indexikalischen Gesten auf cross-modale Weise Kern-Frame-Elemente des *Kleidung*-Frames. Anhand der hier mehrmals modifizierten lexikalischen Einheit ‚Wetter‘ manifestiert sich die Tatsache, dass der spezifizierte *Wetter*-Frame diese Sequenz multimodaler Einzelinitiiierungen von Frame-Elementen, und damit den ‚Turn‘ von GP_{R2}, strukturiert. Er ist auf einer Ebene stärkerer Abstraktion angesiedelt als die illustrierten Teilaspekte wie Schwimmen oder Shorts.

Dass, oder wie, die Elemente in den jeweiligen Frames zusammenhängen bzw. wie die Frames ineinandergreifen, wurzelt hier einerseits in erfahrungsbedingten Korrelationen von Elementen (z.B. wenn es heiß ist, trägt man T-Shirts, keine Pullover). Gleichzeitig leiten größere frame-strukturierte Szenarien und Wissenskomplexe unser Verständnis von Erzählungen und kreieren Assoziationen und Erwartungen (vgl. Skripts nach Schank und Abelson 1977). Daraus lassen sich auch nicht direkt instantiierte Frame-Elemente inferieren, hier z.B. dass die Budapest-Reise im Sommer stattgefunden hat. So ist dann indirekt auch das Kernelement ZEIT des im Hintergrund aktiven *Reisen*-Frame spezifiziert.

Andererseits können wir auch aus der dynamisch emergierenden semiotischen Kontextur Schlüsse ziehen, nämlich der sequenziellen Abfolge (oder temporalen Kontiguität) und simultanen Kombination (vgl. Jakobson 1956) von sprachlichen und gestischen Zeichen. Semiotische Erfahrungen wie diese haben ebenfalls einen Einfluss auf sich dynamisch entwickelnde Wissensstrukturen, insbesondere auf der Seite der Interpretierenden.

Diese Beispielsequenz veranschaulicht zudem, wie im Kontext vorhandene, gegenständliche Ressourcen spontan in die multimodale Beschreibung eingewoben werden. Somit fungieren Körper und Kleidungsstücke von Sprecherin GP_{R1} als sichtbare (bildliche) ‚material anchors‘ (vgl. Hutchins 1995; Sweetser 2012). Der Körper fungiert hier als Anker, der den situativ-materiellen Real Space (vgl. Liddell 2003) und aufgespannten Story Space miteinander verbindet (vgl. Sweetser und Stec 2016: 239). Wir wenden uns nun weniger gegenständlichen Ankerstrukturen (‚imaginary material anchors‘ gemäß Hutchins 2005) und entsprechenden anders gearteten semiotischen Praktiken zu, die im dialogischen Aushandeln von Reiserouten eine wichtige Rolle spielen.

4.3 Abstrakte Wissensstrukturen und verbal-gestische Diagramme

Prozesse multimodaler Bedeutungskonstitution rekrutieren auch vergleichsweise abstrakte und komplexere semantische Strukturen (vgl. Mittelberg 2017). Dazu gehören zum Beispiel solche, von denen angenommen wird, dass sie das konzeptuelle System, abstrakte Wissensbereiche oder ganze Diskurse strukturieren (vgl. Coulson 2001; Ziem 2008, 2014). Unter anderem funktionieren sie als Ordnungsprinzipien von miteinander interagierenden verkörperten kognitiven Strukturen und Prozessen wie Metaphern, Metonymien, Bildschemata, Kräftedynamiken und mentalen Räumen (vgl. Dancygier und Sweetser 2014). Dabei ist ein zentraler Aspekt, dass sich solche größeren schematischen Architekturen in der gestischen Modalität manifestieren können, jedoch nicht müssen. In diesem Abschnitt zeigen wir, wie sich SprecherInnen anhand von bimodal erzeugten Diagrammen in solchen Wissensstrukturen, die nicht in direkter körperlicher Erfahrung und im materiellen Kontext verhaftet sind, bewegen und diese intersubjektiv teilbar machen.

Dabei ist maßgeblich, wie Kern-Elemente des *Reisen*-Frames (eingeführt in Abschnitt 2.2) im Dialog initiiert werden. Dieser Frame besteht einerseits aus einzelnen erfahrungsbedingten Elementen, weist aber eine relative Komplexität und Vieldimensionalität auf: Regionen, Städte, Länder, Fortbewegung, Richtungen, Transportmittel, Essen, Übernachtungsmöglichkeiten, kulturelle und sportliche Aktivitäten usw. Dabei kommen allgemeine enzyklopädische und geographische Wissensbereiche sowie individuelle und gruppenspezifische Gewohnheiten und Präferenzen zum Tragen (z.B. geprägt durch Familien- und Gruppenurlaube).

Es verwundert nicht, dass Frame-Elemente wie die Städte STOCKHOLM und PRAG (siehe Einleitung, Beispiel 1 und Beispiel 3) nicht ikonisch oder durch Symbole repräsentiert werden. Vielmehr werden sie durch in den Raum gesetzte, indexikalische Gesten und die synchron geäußerten Städtenamen instantiiert und mit gestisch gezogenen Linien zueinander in Beziehung gesetzt. So setzt die in dieser Dialogkonstellation im Bild rechts situierte Gesprächspartnerin (GP_{R2}; Abb. 2a) zunächst mit ausgestrecktem Zeigefinger von ihr aus gesehen links auf Schulterhöhe einen Punkt in den Raum, der „Stockholm“ repräsentiert. Von dort aus zieht sie, ohne die Handfiguration zu ändern, eine leicht bogenförmige Linie nach unten bis in die Mitte ihres Gestenraums und verortet dort einen weiteren, verbal mit „Prag“ bezeichneten, Punkt (Abb. 2b).

(3) Von Stockholm nach Prag und Kiew.

- 01 GP_{R2}: Wir fahren [von Stockholm^{G7} nach Prag^{G8}
 02 und von Prag nach Kiew^{G9}
 03 und von da nach unten irgendwo]^{G10}



2a: Verortender Index: „Stockholm“ (G7).

2b: Verortender Index: „Prag“ (G8).

Abb. 2a–b: Verortung und In-Bezug-Setzen zweier Reiseziele.

Durch die Bewegung mit dem ausgestreckten Indexfinger zeichnet die Sprecherin eine Verbindungslinie zwischen den beiden Punkten (G7–G8) und erstellt so bereits ein einfaches Diagramm: ein schematisches Bild einer dyadischen Relation zwischen zwei Elementen bzw. zwei (gestischen) Zeichen (Peirce 1960: 157). Dabei initiieren die Punkte und Linie folgende Kern-Elemente des *Reisen*-Frames: QUELLE, WEG und ZIEL. Dies entspricht gleichzeitig dem in der kognitiven Linguistik prominenten *Weg*-Bildschema: ‚Source-Path-Goal‘ (vgl. Johnson 1987) mit einem Ausgangspunkt (A, ‚Source‘, „Stockholm“), einer Wegstrecke oder einem Vektor (‚Path‘, „nach“) und einem Zielpunkt (B, ‚Goal‘, „Prag“) (vgl. Cienki 2013; Mittelberg 2018). Vektoren dieser Art zeigen die REISERICHTUNG, ein weiteres Kern-Element des *Reisen*-Frame, an:

PATH SCHEMA (Johnson 1987: 28).



Im Vergleich zu solchen klar artikulierten, indexikalisch-ikonischen Gestenstrukturen (‚Index-Icon-Index‘ sind auch eher grobe Richtungsanweisungen im Raum und damit innerhalb des gedachten ‚Europa‘-Reisebereichs zu beobachten, beispielsweise mit „gen Osten“ (im Gestenraum rechts von der Sprecherin angesiedelt) oder „nach Süden“ (im Gestenraum unten angesiedelt). Solche gröberen, an Himmelsrichtungen, d.h. am kardinalen Ordinaten-System (vgl. Levinson 2003: 24–61; Thiering 2015, 2018: 92–188), orientierten Richtungsangaben wurden lautsprachlich mehrfach mit unspezifischen adverbialen Bestimmungen wie „irgendwo“ oder „irgendwie“ begleitet. Ob dies auf undifferenziertes geographisches Wissen oder eine noch nicht weit gediehene Reiseplanung, d.h. noch unklare Vorstellungen, zurückzuführen ist, muss, wie im Abschnitt 5 genauer dargelegt wird, stets anhand des Diskurskontextes bestimmt werden (vgl. Fricke 2007; Hassemer 2015; Wilkens 2003).

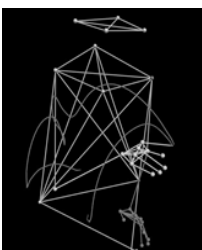


Abb. 3: Gestendiagramm: Stockholm, Prag und Kiew.

Wenn wir den dyadischen Streckenentwurf in Abbildung 3 weiterverfolgen, sehen wir, wie sich ein komplexeres Diagramm mit zusätzlichen Orten und Verbindungslinien bzw. Richtungen aufbaut („von Stockholm nach Prag und

von Prag nach Kiew und von da nach unten irgendwo“). Der Streckenverlauf wird in Abbildung 3 anhand von Motion-Capture-Daten visualisiert. Diese numerischen Daten ermöglichen uns, die Koordinaten von Punkten im dreidimensionalen Raum zu bestimmen und das allmähliche Entstehen von an sich unsichtbaren Verbindungslinien und größeren Diagrammstrukturen im Gestenraum sichtbar zu machen (vgl. Schüller und Mittelberg 2016, 2017).

Wie diese Veranschaulichungen zeigen, sind diese diagrammatischen Darstellungen von Reiserouten im Gestenraum von subjektiver, relativer und approximativer Natur. In der Einleitung wurde anhand von Transkript 1 bereits deutlich, inwiefern dort die zahlreichen indexikalischen Funktionswörter („shifters“ nach Jakobson 1971) ohne gestische Unterstützung nicht leicht disambiguiert werden können. Gleichzeitig wäre es schwierig, die Stationen und Wege dieses doch sehr skizzenartigen Diagramms (Abb. 3) ohne den vorhergehenden Diskurs und die synchrone sprachliche Nennung der Orte und bestimmte Deiktika nachzuvollziehen. Im Laufe der Bedeutungskonstitution kontextualisieren sich gestische und sprachliche Ressourcen gegenseitig (vgl. Jakobson 1956; Mittelberg und Waugh 2009).

Für einfache Konstruktionen des Typus „von X nach Y“ können wir hier bereits Folgendes festhalten: Diagramme bilden eine sich dynamisch aufbauende ikonische Struktur. Sie beginnt mit einer punktuellen Aktivierung X („Index“), gefolgt von einer Verbindungslinie „nach“ („Icon“) zum nächsten Punkt Y („Index“). Auf der Basis von verinnerlichteten Wissensstrukturen fungieren sie als konzeptuell-semiotisches Gerüst, in dem einzelne Frame-Elemente instantiiert und miteinander verbunden werden. Wir nehmen an, dass SprecherInnen kognitive Karten ausschnittartig vor Augen haben, wenn sie kognitiv und pragmatisch bedingte Diagramme in den Gestenraum projizieren, für sich selbst und ihr Gegenüber.

5. Phasen des Ineinandergreifens indexikalischer und ikonischer Modi im dialogischen Konstruieren von Reiserouten

Nachdem die vorhergehenden Abschnitte vereinzelte Einblicke in das multimodale Evozieren von Frames und Diagrammen gegeben haben, ist dieser Abschnitt dem schrittweisen Konstruieren einer Reiseroute in einem der Dialoge gewidmet. Unsere Analyse ergab drei Phasen, in denen indexikalische und ikonische Zeichenmodi mit unterschiedlicher Gewichtung und daraus resultierender semiotischer ‚Leistung‘ ineinandergreifen. Im Folgenden stellen wir die charakteristischen Eigenschaften der jeweiligen hervorstechenden cross-modalen Zeichenprozesse vor: die indexikalische, die indexikalisch-ikonische und die diagrammatisch-ikonische Phase.

5.1 Indexikalische Phase: (Inter-)Subjektive Äußerung und Anerkennung erster Wunschziele.

(4) Beginn der Planungen: Indexikalische Phase (00.41–01.15)

- 01 GP_{R1}: gibt es irgendwas was dir jetzt sofort aufgeploppt ist
 02 wo du so dachtest boah geil
 03 GP_{L1}: [**ich will**]G¹⁰ auf jeden Fall gern nach Barcelona
 04 da war ich noch nie
 05 GP_{R1}: oh [**das**]G¹¹ können wir machen
 06 GP_{L1}: und ich [**würd**]G¹² auch nochmal gern nach Lissabon
 07 weil das ist eigentlich ganz schön
 08 [**ich will**]G¹³ aber auch mal gern nach Stockholm
 09 da warst du auch noch nicht ne
 10 GP_{R1}: Was!
 11 [**du**]G¹⁴ hast jetzt schon drei Ziele genannt
 12 doch in [**Stockholm**]G¹⁵
 13 waren wir doch mit Wimma
 14 GP_{L1}: mit Wimma
 15 aber Wimma war doch in Lund
 16 oder ward ihr auch in Stockholm
 17 GP_{R1}: Doch da sind wir einmal hingefahren
 18 [**Ich möchte**]G¹⁶ gerne in Richtung
 19 GP_{L1}: [...]G¹⁷
 20 GP_{R1}: Osten
 21 ist mir eigentlich egal
 22 Kiew?
 23 Kommen wir dahin?
 24 GP_{L1}: ja
 25 Osten können wir machen
 26 GP_{R1}: ja ne



4a: Citing Geste signalisiert Zustimmung (G11).



4b: Shared Information Delivery PUOH Geste (G15).

Abb. 4a–d: Indexikalische Gesten: (Inter-)Subjektivität



4c: Gestenindex markiert subjektiven Wunsch (G16).



4d: GP_{L1} ahmt Index (G16) von GP_{R1} nach (G17).

Die Gesprächspartnerinnen GP_{L1} (links sitzend) und GP_{R1} (rechts sitzend) sind Arbeitskolleginnen. Sie beginnen ihre Überlegungen mit dem Nennen möglicher Reiseziele innerhalb Europas. GP_{L1} fällt direkt Barcelona als Wunschziel ein ((4), Z01–04). Ihre Reisepartnerin GP_{R1} reagiert positiv auf diesen ersten Vorschlag (Z05). Auf „das (können wir machen)“ zeigt sie mit ihrem ausgestreckten Zeigefinger auf ihre Gesprächspartnerin. Die Geste (Abb. 4a; G11) weist eine indexikalische Beziehung zum Gesagten auf und signalisiert Zustimmung („Acknowledgement Citing Gesture“, vgl. Bavelas 1994). GP_{L1} nennt Lissabon und Stockholm als weitere Ziele (Z06–08). GP_{R1} macht ihre Gesprächspartnerin darauf aufmerksam, dass sie bereits drei Ziele (von fünf möglichen) genannt hat (Z10–11). Auf „du“ öffnet sie kurz ihre rechte Handfläche, wobei ihre Fingerspitzen auf ihr Gegenüber zeigen; so signalisiert sie auf leicht entrüstete und amüsierte Weise, dass GP_{L1} etwas zu schnell vorausprescht. Die beiden lachen kurz.

GP_{R1} fährt fort, indem sie ihre Gesprächspartnerin an eine zurückliegende Reise erinnert (Z12–17). Auf den Ortsverweis „Stockholm“ führt sie mit der rechten Hand eine Zeigegeste mit nach oben gedrehter Handfläche und ausgestrecktem Zeigefinger aus (Abb. 4b; G15). Diese PUOH Geste mit indexikalischer Ausrichtung hat eine interaktiv-diskursive Funktion, indem sie auf einen von der Gesprächspartnerin zuvor genannten Ort anaphorisch Bezug nimmt und ihn dabei mit einer gemeinsamen Reise („waren wir doch“) verknüpft („Shared Information Delivery“ Geste nach Bavelas 1994; „common ground“, vgl. Clark 1996). „Stockholm“ wird als ein neues Kern-Element ZIEL des *Reisen*-Frame eingeführt.

Sich auf die gestellte Aufgabe besinnend, überlegt sich GP_{R1} nun selbst ein Wunschziel (Z18–22). Auf „ich möchte“ (Z18) beugt sie ihren Oberkörper nach vorne, hebt ihren rechten Zeigefinger und zeigt damit auf ihre Gesprächspartnerin (Abb. 4c; G16). Hierdurch kündigt sie das Initiieren eines neuen, ihrem Wunschziel ‚Osten‘ entsprechenden Frame-Elements an, leitet die Aufmerksamkeit ihrer Gesprächspartnerin und bekräftigt gleichzeitig ihre Sprecherrolle. Dann zieht sie ihren Zeigefinger an ihren spitz zugezogenen Mund, was hier Nachdenken und Zögern andeutet. Während GP_{R1} überlegt (Z18), ahmt GP_{L1} die vorhergehende Zeigepose ihrer Partnerin in leicht geänderter Ausführung nach. GP_{L1} zieht ihren rechten Zei-

gefingert auf Augenhöhe und zeigt auf ihre Gesprächspartnerin (Abb. 4d; G17). Mit dieser ‚Acknowledgement Citing‘ Geste (vgl. Bavelas 1994) signalisiert sie ihre Aufmerksamkeit und gibt Rückmeldung, dass sie GP_{R1} zuhört. Sie stimmt so dem Vorschlag von GP_{R1} zu, in den Osten zu fahren (Z24–25).

In dieser ersten Sequenz dienen die indexikalischen Gesten dem Vorschlagen und Zustimmung hinsichtlich erster Reisezielvorschläge, dem Verweis auf gemeinsame Erfahrungen (‚common ground‘; vgl. Clark 1996) sowie dem Signalisieren von Aufmerksamkeit. Somit kommt diesen Gesten eine primär (inter-)subjektive und interaktive Funktion zu. Auch wenn dabei erste Frame-Elemente verbal initiiert werden, spielt die gestische Verortung von Punkten im Raum und ihre Relation zueinander hier noch keine zentrale Rolle.

5.2 Indexikalisch-ikonische Phase: Richtungen und erste Verbindungslinien zwischen Reisezielen

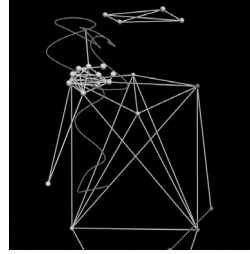
In der folgenden Gesprächssequenz geht es verstärkt um Verortungen und mehr oder weniger grobe Richtungen, zunächst jedoch noch auf approximative Art und Weise.

(5) Zweite Phase: *Reisen*-Frame-Elemente BEREICHE und RICHTUNGEN.

- 01 GP_{L1}: man könnte [**auch**^{G18}
 02 **auch wenn man sich im Osten bewegt**^{G19}
 03 **sich von Norden**^{G20} **nach Süden**
 04 **durch eh**]^{G21}
 05 schlagen
 06 und dann [**zum Schluss**]^{G22} noch so zwei drei Tage an Strand
 07 GP_{R1}: das heißt wir machen quasi [**erst so’n bisschen**]^{G23}
 08 GP_{L1}: [^{G24} nach Kroatien
 09 GP_{R1}: [**erst son bisschen Richtung**]^{G25} [**eh**]^{G26}
 10 Schweden
 11 [**dann rüber**]^{G27}
 12 [**nach**]^{G28}
 13 GP_{L1}: [**Estland Lettland**]^{G29} [^{G30}
 14 GP_{R1}: so in die Richtung
 15 dann könnten wir [**runter**]^{G31}
 16 das heißt wir brauchen noch [**irgendwie**]^{G32}
 17 Griechenland oder so was



5a: Zögernder Index zwecks Verortung (G18).



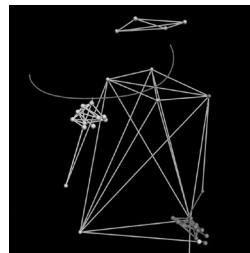
5b: Gestendiagramm: „sich von Norden nach Süden durch eh schlagen“ (G21).

Abb. 5 a–b: Ungefäher Routenentwurf anhand kardinaler Referenzpunkte.

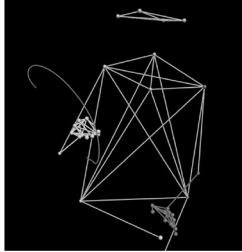
GP_{L1} beginnt einen ersten groben, nun an den Himmelsrichtungen orientierten, Routenverlauf von Nord-Ost nach Süden ((5), Z01–06), welcher durch fünf Handgesten allmählich im Gestenraum entsteht. Auf „auch“ (Z01) zieht sie ihren rechten Zeigefinger nach oben (Abb. 5a; G18). Doch während sie verbal den möglichen Beginn der Reise im Osten anspricht (Z02), zeichnet sie mit ihrer rechten Hand zunächst eine leicht gebogene Linie nach rechts. Dies dient einer zunächst noch groben Richtungsangabe. Sie nimmt dabei auf das verbal geäußerte „Osten“ gestisch Bezug, welches sich auf ihrer mentalen Karte aus ihrer Perspektive („viewpoint“; vgl. McNeill 1992; Sweetser 2012) rechts von ihrem Körper befindet. Auf „Norden“ (Z03) markiert sie mit ihrem Zeigefinger ein wenig höher im Gestenraum einen imaginären Punkt und zieht dann auf „nach Süden durch eh schlagen“ (Z03–04) eine Schlängellinie bis in den unteren Teil des Gestenraums (Abb. 5b; G21). Diese Linienführung drückt nicht nur die verbal erläuterte Reise-richtung von Norden nach Süden aus, sondern impliziert zusätzlich durch die Schängel-Gestalt, dass der genaue Verlauf noch zu klären ist: Welche Städte- und Länder-Elemente noch dazu kommen, muss erst noch ausgehandelt werden. Auf „zum Schluss“ (Z06) dreht GP_{L1} beide offenen Handflächen mit gespreizten Fingern relativ niedrig im Gestenraum nach unten, was einen Prozess des zur-Ruhe-Kommens suggeriert (vgl. Rekittke u.a. 2015).



6a: Geste markiert Hesitation (G26) vor Schweden.



6b: Diagramm: „dann rüber nach“ Estland Lettland (G27/28).



6c: GP_{R1} „runter“ (G31) (Richtung Griechenland).



6d: Hand bleibt in situ auf „irgendwie“ (G32).

Abb. 6 a–d: Routenverlauf in Etappen.

GP_{R1} greift daraufhin die von GP_{L1} vorgeschlagene, noch relativ grobe Route auf: Auf „erst so’n bisschen“ (Z07) zieht sie zunächst mit der rechten Hand eine Linie von links nach rechts mittig in ihren Gestenraum und beginnt dann ein eigenes Diagramm aus ihrer Perspektive. Da schiebt GP_{L1} „nach Kroatien“ (Z08) noch ein; sie ist in Gedanken noch bei ihrem Wunsch nach Entspannung gen Reiseende. Indes führt GP_{R1} fort und wiederholt „erst so’n bisschen Richtung“ (Z09). Während dieser unpräzisen verbalen Äußerung zeichnet sie eine Linie von links unten nach rechts oben in den Gestenraum. Obwohl sie ihr erstes Wunschziel noch nicht ausgesprochen hat, weist ihr Zeigefinger bereits nach oben und somit nach Norden. Sie überlegt kurz: „eh“ (Z09). Hierbei verharrt ihre rechte Hand in derselben Position, sie streckt die Finger nach oben und macht eine wellenförmige Bewegung, welche hier eine kurze Denkpause suggeriert und gestisch füllt (Abb. 6a; G26). Sie greift dann die Idee ihrer Gesprächspartnerin auf und konkretisiert die grob genannten Himmelsrichtungen. Auf „dann rüber“ (Z11) zeichnet sie eine weitere Kurve, nun von oben mittig nach rechts außen (Abb. 6b; G27), und setzt dort einen nächsten Punkt, jedoch ohne ihn verbal zu bezeichnen. Ihre ungenaue Zielvorstellung drückt sich durch eine weitere wellenförmige Bewegung der Finger ihrer rechten Hand aus. GP_{L1} füllt nun die Lücke im dynamisch-diagrammatischen Planungsprozess, indem sie „Estland, Lettland“ (Z13) als mögliche Zwischenstopps im Osten benennt. Hierbei zieht sie mit der rechten Hand eine lange, aufsteigende Linie von links unten nach rechts oben. Sie greift somit die Route ihrer Gesprächspartnerin auf und integriert diese zusätzlichen Informationen aus ihrer eigenen Perspektive. Hierbei zeichnet sie Estland und Lettland in Relation zu Deutschland ein. Objektiv betrachtet, befindet sich Deutschland süd-westlich von den baltischen Ländern. Auf ihre subjektiv ausgerichtete, aber gleichzeitig aus allozentrischer Perspektive dargestellte mentale Karte von Europa projiziert, befinden sich Estland und Lettland relativ zu Deutschland im rechten oberen Bereich, was sie durch eine gerade Linie nach rechts oben illustriert.

GP_{R1} hört und schaut ihrer Gesprächspartnerin aufmerksam zu und signalisiert GP_{L1} durch Kopfnicken auf „so in die Richtung“ (Z14) ihr Einverständnis. Sie fügt dann Griechenland als weitere potentielle Etappe im Süd-Osten von Europa hinzu (Z15–17). Dadurch eröffnet sich in Bezug auf diese

größere geographische Region ein weiterer Reisebereich (Kern-Frame-Element). Während sie „runter“ (Z15) sagt, geht parallel ihre rechte Hand mit ausgestrecktem Zeigefinger nach links mittig unten (Abb. 6c; G31) und visualisiert so im Gestenraum einen Weg, der sich wahrscheinlich auf ihrer kognitiven Karte von Estland oder Lettland in den Süden nach Griechenland erstreckt. Auf „irgendwie“ (Z16) wippen ihre Finger erneut hoch und runter (Abb. 6d; G32), dieses Mal tiefer im Gestenraum als in Abbildung 6a.

An dieser Stelle lässt sich bereits feststellen, dass diese Reiseplanung durch einen kooperativen Konversationsstil geprägt ist: Die Gesprächspartnerinnen, besonders GP_{R1}, greifen die Vorschläge ihres Gegenübers auf und bauen sie in ihre eigenen Reisevorstellungen – und teilweise auch in die entsprechenden diagrammatischen gestischen Streckenskizzen – mit ein. Dadurch, dass sich die GesprächsteilnehmerInnen gegenüber sitzen, müssen sie die gestischen Richtungsangaben, spiegelbildlich verfolgen. Dass dabei ein absoluter Referenzrahmen mit den vier Kardinalpunkten bzw. Himmelsrichtungen veranschlagt wird, hilft sicherlich. Wir gehen davon aus, dass hier mentale Kartenansichten ausschnittshaft in den Raum projiziert werden: Norden wird im oberen und Süden im unteren Gestenraum angesiedelt, Osten jeweils rechts und Westen jeweils links der Sprechenden (vgl. Levinson 2003: 24–61; vgl. Thiering 2015, 2018: 92–188). Dabei werden Informationen unter pragmatischem Druck in ein logisches mentales Raummodell übertragen, dass an die jeweilige Kommunikationssituation adaptiert ist.¹¹ Bei Unentschiedenheit warten sie ab oder schlagen mögliche Ziele und Richtungen (Kern-Frame-Elemente) vor.

5.3 Diagrammatisch-ikonische Phase: Einigung auf komplette Reiseroute

Die folgende Sequenz ist uns bereits in Teilen bekannt: In der Einleitung befindet sich das Transkript ohne Gestenannotation (1), und in Abschnitt 4.2 haben wir uns mit den ersten Diagrammabschnitten beschäftigt (Z01–03 in (3) entspricht Z07–09 in (6)). Hier präsentieren wir nun die komplette, multimodal kontextualisierte Reiseroute, auf die sich die Gesprächspartnerinnen letztendlich einigen.

(6) Letzte Phase: Resümierendes Routendiagramm im peripheren Gestenraum.

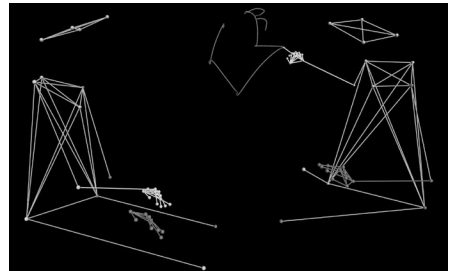
- 01 GP_{R1}: ich war schon in Prag
- 02 Stadt der goldenen Dächer
- 03 großartig
- 04 richtig schön
- 05 **[können wir auch machen]**.^{G33}
- 06 Pass auf!
- 07 Wir fahren **[von Stockholm nach Prag]**

- 08 **und von Prag nach Kiew**
 09 **und von da nach unten irgendwo]^{G34}**
 10 GP_{L1}: ich weiß überhaupt nicht wo Kiew liegt [**relativ zu Prag]**^{G35}
 11 GP_{R1}: []^{G36}
 12 ich auch nicht %lacht
 13 GP_{L1}: %lacht aber gut angetäuscht^{G37}
 14 GP_{R1}: []^{G38} ich denke
 15 **nee]**^{G39} ich denke wir fahren dann so weißte [von
 16 **da**
 17 **nach da**
 18 **rüber**
 19 **runter**
 20 **und dann nach da und dann wieder nach Hause]**^{G40}.
 21 Nein?
 22 GP_{L1}: doch

Nach fortgeschrittenem Gespräch sind sich die Gesprächspartnerinnen einig zunächst nach Schweden und Kiew zu reisen. GP_{L1} fällt Prag als potentieller Zwischenstopp ein, was von GP_{R1} sofort positiv aufgegriffen wird ((6), Z01–05). GP_{L1} gibt kurz zu bedenken, dass ihr die genaue Lage von Kiew nicht bekannt sei (Z10). Wir können somit annehmen, dass sie keine differenzierte mentale Karte von der Region (*Reisen*-Frame-Element BEREICH) hat. In diesem Moment streckt sie ihren rechten Zeigefinger kurz nach oben und deutet dabei – ohne einen Ort zu präzisieren – in den Raum (G35). GP_{R1} gibt dann zu, selbst nicht zu wissen, wo genau Kiew liegt (G36). Die nicht konkret auf ein Ziel deutende Zeigegeste scheint dabei zu signalisieren, dass sie die geographische Lage der beiden Städte auch nicht gut kennt. Angesichts dieser ‚leeren Flecken‘ auf ihrer mentalen Karte bleiben auch ihre gestischen Verweise eher vage und unentschlossen. Als nun beide einsehen, dass sie kein Detailwissen hinsichtlich der anvisierten Länder haben, fangen sie an zu lachen. GP_{L1} begegnet GP_{R1} mit „aber gut angetäuscht“ (Z13), zeigt mit ihrem ausgestreckten Zeigefinger auf sie und lacht weiter (G37).



7a: Beginn und Abbruch des Routendiagramms im geteilten zentralen Gestenraums (G 39).



7b: Geteilte Perspektive auf komplettes Gestendiagramm in der Peripherie (G40).

Abb. 7a–b: Erst gespiegelte, dann geteilte Perspektive auf Diagramm in der Gestenraum-Peripherie.

Abschließend fasst GP_{R1} den geplanten Routenverlauf zusammen (Z15–20). Auf „nee“ setzt sie kurz ihren Zeigefinger links oben in ihrem Gestenraum an (Abb. 7a), nimmt dann aber einen Perspektivwechsel vor, indem sie zunächst ihren Blick nach rechts wendet und auf „von“ (Z15) ihren Zeigefinger etwas oberhalb ihrer Kopfhöhe in der linken äußeren Peripherie neu platziert. Sie lädt ihre Gesprächspartnerin ein, dieselbe, also nicht gespiegelte, Perspektive auf die folgenden multimodalen Erläuterungen einzunehmen, indem sie bisher nicht genutzte Segmente in der Peripherie des gemeinsamen Gestenraums aktiviert. Nun wendet auch GP_{L1} ihren Blick seitwärts und folgt der Zeigegeste von GP_{R1} , die den Anfangspunkt eines kompletten Routendiagramms bildet (Abb. 7b; G40): Letztere zieht zunächst auf „nach da“ (Z17) eine Linie nach unten und auf „rüber“ (Z18) eine kurze, relativ gerade Linie nach rechts. Auf „runter“ (Z19) zeichnet sie entsprechend eine schräg nach links unten verlaufende Linie und auf „und dann nach da“ (Z20) bewegt sie ihren Zeigefinger weiter nach links und schräg nach oben. Da sie die Zielorte kurz zuvor bereits benannt hatte und die Orte, wenn auch skizzenhaft, schon gestisch zueinander in räumliche Beziehung gesetzt wurden, weist GP_{R1} in diesem letzten Teil der multimodalen Synopse lediglich mit sprachlichen Deiktika auf die Punkte ihres Gestendiagramms hin, welche die bereits besprochenen Stationen der Reise repräsentieren. Auf „und dann wieder nach Hause“ (Z20) zieht sie den Zeigefinger noch weiter nach oben und richtet hierbei ihren Blick zurück auf ihre Gesprächspartnerin. Sie hält den Zeigefinger noch kurz an dieser Stelle, wartet auf eine Reaktion und erfragt dann mit „nein?“ Zustimmung oder Ablehnung (Z21). Nachdem GP_{L1} ihr mit einem „doch“ zustimmt, nimmt sie den Finger wieder runter. Der multimodal gestaltete Planungs- und Einigungsprozess ist damit abgeschlossen: Die Reiseroute steht.

Zusammenfassend lässt sich folgendes hervorheben: Im Vergleich zu den eingangs groben Richtungsanzeigen und den ersten vagen Schlängellinien und skizzenhaften Teildiagrammen ist dieses abschließende Diagramm (Abb. 7b) von einer klaren Linienführung gekennzeichnet und lässt sich mit Blick auf die Geographie Europas durchaus nachvollziehen (vgl. Schüller und Mittelberg 2017: 129). Die mentalen Karten der beiden Gesprächspartnerinnen werden so – aus den jeweiligen Perspektiven – beiden auf gestisch transformierte Weise semiotisch habhaft und dienen der geteilten Aufmerksamkeit und Koordination von Vorschlägen und Präferenzen. Dabei scheinen die ineinandergreifenden Diagramme und Framestrukturen das Gesprächsgeschehen zu unterfüttern, indem sich die Gesprächspartnerinnen jeweils auf die multimodalen Aussagen der anderen beziehen und ihre eigenen Ausführungen darauf aufbauen. Das Ergebnis dieses kooperativ geführten Gesprächs ist eine Reiseroute, die sich nachvollziehen und umsetzen lässt.

6. Diskussion: Indexikalische Zeichenmodi orchestrieren dynamische multimodale Gesprächskontexturen

Die indexikalische Bedingtheit multimodaler Interaktion wurde in diesem Beitrag als ein den Gebrauch gestischer und sprachlicher Ressourcen stark motivierendes Prinzip erörtert und auf ihre materiellen und kontextuellen Grundlagen zurückgeführt. Unsere Analyseergebnisse bestätigen nun, dass der Gebrauch von Inhalts- und Funktionswörtern sowie von primär ikonischen und primär indexikalischen Gesten prinzipiell perspektiviert (*view-pointed*), teilhaft (metonymisch reduziert) und situiert verankert (pragmatisch bedingt) ist. Dabei haben sich drei unterschiedlich wirkende, an der Orchestrierung von Diskursinhalten, Einigungsprozessen und Ausdrucksmodalitäten teilhabende, indexikalische Modi herauskristallisiert, die im Folgenden kurz dargelegt werden: semantische Bezüge, interaktive Indizes, sowie intra- und intermodale Bezüge.

6.1 Verkörperte Referenzpunkte

Das mentale Aufspannen von Framestrukturen durch das gestische Andeuten und Verbinden von Frame-Elementen wird wenigstens zum Teil durch verkörperte Referenzpunkte (*embodied reference points*, vgl. Mittelberg 2017) angestoßen und dabei cross-modal und situativ kontextualisiert. Unter verkörperten Referenzpunkten verstehen wir in Anlehnung an Langackers Konzept des Referenzpunktes (*reference-point constructions*; vgl. Langacker 1993) gestische Indizes, die zentrale Punkte in der Bedeutungskonstruktion darstellen und mentalen Zugang zu Zielbereichen gewähren (vgl. Mittelberg 2006; Talmy 2018).¹² Gesten ermöglichen einen dynamisch-körperlich geleiteten mentalen Zugang; wir nennen dies einen kognitiv-aktionalen Zugang. In unseren Daten dienen Gesten beispielweise als Zugangsmöglichkeiten zu kognitiven Karten und semantischen Framestrukturen (z.B. ‚Europa‘, oder ‚Süden‘). Durch das Markieren von Punkten und Regionen im Gestenraum, auf die sie teils recht präzise und teils recht vage verweisen (z.B. ‚irgendwie nach Süden‘), können indexikalische Gesten körperbasierte Ansatzpunkte von und ‚Haltepunkte‘ in sich entfaltenden semiotischen Strukturen wie etwa virtuellen Diagrammen darstellen.

Bei diesen Praktiken der multimodalen Verortung und Andeutung von zu inferierenden semantischen Zielbereichen bzw. Netzwerken kommen Prozesse der pragmatischen Inferenz und somit metonymische Modi zum Tragen, die tendenziell auf Basis von Kontiguitätsbeziehungen operieren (vgl. Mittelberg 2017; Panther und Thornburg 2003). Diese ermöglichen es den GesprächspartnerInnen, weiterführende Assoziationen, z.B. zu anderen Frame-Elementen oder auch anderen Frames, gedanklich zu verfolgen, multimodal weiterzuentwickeln und dabei zur gemeinsamen Betrachtung anzubieten. Sie bieten Schnittstellen für neue ‚Brückenschläge‘ und damit für kreatives Denken und semiotisches Handeln während des Dis-

kursgeschehens (vgl. Coulson 2001; Dancygier und Sweetser 2014; Ehmer 2011; Ziem 2014). Unsere Analysen unterstützen gleichzeitig die Annahme, dass durch die indexikalische Annäherung an im Kontext oder in der Vorstellung existierende Strukturen und Objekte (siehe ‚imaginary material anchors‘; vgl. Hutchins 2005; Abschnitt 3.2) diese subjektiven und intersubjektiven Praktiken des diagrammatischen Denkens und Gestikulierens an dialogischen Prozessen der ‚distributed cognition‘ teilhaben (vgl. Streeck u.a. 2011).

6.2 Interaktive Indizes

Indexikalische Handformen und Handbewegungen sind, wie bereits betont, die prädominanten gestischen Formen in den hier analysierten Daten. Ein zentrales Ergebnis dieser Studie ist, dass unterschiedlich ausgeführte indexikalische Gesten im dialogischen Konstruieren von Reiserouten ein breites Spektrum an interaktiven und diskursiven Funktionen aufweisen (vgl. Bavelas u.a. 1995; Kendon 2004). Dies wird besonders deutlich im allmählichen Aushandeln von Kompromissen, indem die DialogpartnerInnen zum Beispiel durch punktuelle, grob die Richtung anzeigende sowie linienproduzierende Zeigefinger verschiedene virtuelle Orte und Routenabschnitte im Gestenraum produzieren (vgl. Fricke 2007).

Zusammenfassend lässt sich sagen, dass es Gesprächsphasen gibt, in denen bestimmte indexikalische Praktiken stärker vertreten sind als andere: So ist zum Beispiel die erste Phase (Abschnitt 5.1) prädominant indexikalisch hinsichtlich des multimodalen Ausdrucks subjektiver Wünsche und intersubjektiver Reaktionen. Interaktive, direkt auf die Gesprächspartnerin gerichtete Gesten, verstärken das Äußern von Wünschen und Vorschlägen („ich will auf jeden Fall“^{G10}, „ich will aber auch mal gern“^{G13}, „ich möchte gern in Richtung“^{G16}). Dazu kommen die gestischen Indices, die sich auf gemeinsame Erfahrungen und geteiltes Wissen und damit vergleichbare Instantiierungen des *Reisen*-Frame („waren wir doch“^{G15}; ‚Shared Information Delivery‘ Geste nach Bavelas 1994) oder auf bereits gemachte Vorschläge beziehen. Letztere gehen zum Beispiel mit einem lautsprachlichen Anerkennen und Zustimmung („das können wir machen“^{G11}, „können wir auch machen“^{G33}) einher („Acknowledgement Citing‘ Gesten nach Bavelas 1994).

Interaktive Gesten, die dem Gesprächsmanagement dienen, konnten wir ebenfalls beobachten: So zum Beispiel ‚Giving Turn‘-Gesten oder PUOH-Gesten (vgl. Müller 2004), mit denen Sprechende signalisieren, dass ihre Hände leer sind und sie eine Antwort vom Gesprächspartner erbitten („Seeking‘-Geste in Abb. 2 nach Bavelas 1994). Hier schließen sich zahlreiche weiterführende Fragestellungen hinsichtlich interagierender Dimensionen wie Einstellung („stance“), Unentschiedenheit, Ironie, gegenseitige Herausforderung oder Besänftigung an.

6.3 Intramodale und intermodale Bezüge

Wie die Analysen gezeigt haben, sind die multimodalen Daten nicht nur reich an indexikalischen Elementen, die der einmaligen Verortung oder dem punktuellen Bezug dienen, sondern weisen zudem sprachliche und gestische Wiederaufnahmen und Rückverweise auf. Dabei lassen sich intramodal bedeutungsstiftende Prozesse von intermodal operierenden Modi indexikalischer und ikonischer Natur wie folgt unterscheiden.

Im sprachlichen Diskurs stellen anaphorische Verweise, etwa durch Personal- und Demonstrativpronomina, *intramodal* Rückbezüge her. In der gestischen Modalität beziehen sich Zeichen aufeinander, indem zum Beispiel Zeigegesten auf bereits gestisch gesetzte Positionen im Gestenraum erneut verweisen (vgl. Fricke 2007). Auch das Wiederaufnehmen von ikonischen Gestenformen, die in derselben Gesprächssequenz bereits von der Sprecherin selbst oder vom Gegenüber in ähnlicher Weise erzeugt wurden, wirkt kohäsiv (vgl. ‚catchments‘, ‚cohesives‘ oder ‚mimicry‘, McNeill 2005). Wir erinnern uns an die in Abb. 1e–f wiederholte Schwimmgeste als Beispiel für das ikonische Nachahmen einer von der Gesprächspartnerin ausgeübten Geste, die dem ‚Alignment‘ diene.

Intermodale (‚cross-modal‘) bedeutungsstiftende Prozesse kommen besonders dann zum Tragen, wenn potentiell vieldeutige gestische oder sprachliche Formen nur durch in der anderen Modalität synchron veräußerte Hinweise disambiguiert und verstanden werden können. Sprachliche Deiktika wie „da“, „darunter“ und „von dort“ beziehen ihre lokale Bedeutung aus dem sprachlichen und außersprachlichen Kontext, nämlich ‚semiotische Kontexturen‘ (vgl. Jakobson 1956) oder ‚contextures of action‘ (vgl. Goodwin 2011). Wie wir sehen konnten, funktionieren in den Raum gesetzte Zeigegesten, welche bestimmte, in der Rede genannte Orte auf einer erst im Gesprächsfluss entstehenden virtuellen Karte evozieren, ähnlich wie sprachliche Deiktika, indem sie ihre Bedeutung zum großen Teil aus der multimodalen Gesamtäußerung ziehen (vgl. Kendon 2004). Sie können so zu den kontextbedingten Funktionselementen gezählt werden, die Jakobson (1971), wie oben erwähnt, als Verschieber (‚shifters‘) und Talmy (2000) als ‚closed-class items‘ (im Gegensatz zu Inhaltswörtern und ‚open-class items‘) bezeichnen. Um eine Aussage über die innere oder äußere Welt zu machen, sind bekanntlich beide Wortarten bzw. Zeichentypen, eine Kombination von Indizes und Ikonen, nötig (vgl. ‚dicisigns‘, Peirce 1960): Ohne einen Index kann ein Ikon keine lokale Bedeutung entfalten, was für bildliche wie sprachliche und gestische Zeichen gilt (vgl. Mittelberg 2006; Stjernfelt 2014).

Nun konnten wir hier eine zusätzliche Variante bimodal erzeugter Indexikalität beobachten, welche sich als konstitutiv für Diagramme erwies, nämlich beim Erstellen des durch eine klare Linienführung hervorstechenden Routendiagramms (Abb. 7b). Das von GP_{R1} in die Luft gezeichnete diagrammatisch-ikonische Gerüst unterfüttert virtuell durch indexikalisch gesetzte Punkte und die Verbindungslinien eine zeitgleiche Aneinanderkettung von

indexikalischen Funktionswörtern. Hier handelt es sich also nicht um eine gegenseitige Stütze von indexikalischen Gestenpunkten und in der Lautsprache genannten Eigennamen von Städten oder Ländern oder anderen inhaltlichen Frame-Elementen. Es sind Kombinationen bestehend aus Präpositionen, Demonstrativpronomina, Konjunktionen sowie Lokal- und Zeitadverbien („darunter“, „rüber“, „dann“, „und“), und entweder primär deiktischen Gesten (Punkte, die durch das Verharren auf einer Position im Gestenraum entstehen) oder eben das Ziehen von Verbindungslinien. Letztere sind die einzigen ikonischen Elemente in dieser Routenbeschreibung. Um die anaphorischen Prozesse in beiden Modalitäten nachzuvollziehen, muss die Gesprächspartnerin den bisherigen cross-modalen Erklärungen unter Berücksichtigung räumlicher Koordinaten, wechselnden Perspektiven und zeitlichen Abfolgen oder Korrelationen von sprachlichen und gestischen Zeichen sehr genau gefolgt sein.

7. Abschließende Bemerkungen

Angesichts der sprachlichen und gestischen Strategien, die DialogpartnerInnen in der hier präsentierten Studie anwandten, scheinen Diagramme und Framestrukturen nicht nur das Entstehen von gestischen Darstellungen zu motivieren. Sie dienen auch dem kollaborativen, multimodalen Entwickeln weitergehender Assoziationen und dem gegenseitigen Verstehen und Verständigen der GesprächspartnerInnen. Aus den sich herauskristallisierten Verfahren der dialogischen Bedeutungskonstitution erwiesen sich, neben ikonisch-diagrammatischen Praktiken, insbesondere verschiedene Modi und Grade von Indexikalität als die orchestrierenden Prinzipien in solchen dynamischen multimodalen Kontexturen. In dieser Studie lag der Fokus auf semantischen Frames und wie deren Elemente gestisch und/oder verbalsprachlich instantiiert werden. Der nächste Schritt besteht darin, in denselben Daten die syntaktischen Realisationen (nicht nur) der hier besprochenen Frames cross-modal zu untersuchen. Das FrameNet und Konstruktikon des Deutschen sind dafür reichhaltige Ressourcen.

Die untersuchten Reiseplanungen geben zudem Anhaltspunkte, inwiefern die GesprächspartnerInnen über ähnlich ‚gefüllte‘ bzw. ausdifferenzierte Wissensstrukturen von Städten, Ländern und Kulturen verfügen. So konnten wir den Dialogen nicht nur unterschiedliche Wissens- und Erfahrungsbestände seitens der GesprächsteilnehmerInnen entnehmen, sondern auch Präferenzen, erneut in bereits selbst bereiste Städte und Länder zu fahren (d.h. bezüglich derer Vorwissen und durch persönliche Erfahrungen angereicherte Wissensstrukturen bestehen) oder lieber Neuland zu erkunden. Letzteres würde nicht nur andere Instantiierungen von Frame-Elementen ermöglichen, sondern auch neue Verbindungen zwischen Frames.

Auf der Basis der hier dargelegten, auf qualitativen Analysen beruhenden Beobachtungen bietet es sich an, die in dieser Studie offengelegten Tendenzen unter Zuhilfenahme quantitativer Methoden im gesamten *MuSKA*

Datenkorpus, insbesondere in den numerischen Motion-Capture Daten, auf intramodale Muster und crossmodale Clusterbildung hin zu untersuchen.

Anmerkungen

- 1 Gemäß FrameNet des Deutschen: (<https://gsw.phil.hhu.de/framenet/frame?id=974>).
- 2 GP steht für „Gesprächspartner/in“, das tiefgesetzte _{R1} für „rechts“ im Dialog 1 sitzend (mit Referenz auf Abbildungen im Beitrag. (1) verweist auf „Transkript (1)“ und Z01 auf „Zeile 01“.
- 3 Zu symbolischen „Emblemen“ und Gestenformen mit relativ hohem Konventionalisierungsgrad, siehe Bressen und Müller 2017; Ladewig 2020; Kendon 2004; McNeill 1992; Müller 2017.
- 4 Die Teilhabe von Gesten an kognitiven Operationen wie Blending oder konzeptueller Integration (vgl. Fauconnier und Turner 2002) hat bereits wachsende Beachtung gefunden: vgl. z.B. Liddell (2003); Parrill und Sweetser (2004); Sweetser (2007, 2012); Sweetser und Stec (2016).
- 5 Siehe dazu FrameNet UC Berkeley (<https://framenet.icsi.berkeley.edu> (31.5.2021)) und FrameNet des Deutschen (<https://gsw.phil-fak.uni-duesseldorf.de> (31.5.2021)).
- 6 Siehe zur Konstruktion des Deutschen (vgl. <https://gsw.phil.hhu.de/constructiconold> (31.5.2021)).
- 7 Siehe Bühlers (1934/1982) Begriff der „Deixis am Phantasma“: siehe Fricke (2007) für einen Überblick verschiedener Erscheinungsformen der Deixis (z.B. am Vorstellungs-, Wahrnehmungs- oder Zeichenraum); siehe McNeill u.a. (1993) zum Begriff der abstrakten Deixis.
- 8 *Niederschlag*-Frame: „Wasser in fester oder flüssiger Form (NIEDERSCHLAG) fällt an einem bestimmten ORT und zu einer bestimmten ZEIT vom Himmel, was für eine bestimmte DAUER anhält. Die MENGE_PRO_EINHEIT oder die MENGE des Niederschlags kann auch angegeben werden“ (<https://gsw.phil.hhu.de/framenet/frame?id=943> (31.5.2021)).
- 9 *Umgebungstermperatur*-Frame: „Die Temperatur in einer bestimmten UMGEBUNG, bestimmt durch ZEIT und ORT, wird angegeben“ (<https://gsw.phil.hhu.de/framenet/frame?id=25> (31.5.2021)).
- 10 *Kleidung tragen*-Frame: „Die Wörter, die diesen Frame evozieren, beziehen sich auf die KLEIDUNG, die der TRÄGER (oder ein bestimmtes KÖRPERTEIL des TRÄGERS) anhat“ (<https://gsw.phil.hhu.de/framenet/frame?id=280> (31.5.2021)).
- 11 Wir danken Martin Thiering (pers. Kommunikation) für diesen Hinweis.
- 12 Langacker (2009: 52) definiert die menschliche Fähigkeit, Referenzpunkte zu kreieren, wie folgt: „This is our ability to invoke one conceived entity as a reference point in order to establish mental contact with another, i.e. to mentally access one conceived entity through another“.

Literatur

- Bavelas, Janet (1994). Gestures as part of Speech. Methodological implications. *Research on Language and Social Interaction* 27, 201–221.
- Bavelas, Janet, Nicole Chovil, Jan L. Coates und Lori Roe (1995). Gestures specialized for dialogue. *Personality and Social Psychology Bulletin* 21, 394–405.
- Brenger, Bela und Irene Mittelberg (2015). Shakes, Nods and Tilts: Motion-Capture Data Profiles of Speakers' and Listeners' Head Gestures. In: Ferré Gaëlle und Mark Tutton (eds.). *Proceedings of the 3rd Gesture and Speech in Interaction (GESPIN) Conference*, Nantes, 43–49.
- Bressem, Jana (2013). A linguistic perspective on the notation of form features in gestures. In: Cornelia Müller, Alan Cienki, Ellen Fricke, Silvia H. Ladewig, David McNeill und Sedinha Tessendorf (eds.). *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction* (38.1). Berlin und Boston: de Gruyter Mouton, 1079–1098.
- Bressem, Jana und Cornelia Müller (2017). The “Negative-Assessment-Construction” – A multimodal pattern based on a recurrent gesture? *Linguistics Vanguard* 3, 1, 1–9.
- Bühler, Karl (1934/1982). *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Stuttgart und New York: Fischer.
- Cienki, Alan (2013). Mimetic schemas and image schemas in cognitive linguistics and gesture studies. *Review of Cognitive Linguistics* 11, 417–441.
- Clark, Herbert H. (1996). *Using Language*. Cambridge, UK: Cambridge University Press.
- Coulson, Seana (2001). *Semantic Leaps: Frame-shifting and Conceptual Blending in Meaning Construction*. Cambridge, UK: Cambridge University Press.
- Dancygier, Barbara und Eve E. Sweetser (2014). *Figurative Language*. Cambridge, UK: Cambridge University Press.
- Downs, Roger und David Stea (1982). *Kognitive Karten. Die Welt in unseren Köpfen*. New York: Harper & Row.
- Du Bois, John W., Stephan Schuetze-Coburn, Susanna Cumming und Danae Paolino (1993). Outline of discourse transcription. In: Jane A. Edwards und Martin D. Lampert (eds.). *Talking Data: Transcription and Coding in Discourse Research*. Hillsdale, NJ: Erlbaum, 45–89.
- Ehmer, Oliver (2011). *Imagination und Animation. Die Herstellung mentaler Räume durch animierte Rede*. Berlin und New York: de Gruyter Mouton.
- Enfield, Nick (2009). *The Anatomy of Meaning. Speech, Gestures and Composite Utterances*. Cambridge, UK: Cambridge University Press.
- Fauconnier, Gilles (1985/1994). *Mental Spaces. Aspects of Meaning Construction in Natural Language*. Cambridge, MA: MIT Press.
- Fauconnier, Gilles (1997). *Mappings in Thought and Language*. Cambridge, UK: Cambridge University Press.
- Fauconnier, Gilles und Mark Turner (2002). *The Way We Think*. New York: Basic Books.
- Fillmore, Charles (1977). Scenes-and-frames semantics. In: Antonio Zampolli (eds.). *Linguistic Structures Processing* 4. Amsterdam, New York und Oxford: North Holland, 55–88.

- Fillmore, Charles (1982). Frame semantics. In: Linguistic Society of Korea (eds.). *Linguistics in the Morning Calm*. Seoul: Hanshin, 111–137.
- Fillmore, Charles (1985). Frames and the semantics of understanding. *Quaderni di Semantica* 6, 222–254.
- Fricke, Ellen (2007). *Origo, Geste und Raum. Lokaldeixis im Deutschen*. Berlin und New York: de Gruyter Mouton.
- Fricke, Ellen (2012). *Grammatik multimodal. Wie Wörter und Gesten zusammenwirken*. Berlin und New York: de Gruyter Mouton.
- Gibbs, Raymond W. Jr. (2006). *Embodiment and Cognitive Science*. New York: Cambridge University Press.
- Gibson, James (1977). The theory of affordances. In: Robert Shaw und John Bransford (eds.). *Perceiving, Acting, and Knowing*. Hillsdale: Erlbaum, 67–82.
- Goodwin, Charles (2007). Environmentally coupled gestures. In: Susan Duncan, Justine Cassell und Elena Levy (eds.). *Gesture and the Dynamic Dimensions of Language*. Amsterdam und Philadelphia: John Benjamins, 195–212.
- Goodwin, Charles (2011). Contextures of action. In: Jürgen Streeck, Charles Goodwin und Curtis LeBaron (eds.). *Embodied Interaction: Language and Body in the Material World*. Cambridge, UK: Cambridge University Press, 182–193.
- Hassemer, Julius (2015). *Towards a theory of Gesture Form Analysis: Principles of gesture conceptualisation, with empirical support from motion-capture data*. Ph.D. thesis, RWTH Aachen University.
- Haviland, John (2000). Pointing, gesture spaces and mental maps. In: David McNeill (ed.). *Language and Gesture*. Cambridge, UK: Cambridge University Press, 13–46.
- Holler, Judith und Kati Wilkin (2011). Co-Speech gesture mimicry in the process of collaborative referring during face-to-face dialogue. *Journal of Nonverbal Behavior*.
- Hutchins, Edwin (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.
- Hutchins, Edwin (2005). Material anchors for conceptual blends. *Journal of Pragmatics* 37, 1555–1577.
- Jakobson, Roman (1956). Two aspects of language and two types of aphasic disturbances. In: Linda R. Waugh und Monique Monville-Burston (eds.). *Roman Jakobson. On Language*. Cambridge und London: Harvard University Press, 115–133.
- Jakobson, Roman (1960). Language and poetics. In: Krystyna Pomorska und Stephen Rudy (eds.). *Roman Jakobson – Language in Literature*. Cambridge und London: Harvard University Press, 62–94.
- Jakobson, Roman (1966). Quest for the essence of language. In: Linda R. Waugh und Monique Monville-Burston (eds.). *Roman Jakobson – On Language*. Cambridge und London: Harvard University Press, 407–421.
- Jakobson, Roman (1971). Shifters, verbal categories, and the Russian verb. In: *Roman Jakobson, Selected Writings, Volume II: Word and Language*. The Hague, Paris: Mouton, 130–147.
- Johnson, Mark (1987). *The Body in the Mind. The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: University of Chicago Press.
- Kendon, Adam (2004). *Gesture: Visible Action as Utterance*. Cambridge, UK: Cambridge University Press.
- Kita, Sotaro, Ingeborg van Gijn und Harry van der Hulst (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In: Ipke

- Wachsmuth und Martin Fröhlich (eds.). *Gesture and Sign language in Human-Computer Interaction*. Berlin: Springer, 23–35.
- Kita, Sotaro (2003). *Pointing: Where Language, Culture, and Cognition Meet*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Ladewig, Silva H. (2020). *Integrating Gestures. The Dimension of Multimodality in Cognitive Grammar*. Berlin und Boston: de Gruyter Mouton.
- Langacker, Ronald W. (1993). Reference-point constructions. *Cognitive Linguistics* 4, 1–38.
- Langacker, Ronald W. (2009). Metonymic grammar. In: Karl-Uwe Panther, Linda L. Thornburg und Antonio Barcelona (eds.). *Metonymy and Metaphor in Grammar*. Amsterdam: John Benjamins, 5–71.
- Levinson, Steven C. (2003). *Space in Language and Cognition*. Cambridge, UK: Cambridge University Press.
- Liddell, Scott (2003). *Grammar, Gesture and Meaning in American Sign Language*. Cambridge, UK: Cambridge University Press.
- McNeill, David (1992). *Hand and Mind. What Gestures Reveal about Thought*. Chicago: Chicago University Press.
- McNeill, David (2005). *Gesture and Thought*. Chicago: Chicago University Press.
- McNeill, David, Justine Casell, Elena Levy (1993). Abstract deixis. *Semiotica* 95 1/2, 5–19.
- Mittelberg, Irene (2006). *Metaphor and metonymy in language and gesture. Discourse evidence for multimodal models of grammar*. Ph.D.-Dissertation, Cornell University. Ann Arbor: UMI.
- Mittelberg, Irene (2008). Peircean semiotics meets conceptual metaphor. Iconic modes in gestural representations of grammar. In: Alan Cienki und Cornelia Müller (eds.), *Metaphor and Gesture*. Amsterdam und Philadelphia: John Benjamins, 115–154.
- Mittelberg, Irene (2010a). Geometric and image-schematic patterns in gesture space. In: Vyvyan Evans und Paul Chilton (eds.). *Language, Cognition, and Space: The State of the Art and New Directions*. London: Equinox, 351–385.
- Mittelberg, Irene (2010b). Interne und externe Metonymie. Jakobsonsche Kontiguitätsbeziehungen in redebegleitenden Gesten. *Sprache und Literatur* 41, 1, 112–143.
- Mittelberg, Irene (2013). The ex-bodied mind. Cognitive-semiotic principles as motivating forces in gesture. In: Cornelia Müller, Alan Cienki, Ellen Fricke, Silvia H. Ladewig, David McNeill und Sedinha Tessendorf (eds.). *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction* (38.1). Berlin und Boston: de Gruyter Mouton, 750–779.
- Mittelberg, Irene (2014). Gestures and iconicity. In: Cornelia Müller, Alan Cienki, Ellen Fricke, Silvia H. Ladewig, David McNeill und Jana Bressemer (eds.). *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction* (38.2). Berlin und Boston: de Gruyter Mouton, 1712–1732.
- Mittelberg, Irene (2017). Embodied frames and scenes: Body-based metonymy and pragmatic inferencing in gesture. *Gesture* 16, 2, 203–244.
- Mittelberg, Irene (2018). Gestures as image schemas and force gestalts: A dynamic systems approach augmented with motion-capture data analyses. *Cognitive Semiotics* 11, 1, 1–21.

- Mittelberg, Irene (2019a). Peirce's universal categories: On their potential for gesture theory and multimodal analysis. *Semiotica* 228, 193–222.
- Mittelberg, Irene (2019b). Visuo-kinetic signs are inherently metonymic: How embodied metonymy motivates form, function and schematic patterns in gesture. *Frontiers in Psychology* (Language Sciences).
- Mittelberg, Irene und Linda R. Waugh (2009). Metonymy first, metaphor second: A cognitive-semiotic approach to multimodal figures of thought in co-speech gesture. In: Charles Forceville und Eduardo Urios-Aparisi (eds.). *Multimodal Metaphor*. Berlin und New York: de Gruyter Mouton, 329–356.
- Mittelberg, Irene und Linda R. Waugh (2014). Gestures and metonymy. In: Cornelia Müller, Alan Cienki, Ellen Fricke, Silvia H. Ladewig, David McNeill und Jana Bressem (eds.). *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction* (38.2). Berlin und Boston: de Gruyter Mouton, 1747–1766.
- Müller, Cornelia (1998). *Redebegleitende Gesten: Kulturgeschichte – Theorie – Sprachvergleich*. Berlin: Spitz Verlag.
- Müller, Cornelia (2004). Forms and uses of the palm up open hand: a case of a gesture family? In: Cornelia Müller und Roland Posner (eds.). *The Semantics and Pragmatics of Everyday Gesture*. Berlin: Weidler, 233–256.
- Müller, Cornelia (2017). How recurrent gestures mean: Conventionalized contexts-of-use and embodied motivation. *Gesture* 16, 277–304.
- Panther, Klaus-Uwe und Linda L. Thornburg. (2003). *Metonymy and Pragmatic Inferencing*. Amsterdam: John Benjamins.
- Parrill, Fey und Eve E. Sweetser (2004). What we mean by meaning. Conceptual integration in gesture analysis and transcription. *Gesture* 4, 197–219.
- Peirce, Charles S. (1955). Logic as semiotic: The theory of signs. In: Justus Bucher (ed.). *Philosophical Writings of Peirce*. New York: Dover, 98–119.
- Peirce, Charles S. (1960). *Collected papers of Charles Sanders Peirce* (1931–1958). Vol. I.: *Principles of Philosophy*, Vol. II: *Elements of Logic*. Charles Hartshorne und Paul Weiss (eds.). Cambridge, MA: Harvard University Press.
- Pombo, Olga und Alexander Gerner (eds.) (2010). *Studies in Diagrammatology and Diagrammatic Praxis*. London: College Publications Rosch.
- Rekittke, Linn-Marlen, Irene Mittelberg und Dana Wolf (2015). How sound moves the body: Felt qualities of experience in gestural enactments of film sound. In: Ferré Gaëlle und Mark Tutton (eds.). *Proceedings of the 3rd Gesture and Speech in Interaction (GESPIN) Conference*, Nantes, 199–205.
- Schank, Roger C. und Robert Abelson (1977). *Scripts, Goals, Plans, and Understanding*. Hillsdale: Erlbaum.
- Schüller, Daniel und Irene Mittelberg (2016). Diagramme von Gesten. Eine zeichentheoretische Analyse digitaler Bewegungsspuren. *Zeitschrift für Semiotik* 38, 3–4, 3–34.
- Schüller, Daniel und Irene Mittelberg (2017). Motion-Capture-gestützte Gestenforschung. Zur Relevanz der Notationstheorie in den Digitalen Geisteswissenschaften. *Zeitschrift für Semiotik* 39, 1–2, 109–146.
- Stjernfelt, Frederik (2007). *Diagrammatology: An Investigation on the Borderlines of Phenomenology, Ontology and Semiotics*. Dordrecht: Springer.

- Stjernfelt, Frederik (2014). *Natural Propositions: The Actuality of Peirce's Doctrine of Dicisigns*. Boston: Docent Press.
- Streeck, Jürgen, Charles Goodwin and Curtis LeBaron (eds.) (2011). *Embodied Interaction: Language and Body in the Material World*. New York: Cambridge University Press.
- Sweetser, Eve E. (2007). Looking at space to study mental spaces. Co-speech gesture as a crucial data source in cognitive linguistics. In: Monica Gonzalez-Marquez, Irene Mittelberg, Seana Coulson und Michael Spivey (eds.). *Methods in Cognitive Linguistics*. Amsterdam und Philadelphia: John Benjamins, 201–224.
- Sweetser, Eve E. (2012). Viewpoint and perspective in language and gesture. From the ground down. In: Barbara Dancygier und Eve E. Sweetser (eds.). *Viewpoint in Language: A Multimodal Perspective*. Cambridge, UK: Cambridge University Press, 1–24.
- Sweetser, Eve und Kashmiri Stec (2016). Maintaining multiple viewpoints with gaze. In: Barbara Dancygier, Lu Wei-Lun und Arie Verhagen (eds.). *Viewpoint and the Fabric of Meaning: Form and Use of Viewpoint Tools across Languages and Modalities*. Berlin und Boston: de Gruyter Mouton, 237–257.
- Talmy, Leonard (2000). *Toward a Cognitive Semantics, Vol. 1: Concept Structuring Systems*. Cambridge, MA: MIT Press.
- Talmy, Leonard (2018). *The Targeting System of Language*. Cambridge, MA: MIT Press.
- Thiering, Martin (2015). *Spatial Semiotics and Spatial Mental Models: Figure-Ground Asymmetries in Language*. Berlin und New York: de Gruyter Mouton.
- Thiering, Martin (2018). *Kognitive Semantik und kognitive Anthropologie: Eine Einführung*. Berlin und New York: de Gruyter Mouton.
- Tomasello, Michael (1995). Joint attention as social cognition. In: Chris Moore und Philip Dunham (eds.). *Joint attention: Its Origins and Role in Development*. Hillsdale, NJ: Lawrence Erlbaum, 103–130.
- Tomasello, Michael und Malinda Carpenter (2007). Shared Intentionality. *Developmental Science* 10, 1, 121–125.
- Wilkens, David (2003). Why pointing with the index finger is not universal (in sociocultural and semiotic terms). In: Sotaro Kita (ed.). *Pointing: Where Language, Culture and Cognition Meet*. Mahwah, NJ: Lawrence Erlbaum.
- Ziem, Alexander (2008). *Frames und sprachliches Wissen. Kognitive Aspekte der semantischen Kompetenz*. Berlin und New York: de Gruyter.
- Ziem, Alexander (2014). *Frames of Understanding in Text and Discourse*. Amsterdam, Philadelphia: John Benjamins.

Univ.-Prof. Dr. Irene Mittelberg
 RWTH Aachen University
 Institut für Anglistik, Amerikanistik und Romanistik
 Kompetenzzentrum für Gebärdensprache und Gestik (SignGes)
 Theaterplatz 14
 D-52062 Aachen
 E-Mail: irene.mittelberg@ifaar.rwth-aachen.de

*Linn-Marlen Rekittke
RWTH Aachen University
Institut für Anglistik, Amerikanistik und Romanistik
Kompetenzzentrum für Gebärdensprache und Gestik (SignGes)
Theaterplatz 14
D-52062 Aachen
E-Mail: linnrekittke@gmail.com*