

Kognitiv-aktionale Strukturen und Praktiken multimodaler Bedeutungskonstitution. Frames und Diagramme im dialogischen Aushandeln von Reiserouten

Irene Mittelberg und Linn-Marlen Rekittke

Deutsches Abstract

Dieser Beitrag untersucht, mit welchen lautsprachlichen und gestischen Ressourcen Dialogpartner mentale *Frames* (Fillmore 1982) und mentale Karten beim gemeinsamen Entwickeln von Reiserouten evozieren. Auf der Basis von multimodalen deutschen Konversationsdaten gilt besondere Aufmerksamkeit den verschiedenen bezugnehmenden und interaktiven Funktionen indexikalischer Ausdrucksformen und Praktiken. Einerseits wird herausgearbeitet, wie in beiden Modalitäten indexikalische und ikonische Elemente ineinander greifen und dabei multimodale Diagramme (*icons of relations*; Peirce 1960) im Gestenraum emergieren. Ein weiterer Schwerpunkt liegt auf interaktiven Praktiken des Sich-aufeinander-Beziehens und Sich-Einigens der Gesprächspartner (Bavelas et al. 1995). Um den dynamischen kognitiv-aktionalen Prozessen und Strukturen multimodaler Bedeutungskonstitution Rechnung tragen zu können, werden in der Verkörperungstheorie wurzelnde kognitiv-semantische Ansätze mit pragmatisch ausgerichteten semiotischen Theorien (Jakobson 1971; Peirce 1960) verknüpft. Ziel ist dabei, die situierte Bedingtheit multimodaler semiotischer Handlungen und das Einbeziehen von *imaginary material anchors* (Hutchins 2005) im dialogischen Aushandeln von Reiserouten vor Augen zu führen.

Keywords: Gesten, multimodale Interaktion, Dialog, *Embodiment*, *Frames*, Semiotik, Indexikalität, Ikonizität, Diagramme, mentale Karten, mentale Räume

English abstract

In this paper, we examine how dialogue partners evoke mental frames (Fillmore 1982) and mental maps through verbal and gestural means when jointly planning an interrail trip through Europe. On the basis of multimodal German conversational data, special attention is paid to the different referential and interactive functions indexical forms and practices may assume in such exchanges. On the one hand, we examine how in both modalities indexical and iconic elements and movements interact in constituting multimodal diagrams (*icons of relations*; Peirce 1960) in gesture space. Another focus is on how interlocutors employ interactive practices of referring to one another and coming to an agreement (Bavelas et al. 1995). In order to account for the dynamic cognitive-semiotic processes and structures that give rise to multimodal meaning constitution, we combine embodied cognitive semantic approaches with pragmatist semiotic theories (Jakobson 1971; Peirce 1960). The overarching goal is to put into relief the situated conditioning of multimodal semiotic acts involving *imaginary material anchors* (Hutchins 2005) during dialogic travel planning.

Keywords: Gesture, multimodal interaction, dialogue, embodiment, frames, semiotic acts, mental maps, mental spaces, indexicality, iconicity, diagrams

1. Einleitung
- 1.1 Theoretische Ausgangspunkte: Peirce, Jakobson und kognitive Semantik
- 1.2 Multimodale Datengrundlage
2. Multimodales Embodiment: Gesten als Vermittler zwischen aktionaler, kontextueller Verhaftung und kognitiven Vorgängen
 - 2.1. „Experiential Grounding“ und situierte Abstraktionsprozesse
 - 2.2. Verankerung, Bezugnahme und Interaktion: Zur indexikalischen Bedingtheit multimodaler dialogischer Bedeutungskonstitution
 - 2.3. *Imaginary material anchors*: Geteilte Intentionalität durch kognitiv-semiotische Handlungen im Gestenraum
3. Multimodal evozierte Frames und Diagramme im interaktiven Diskursgeschehen
 - 3.1. Frames (Fillmore)
 - 3.2. Tendenzen im sprachlich-gestischen Evozieren von Frames
 - 3.2.1. Basale Frames: Körperliche Handlungen und Gegenständliches
 - 3.2.2. Komplexe, abstraktere Framestrukturen
 - 3.2.3. Frame-Evozierung bezüglich Orte und Länder in der Reiseplanung
 - 3.3. Multimodal erwirkte Gestendiagramme in dialogischen Planungsprozessen
4. Multimodale Bedeutungskonstitution: Eine qualitative Studie zum dialogischen Aushandeln von Reiserouten
 - 4.1. Studiendesign, Korpus, Annotation
 - 4.2. Qualitative Analyse multimodaler Interaktion mit Fokus auf Reiseroutendiagramme
 - 4.2.1. Indexikalische Phase: Subjektive Äußerung und intersubjektive Anerkennung erster Wunschziele
 - 4.2.2. Indexikalisch-ikonische Phase: Grobe Richtungen und erste Verbindungslien zwischen Städten und Regionen
 - 4.2.3. Diagrammatisch-ikonische Phase: Einigung auf komplett Reiseroute
 5. Diskussion: Indexikalische Bezüge orchestrieren dynamische multimodale Gesprächskontexturen
 - 5.1. Verkörperte Referenzpunkte
 - 5.2. Interaktive Indizes
 - 5.3. Intramodale und intermodale Bezüge
 6. Abschließende Bemerkungen
 7. Literaturverzeichnis

1. Einleitung

Dieser Beitrag untersucht, wie Dialogpartner mithilfe von verbalen und gestischen Ressourcen alternative Reiserouten entwerfen und sich schließlich auf eine Strecke einigen. Sein Anliegen ist es aufzuzeigen, inwiefern von den Gesprächsteilnehmern multimodal gestaltete Entwürfe verschiedener Reistestrecken von mentalen *Frames* (Fillmore 1982) und multimodal realisierten Diagrammen (Peirce 1960) unterfüttert und vorangetrieben werden. Es soll demonstriert werden, inwiefern im Diskursverlauf dynamische konzeptuelle Strukturen und Prozesse, subjektiv wie intersubjektiv, eine kurzlebige semiotische Realität im Zentrum des Konversationsgeschehens erlangen und in die laufende multimodale Bedeutungskonstitution sinn- und strukturstiftend einfließen.

Unser Interesse zielt insbesondere darauf, mit welchen sprachlichen und gestischen Mitteln mentale Karten und Frames evoziert und exemplifiziert werden. Wir nehmen hier besonders indexikalische Ausdrucksformen in den Blick, mit dem Ziel, ihre referentiellen und interaktiven Funktionen hinsichtlich der Reiseplanung und Verständigung zwischen den Konversationspartnern herauszuarbeiten. Einerseits geht es darum, wie in den beiden Modalitäten

indexikalische Elemente (z.B. Funktionswörter und deiktische Gesten) und ikonische Elemente (z.B. Inhaltswörter und gestisch evozierte Linien und Figurationen) ineinander greifen, um multimodale Diagramme (*icons of relations*; Peirce 1960) im Gestenraum entstehen zu lassen. Dabei soll durch das Visualisieren solcher virtuellen Referenzstrukturen illustriert werden, wie sie im Raum entstehen und dabei den Gesprächspartnern als kognitiv-semiotische Ressource dienen. Ein weiterer Schwerpunkt liegt auf indexikalischen interaktiven Praktiken des Sich-aufeinander-Beziehens und Sich-Einigens zwischen den Gesprächspartnern (Bavelas 1994; Bavelas et al. 1995).

Gemäß Fillmore (1982, 1985) sind Frames semantische Systeme, die in Erfahrungsmustern und interaktionalen Praktiken wurzeln. Wörter evozieren automatisch Frames. Frames bestimmen die Bedeutung von sprachlichen Ausdrücken und leiten sowohl Sprachgebrauch als auch Diskursverständhen unter Einbeziehung von enzyklopädischen Wissensbeständen und verinnerlichten Erfahrungskontexten.

By the word ‘frame’ I have in mind any system of concepts related in such a way that to understand any of them you have to understand the whole structure in which it fits; when one of the things in such a structure is introduced into a text or into a conversation, all of the others are automatically made available.

(Fillmore 1982:111)

Frames sind demnach metonymisch organisierte Strukturen; sie bestehen aus miteinander korrelierten Teilen, wobei ein Element oder ein Aspekt einen anderen Aspekt sowie das ganze Konstrukt aktivieren kann (Dancygier/Sweetser 2014). Innerhalb der größeren Framestrukturen EUROPA und INTERRAIL TOUR werden, so nehmen wir an, anhand von sprachlich-gestisch realisierten Verortungen Ziele gesetzt und Reiserouten schematisch gezeichnet bzw. grobe Richtungen flüchtig angedeutet. Es kann davon ausgegangen werden, dass diesen Diagrammen einerseits kognitive Karten (Downs/Stea 1982) im Bewusstsein der Gesprächsteilnehmer mehr oder weniger deutlich zugrunde liegen. Gleichzeitig werden durch die sprachliche und gestische Evozierung der entsprechenden Frames von bestimmten Orten und Ländern enzyklopädische Wissensbestände sowie persönliche Erinnerungen und Vorstellungen aufgerufen und auch Anknüpfungspunkte für weiterführende Überlegungen kreiert. Solche kollaborativen multimodalen Diskursprozesse des Hervorbringens und Aushandlens von Ideen und Vorstellungen sollen in diesem Beitrag anhand von deutschen Konversationsdaten exemplarisch untersucht und durch das Zusammenführen verschiedener Ansätze theoretisch untermauert werden.

1.1. Theoretische Ausgangspunkte: Peirce, Jakobson und kognitive Semantik

Von einem multimodalen und verkörperten Verständnis von Sprache, Interaktion und Kognition ausgehend, nähert sich dieser Beitrag *cross-modalen* Strategien der Bedeutungskonstitution mit einem Ansatz, der kognitiv-linguistische Ansätze (u.a. Coulson 2001; Fillmore 1982; Johnson 1987; Lakoff 1987; Sweetser 2012) und klassische semiotische Theorien von Charles Sanders Peirce (1955, 1960) und Roman Jakobson (1956, 1960, 1966, 1971) verknüpft. Diese verschiedenen

Perspektiven sind in dem Sinne kompatibel, als sie bei der Analyse von bedeutungsstiftenden Prozessen besonders die Motivation, Funktion und Struktur von Sprachformen und Sprachgebrauch berücksichtigen (e.g. Croft/Cruse 2004; Jakobson 1966, 1990). Beide Traditionen räumen verinnerlichten wahrnehmungs- und handlungsbasierten Mustern von Erfahrung und Interpretation einen hohen Stellenwert ein (vgl. Mittelberg 2008, 2010b, 2013).¹ Auch wird davon ausgegangen, dass Bedeutung einem Zeichenträger, etwa einem Wort oder einer Geste, nicht innewohnt, sondern im Moment des Sprach- und Zeichengebrauchs dynamisch im Bewusstsein entsteht und bestimmte, durch pragmatische Kräfte, kognitiv-semiotische Prinzipien und Erfahrung geleitete Assoziationen hervorruft. Dies kann beispielsweise durch metonymische und metaphorische Prozesse (Dirven & Pörings 2002; Jakobson 1956) und/oder durch das Aufspannen von mentalen Strukturen wie mentalen und Frames geleitet werden (Coulson 2001).

Indem diese Ansätze in ihrer Zusammenschau einen weiteren semiotischen Geltungshorizont aufspannen als rein sprachbasierte Theorierahmen, scheinen sie im besonderen Maße geeignet, auf Gesten und allgemein auf solche Kommunikationsformen angewandt zu werden, die Modalitäten mit einbeziehen, denen im Kontext natürlicher Konversation im Vergleich zur Lautsprache kein hochgradig konventionalisierter symbolischer Kode zugrunde liegt. Gesten sind zwar zu unterschiedlichen Graden konventionalisiert, sei es etwa bedingt durch kulturelle Praktiken, Bewegungsmuster oder konzeptuelle Strukturen und Operationen.² Da in redebegleitenden Gesten jedoch das hierarchisierte Verhältnis von semiotischen Modi, wie Ikonizität, Indexikalität und Konventionalität (oder Gewohnheit; Peirce 1960), eine größere Varianz aufweist und in jedem multimodalen Performanzakt unter Einfluss von pragmatischen Faktoren sich stets neu austariert, erscheint es sinnvoll, semiotische Perspektiven miteinzubeziehen (Mittelberg 2008, 2013a, 2014).³ Das Ziel ist dabei, der körperlich-aktionalen 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 eine Bandbreite von kommunikativen Funktion(en) haben (vgl. Kendon 2004). In Anlehnung an Peirce (1960) und Jakobson (1960) gehen wir von einer hierarchisierten Schichtung von mehreren Funktionen in einem semiotischen Akt aus. In der hier eingenommenen hierarchischen Sichtweise wird davon ausgegangen, dass die prädominante Funktion die Bedeutung einen multimodalen semiotischen Akt letztendlich bestimmt und andere Funktionen gleichzeitig in unterschiedlichem Grade mitwirken (Mittelberg 2013 zum *hierarchical view* und Mittelberg/Waugh 2009).

¹ Neuere Arbeiten in den Bildwissenschaften sind ebenfalls an der theoretischen Schnittstelle von Peirces Semiotik und der Verkörperungstheorie (*Embodiment*, Gibbs 2006; Johnson 1987; siehe Abschnitt 2.1) angesiedelt (vgl. Bredekamp 2010; Krois et al. 2007).

² Hinsichtlich sogenannter ‚Embleme‘ und anderer stark konventionalisierter Gestenformen, die bestimmte Kriterien der Wohlgeformtheit erfüllen müssen und eine kulturell definierte Form-Bedeutung-Relation aufweisen, siehe Calbris 1990; Kendon 2004; McNeill 1992; Müller/Posner 2004.

³ Siehe z. B. Ansätze, die auf dem Modell des sprachlichen Zeichens von Saussure basieren (Calbris 1990).

Gemäß des Themenschwerpunkts dieses Sonderheftes bildet die Theorie der *mentalalen Räume* (Fauconnier 1994/1985; Fauconnier/Sweetser 1996) den globalen konzeptuellen Rahmen für unsere Untersuchungen, inwiefern sprachliche und gestische Elemente mentale Räume im Sinne von *space builders* (Coulson 2001:22) eröffnen und insbesondere semantische Frames aktivieren. Wir möchten zeigen, inwiefern man annehmen kann, dass diese Frames und mentalen Karten im Bewusstsein der Gesprächspartner aufgerufen und weiter aufgespannt werden. Hier legen wir Ansätze zugrunde, die eine framebasierte Struktur von mentalen Räumen annehmen (Coulson 2001; Dancygier/Sweetser 2005, 2014). Zudem wird das Konzept der Referenzpunkte und -konstruktionen (*reference-point constructions*; Langacker 1993) für unsere Zwecke nutzbar gemacht (Mittelberg erscheint). Da es bereits eine wachsende Literatur zur Anwendung der *Mental-Spaces-Theorie* auf multimodale Diskursdaten gibt (Cienki 2008, 2013; Fricke 2007; Parrill/Sweetser 2004; Sweetser 2007, 2012), die Rolle semantischer Frames in der gestischen und multimodalen Bedeutungskonstitution bisher jedoch nur ansatzweise eruiert wurde (vgl. Ehmer 2011; Mittelberg 2013a, erscheint; Mittelberg/Waugh 2014), liegt der Fokus in diesem Beitrag auf dem multimodalen Evozieren von Framestrukturen in der Face-to-Face-Interaktion.

1.2. Multimodale Datengrundlage

Die diesen Ausführungen zugrunde liegenden deutschen Konversationsdaten stammen aus dem MuSKA (Multimodal Speech and Kinetic Action) Korpus des Natural Medial Lab (HumTec, RWTH Aachen). Sie bestehen aus mit Audio, Video und einem Bewegungserfassungssystem (*Motion-Capture-System*) aufgenommenen Sequenzen, in denen jeweils zwei Gesprächsteilnehmer zusammen eine fiktive Interrailtour durch Europa planen.

Als ein erster Eindruck von unserem Untersuchungsgegenstand dient das folgende Transkript der verbalsprachlichen Äußerung einer Sequenz aus dem Korpus. Die von Gesprächspartner GP_{R1} zunächst gemachten Vorschläge hinsichtlich einer möglichen Route lassen sich noch ohne Berücksichtigung der simultan produzierten Gesten nachvollziehen, da Ortsnamen genannt werden, die man auf der eigenen mentalen Karte verorten kann und von denen man eine gewisse Vorstellung aufgrund von bestehenden und differenzierten Frames hat (z.B. ein bestimmtes Wissen oder auch eigene Reiseerfahrungen hinsichtlich Stockholm und Prag). Hinsichtlich der anschließend formulierten ‚wieder nach hause‘ führenden Routenidee ist dies nicht der Fall. Letztere besteht aus einer Aneinanderreihung von Funktionswörter wie Präpositionen, lokalen und temporalen Adverbien:

(1) Routenplanung Stockholm-Prag-Kiew (02:51-03:20):

```

01 GPR1: Pass auf!
02      Wir fahren von Stockholm nach Prag
03      und von Prag nach Kiew
04      und von da nach unten irgendwo
05 GPL1: Ich weiß überhaupt nicht, wo Kiew liegt relativ zu Prag
06 GPR1: ich auch nicht %lacht
07 GPL1: aber gut vorgetäuscht

```

```

08 GPR1:ich denke
09     nee ich denke wir fahren dann so weiße von
10     da
11     nach da
12     rüber
13     runter
14     und dann nach da und dann wieder nach Hause

```

Im Verlauf des Artikels wird detailliert dargelegt, wie sich zumeist erst im Verbund mit den synchron produzierten Gesten sowie den vorherigen multimodalen Ausführungen jeweils lokal erzeugte Form-Bedeutung-Korrelationen ergeben. Gestische Verortungen und Bewegungen, und die daraus schrittweise entstehenden Diagramme, nehmen so teil am kollaborativen multimodalen Konzeptualisieren und Mediatisieren von einzelnen Reiseabschnitten sowie der Reiseroute als ganzes Konstrukt.

2. Multimodales Embodiment: Gesten als Vermittler zwischen aktionaler, kontextueller Verhaftung und kognitiven Vorgänge

Der hier gewählte Zugang wurzelt in der Verkörperungstheorie (*embodiment*), die von einem sich gegenseitig konstituierenden Verhältnis von menschlichen kognitiven Fähigkeiten und der Morphologie des Körpers und seiner routinemäßigen Interaktion mit der materiellen und sozialen Umwelt ausgeht. Sprache, *Mind* und Körperlichkeit werden – konträr zu kartesianischen Auffassungen – nicht getrennt voneinander, sondern als sich wechselseitig bedingend verstanden (u.a. Gibbs 1994, 2006; Johnson 1987, 2007; Lakoff 1987; Langacker 1987; Lakoff/Johnson 1999). Von den essentiellen verinnerlichten Wahrnehmungs- und Erfahrungsmustern, die gemäß des Konzepts des *embodied mind* die menschliche Kognition und Sprache maßgeblich strukturieren, sind hier mit Blick auf den multimodal kommunizierenden Körper besonders folgende form- und bedeutungsstiftende Elemente und Konzeptualisierungsverfahren zu nennen (für einen Überblick siehe Croft/Cruse 2004; Dancygier/Sweetser 2014):

- Kategorien und *Idealized Cognitive Models* (ICM; Lakoff 1987)
- Prototypen (Rosch 1977)
- konzeptuelle Bildschemata (*image schemas*; Johnson 1987)
- Kräfteverhältnisse (*force dynamics*, Talmy 1988)
- Metaphern und Metonymien (Lakoff 1993; Panther/Thornburg 2003)
- Perspektive (*viewpoint*; McNeill 1992)
- Referenzpunkte (*reference points*, Langacker 1993)
- perzeptive Symbolsysteme (*perceptive symbol systems*, Barsalou 1999)
- Frames (Fillmore 1982, 1984)
- Konstruktionen (Goldberg 1995)
- kognitive Grammatik (Langacker 1987)
- Simulation von senso-motorischen Mustern (Bergen 2007, 2012).

Obwohl nur der geringere Teil dieser Ansätze auf mündlichen Gebrauchsdaten basiert, werden sie dennoch als gebrauchsisierte (*usage-based*) Modelle verstanden (*usage events*, Langacker 1991; siehe auch Cienki 2015; Gonzalez-

Marquez et al. 2007). Stark kognitiv orientierte Ansätze zeigen oft (nicht nur) aus soziolinguistischen, anthropologischen und ethnographischen Perspektiven ein gewisses Desiderat an materieller Verhaftung, soziokultureller Einbettung und systematischer Berücksichtigung der zwischenmenschlichen Interaktion im Diskursverlauf. Diesem Desiderat steuern unter anderem Arbeiten aus der Konversationsanalyse (*Conversation Analysis (CA)*) und linguistischen Anthropologie entgegen, indem sie in detaillierten Analysen rekonstruieren, wie Faktoren der materiellen, räumlichen und soziokulturellen Bedingtheit von Sprachgebrauch in die multimodale Bedeutungskonstitution und körperliche Koordination zwischen Gesprächspartnern einfließen und dabei diskurs- und interaktionsstrukturierende Funktion haben (vgl. u.a. Andrén 2012; Auer/Pfänder 2012; Clark 1996; Deppermann 2011, 2013; Deppermann/Schmitt 2007; Enfield 2009; Goodwin 2011; Haviland 2009; Hutchins/Nomura 2011; Imo 2013; Jäger/Linz 2004; Kendon 1990, 2004; Mittelberg/Schüller im Druck; Mondada 2013; Müller 1998, Müller/Bohle 2007; Schmitt 2007; Streeck 2009, 2013; Streeck/Goodwin/LeBaron 2011). Im Folgenden werden einige Grundannahmen skizziert, die für den hier vertretenen Ansatz und unsere Erkenntnisinteresse von Belang sind.

2.1. ,Experiential Grounding' und situierte Abstraktionsprozesse

In der empirischen Gesten- und Gebärdensprachforschung bilden beobachtbare, körperliche Artikulationen und die sich daraus ergebenden Form- und Verhaltensmuster die Basis für linguistische und semiotische Untersuchungen der Face-to-Face-Kommunikation. So kommt es nicht von ungefähr, dass diese Zweige der Multimodalitätsforschung bereits zusätzliche Evidenz für die Annahme verkörperter Erfahrungs- und Handlungsmuster liefern konnten, besonders innerhalb der kognitiven Linguistik (vgl. Bressem 2012; Cienki 2013, 2015; Ladewig 2010; Mittelberg 2010a, 2008; Müller 1998, 2010; Schoonjans/Sambre/Brône/Feyaerts erscheint; Sweetser 2007; Wilcox 2015; Zima 2014; Zima/Brône 2015). Dabei ergibt sich ein interessantes Spannungsverhältnis zwischen materieller bzw. soziokultureller Verhaftung (*grounding*; Barsalou 2010) und Abstraktion, in dem Gesten eine dynamische Vermittlerfunktion einzunehmen scheinen. Einerseits sind die während der kindlichen Entwicklung durch körperliche Wahrnehmung und Erfahrung erwirkten, konzeptuellen Muster und kognitiven Operationen, wie sie in der kognitiven Linguistik von zentraler Bedeutung sind, auf einem sehr hohen Niveau von Abstraktion und Schematisität anzusiedeln. Gleichzeitig abstrahieren Handformen und Bewegungen, die genuin an die Konkretheit des menschlichen Körpers und den ihn umgebenen Kontext gebunden sind, tendenziell essentielle, prototypische und/oder im jeweiligen Ausdrucksmoment saliente, lokal diskursrelevante Eigenschaften. Gesten reflektieren Teilespekte nicht nur von Bewegungsabläufen, Handlungen und der beobachtbaren gegenständlichen Umwelt, sondern auch von Erinnerungen, neuen Ideen und kognitiven Modellen (Arnheim 1969; Fricke 2010; Mittelberg 2013a; Müller 1998; Streeck 2009).

In gestischen Andeutungen von Objekten, Handlungen und Relationen werden prinzipiell Aspekte ausgelassen, ob bedingt durch partielle Wahrnehmung,

partielles Gedächtnis oder mediale Bedingungen der Zeichenkonstitution, welche stets auf metonymischen Verfahren beruhen (Mittelberg 2006, 2010b; Mittelberg & Waugh 2009, 2014). So implizieren Gesten meistens ein Loslassen von der materiellen Welt, wenn auch nicht so sehr wie die Lautsprache, welche wiederum trotz ihrer symbolischen Natur die Dinge und Geschehnisse zum Teil auch ikonisch nachahmt, ob lautlich in einzeln Wortformen oder strukturell in syntagmatischen Gefügen (Haiman 1980; Jakobson 1966; Jakobson/Waugh 1979). Sowohl der Kontakt mit Personen, als auch das Manipulieren von Gegenständen oder andere Handlungen schwingen in Gesten oft noch mit oder werden ad hoc semiotisch erzeugt (Mittelberg/Schmitz/Groninger im Druck). Im Vergegenwärtigen und Verkörpern abstrakter Diskursinhalte kommen dazu metaphorische Prozesse ins Spiel, die nur in der Sprache, nur gestisch oder eben multimodal zum Ausdruck kommen können (vgl. Cienki 1998; Cienki/Müller 2008; Mittelberg 2006, 2008, 2013a; Müller 1998; Müller/Tag 2010; Sweetser 1998). Da metaphorische Prozesse in den hier zugrunde gelegten Diskursdaten nicht prominent sind, werden sie hier nicht vertiefend behandelt. Wie wir im Folgenden sehen werden, tritt das Verhältnis von materiellen und konzeptuellen Strukturen und kognitiv-aktionalen Prozessen hier anders zutage.

2.2. Verankerung, Bezugnahme und Interaktion: Zur indexikalischen Bedingtheit multimodaler dialogischer Bedeutungskonstitution

Indexikalität ist dem kommunizierenden Körper und somit den von ihm ausgehenden Gesten inhärent. Sie bedingt verschiedene Dimensionen und Praktiken gestischer Zeichenkonstitution sowie multimodaler Pragmatik allgemein (Mittelberg 2013, erscheint; Sweetser 2012). Neben gestisch-sprachlicher Prozesse der Bezugnahme durch prototypische Zeigegesten auf gegenständliche oder abstrakte, gedachte Referenzobjekte oder durch richtungsangebendes Zeigen, sind hier weitere, abgeschwächte Grade von Indexikalität zu berücksichtigen (Bühler 1934; Fricke 2007; Kita 2003; Mittelberg/Waugh 2014).⁴ Zu letzteren gehört, dass Sprecher durch Gesten den Kontakt mit der materiellen Kultur und Umwelt auch bewusst und aktiv herstellen können, indem sie Oberflächen, Gegenstände und Instrumente involvieren und mit ihnen interagieren. *Environmentally-coupled gestures* (Goodwin 2007) im Kontext archäologischer Ausgrabungsarbeiten, andere Expertenerklärungen am Objekt (Streeck 2009, 2013), das Platzieren von Gegenständen (Clark 2003) oder das fokussierte, kollaborative Agieren wie das Durchführen von Operationen im Team (Mondada 2011) implizieren neben dem ikonischen Nachahmen von Handlungen immer auch indexikalische Prinzipien und Zeigegesten als situative Praktiken (Goodwin 2003).

Materielle Strukturen und ihre Affordanzen (Gibson 1977) fungieren als *material anchors* (Hutchins 1995) für kognitive Operationen und diskursive, interaktive Erklärungs- und Verstehensprozesse. Zum Beispiel funktionieren

⁴ Siehe Bühlers (1934) Begriff der *Deixis am Phantasma*, siehe Fricke (2007) für einen Überblick verschiedener Erscheinungsformen der Deixis (z.B. am Vorstellungs-Wahrnehmungs- oder Zeichenraum); siehe McNeill/Cassell/Levy 1982 zum Begriff der abstrakten Deixis.

Uhren und Modelle von Uhren als materielle Anker, anhand derer Kinder durch Prozesse der konzeptuellen Integration (und *blending*; Coulson 2001; Fauconnier/Turner 2012) ein Zeitverständnis entwickeln und lernen zu sagen, wie spät es ist (Williams 2008). Zwecks Orientierung und Streckenplanung wird typischerweise mit Land- und Straßenkarten interagiert: Durch das Zeigen auf bestimmte Orte und das Entlanggehen an eingezeichneten Wegen, sei es für einen selbst, oder im Dialog mit Gesprächspartnern.

Der multimodal kommunizierende Körper lässt sich als indexikalischer Anker verstehen, von dem aus sich die raumzeitlichen Koordinaten multimodaler Diskursentfaltung im Sinne von Bühlers Konzept der *Origo* (Bühler 1982; Fricke 2007, 2012) stets von Neuem aufspannen und dabei materielle und interaktive Gesprächskontexte wie auch imaginäre Welten, d.h. mentale Räume, eröffnen. Mit Hinblick auf die dem kommunizierenden Körper, aber auch Vorstellungs- und Diskursprozessen einverlebte Indexikalität muss auch berücksichtigt werden, dass sich Sprecher in andere Personen und Kontexte hineinversetzen und dabei aus verschiedenen Perspektiven (*viewpoint*; McNeill 1992) Szenarien beschreiben und auch (nach-) erleben können.⁵ Somit sind Beschreibungen tendenziell perspektiviert. Sweetser (2012) 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)

Im Dialog interagieren mindestens zwei indexikalisch verankerte und zumeist unterschiedlich perspektivierte kommunizierende Körper. In konversationellen Interaktionsformen spielen indexikalische Elemente und Praktiken wie Funktionswörter (z.B. Personal- und Demonstrativpronomen) und gestisches Auf-Einander-Beziehen eine zentrale Rolle (vgl. *interactive discourse index* nach Mittelberg/Waugh 2014; Kendon 2004; Müller 2004). Hinsichtlich der interaktiven Funktionen von Gesten stützen wir uns insbesondere auf die Arbeiten von Bavelas und Kollegen (Bavelas 1994; Bavelas et al. 1995). Bavelas (1994: 213) unterscheidet vier Typen von *interaktiven Gesten*, die jeweils weitere Unterkategorien aufweisen. Die *Delivery Gesture* bezieht sich auf von Gesprächspartnern ausgetauschte Informationen: Sie wird unterteilt in eine *General Delivery Gesture*, welche die Sprecher-Zuhörer Beziehung markiert, indem neue Information sozusagen (metaphorisch) an den Gesprächspartner weitergereicht wird. Eine *Shared Information Gesture* bezieht sich auf Inhalte, die den Gesprächspartnern bereits bekannt sind (*common ground*; Clark 1996). *Citing*

⁵ Diese indexikalische Dimensionen ikonischer Verkörperungen (*character viewpoint*; McNeill 1992) und gestischer Darstellungen aus der Beobachterperspektive (*observer viewpoint*) sind in der Literatur für multimoden Sprachgebrauch bereits gut beschrieben, inklusive dualer Perspektive, d.h. das verschiedene Artikulatoren gleichzeitig unterschiedliche Perspektiven reflektieren (Dancygier/Sweetser 2012; Mittelberg 2013).

Gestures sind anaphorische Verweise auf eine im Gespräch vorausgegangene Äußerung, wobei zwischen einer ein Zitat kennzeichnende *General Citing Gesture* und einer *Acknowledgment Gesture* unterschieden wird. Durch letztere signalisieren Zuhörer, dass sie die Äußerung ihres Gesprächspartners vernommen bzw. verstanden haben. Die *Seeking Gesture* evoziert eine Antwort, Hilfe oder Zustimmung. Schließlich regeln *Turn Gestures* den Sprecherwechsel: Mit einer *Giving Turn Gesture* vergibt der Sprecher die Sprecherrolle an einen Gesprächspartner, und mit einer *Taking Turn Gesture* wird die Sprecherrolle übernommen (siehe Abschnitte 3 und 4 für Beispiele).

2.3. *Imaginary material anchors*: Geteilte Intentionalität durch kognitiv-semiotische Handlungen im Gestenraum

In unseren Daten besteht die kognitive Leistung im kollaborativen Erstellen von fiktiven Reiserouten im geteilten Gestenraum. Da die Gesprächspartner in den hier zugrunde gelegten Daten gemeinsam ihnen vorgegebene kognitive und kommunikative Aufgaben bewältigen müssen, benötigen wir für die Analyse Konzepte und Methoden, die diesen Phänomenen gerecht werden können. Es geht nicht um im Kontext vorhandene Dinge, Personen oder Ereignisse, die koordiniert und multimodal kontextualisiert werden, sondern 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, joint intentionality*; Tomasello 1995; Tomasello/Carpenter 2007). Dies geschieht nicht zuletzt durch routinemäßige semiotische Handlungen wie das händische Schreiben und Zeichnen mithilfe von Artefakten (Mittelberg/Schmitz/Groninger erscheint; Mittelberg erscheint).

Hier können wir uns auf Hutchins' (2005) Konzept der ‚imaginary material anchors‘ stützen, d. h. solche Anhaltspunkte und Elemente, die nicht als gegenständliche Strukturen im Kontext vorhanden sind, sondern die man sich anhand von Wissen und verinnerlichter Erfahrung (mehr oder weniger genau) vorstellen kann. In unserem Falle geht es um imaginäre Kartenstrukturen, und das Manipulieren besteht hier zum Beispiel im Zeigen auf in der Vorstellung befindliche Orte, sowie im simulierten Entlangstreichen an Strecken mit dem Zeigefinger.

In some cases it is possible to do the cognitive work by imagining the manipulation of a physical structure. (...) The cultural process of crystallizing conceptual models in material structure and saving those up through time puts modern humans in a world where thinking depends in significant measure on the availability of a set of physical structures that can be manipulated in this way. A final turn on this path is that when a material structure becomes very familiar, it maybe possible to imagine the material structure when it is not present in the environment.

(Hutchins 2005:1575)

Aus den vorausgehenden Erörterungen hat sich die Vermittlerfunktion von gestischen Artikulationen und Praktiken zwischen dem Kognitiven, Abstrakten und Imaginären einerseits und dem Körperlichen, Materiellen, objektfokussierten,

interaktiven und semiotischen Handeln andererseits herauskristallisiert. Diese unterschiedlich gelagerten Faktoren können in variabler Weise und Gewichtung die gestische Zeichenkonstitution motivieren und strukturieren, sowie die Interpretation von multimodalen semiotischen Akten leiten (siehe Mittelberg 2013a zum Begriff des *exbodied mind*; siehe auch Ladewig 2010; Müller 2010; Streeck 2009, 2013).⁶ Wie wir im nächsten Abschnitt sehen werden, geben uns die kognitive Semantik und Semiotik Instrumente an die Hand, die uns erlauben, Vorstellungswelten in die Analysen mit einzubeziehen und so aufzuzeigen, inwiefern Dialogpartner imaginäre und irreale Welten, sprich verschiedene *mental spaces*, im Gesprächsverlauf erschaffen bzw. sich zwischen ihnen hin und her bewegen.⁷

3. Multimodal evozierte Frames und Diagramme im interaktiven Diskursgeschehen

Die Teilhabe von Gesten an kognitiven Operationen wie Blending und konzeptueller Integration hat in der Gestenforschung in den letzten zehn Jahren wachsende Beachtung gefunden (vgl. Cienki 2008; 2013; Dancygier/Sweetser 2012; Ehmer 2011; Liddell 2003; Parrill/Sweetser 2004; Sweetser 2007, 2012). Aufgrund ihrer dynamisch-räumlichen Medialität sind gestische Konzeptualisierungs- und Darstellungsprozesse besonders dazu geeignet, die dynamische Natur dieser Vorgänge aufzuzeigen (Fauconnier/Turner 2002). Wie oben bereits dargelegt, bildet die *Mental-Spaces*-Theorie den allgemeinen Rahmen für unseren Fokus auf das multimodale Evozieren und Exemplifizieren mentaler Framestrukturen im Erstellen von Diagrammen. Zunächst wenden wir uns den theoretischen Prämissen mentaler Frames und Praktiken ihrer sprachlich-gestischen Evokation zu. Darauf folgend wird das peircesche Diagrammverständnis vor dem Hintergrund multimodaler Kommunikation umrissen und für unsere Zwecke nutzbar gemacht.

3.1. Frames (Fillmore)

Frames, eingeführt von Fillmore (1977, 1982, 1985), nehmen einen zentralen Platz in verschiedenen Gebieten der kognitiven Linguistik ein. Im engeren Sinn beziehen sie sich auf lexikalische Strukturen, doch wie zahlreiche Arbeiten zeigen, unterfüttern Frames auch *mental spaces*, Konstruktionen, Metaphern, Metonymien und Prozesse des Diskursverständens, um nur einige Dimensionen von sprachlichen Strukturen und Bedeutungskonstruktion zu nennen (Coulson 2001; Croft & Cruse 2004; Dancygier & Sweetser 2014; Ziem 2014).

⁶ Streeck (2009: 151) unterscheidet zwischen: *depicting* (e.g. eine ikonische Geste stellt einen konkreten Gegenstand dar) und *ceiving* (i.e. eine Geste konzeptualisiert einen thematischen Gegenstand). “When ‘they think with their hands’, speakers rely on their bodies to provide conceptual structure” (S. 152).

⁷ Siehe Deppermann (2011) hinsichtlich der Einbeziehung kognitiver Prozesse in die Konversationsanalyse.

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)

Zwecks einer ersten Vergegenwärtigung von Framestrukturen beginnen wir mit einem oft zitierten Beispiel, dem WEEK Frame (Fillmore, 1982). Nehmen wir an, ein Freund schreibt Folgendes per Email: „Habt Ihr Lust und Zeit, Freitag Abend mit uns im thailändischen Restaurant auf der Theaterstraße zu essen? Wir treffen uns um 20 Uhr“. Die zeitliche Information bezüglich des Tages „Freitag“ verstehen wir mit der ganzen Woche im Sinn, d.h mit Bezug auf den WEEK Frame und dem entsprechenden kulturellen Hintergrundwissen bezüglich einer fünftägigen Arbeitswoche und einem zweitägigen Wochende. „20 Uhr“ bekommt seine Bedeutung mit Bezug auf den TAG Frame, der bekanntlich aus 24 Stunden besteht. Dazu kommt die Ortsangabe „Theaterstraße“, die den urbanen Kontext und geographischen Gegebenheiten einer bestimmten Stadt und die entsprechenden mentalen Karten und Erinnerungen partiell aufruft. Wenn wir uns jetzt eine gestische Darstellung des Konzepts ‘Freitag’ versuchen vorzustellen, wird schnell deutlich, dass angesichts seiner abstrakten Natur das gestische Charakterisieren nur sinnstiftend sein kann, wenn wir innerhalb des größeren WOCHEN Frames agieren, in welchem jeder Wochentag eine klar definierte Stelle hat. Eine Sprecherin könnte also zum Beispiel eine Art virtuelles Schema im Gestenraum auslegen, indem sie mehrere Punkte nebeneinander vor sich in den Raum setzt, von denen jeder einen bestimmten Tag in der Woche darstellt. Dabei könnte sie den Punkt, der Freitag repräsentiert, sprachlich entsprechend benennen (vgl. Calbris 2011 zu gestischen Darstellungen von Zeitsträngen und Mittelberg 2006, 2008 zu gestischen Indikationen der Position von sprachlichen Elementen in einem SATZ Frame).

Das oben genannte Beispiel involviert einen weiteren, in der Literatur prominent diskutierten Frame (oder *Script* gemäß Schank/Abelson 1977).⁸ Der RESTAURANT DINING Frame strukturiert komplexere Szenarien mit kulturell bedingten Elementen, Handlungen und Phasen. In der gestischen Modalität können einzelne routinierte Handlungen nachgeahmt werden, die ein Abendessen in einem Restaurant typischerweise ausmachen oder die man selbst so ausgeführt oder beobachtet hat, zum Beispiel: Die Menükarte in der Hand zu halten, die Aufmerksamkeit des Servicepersonals zu suchen, das vor einem platzierte Gericht zu inspizieren, mit Messer und Gabel oder mit Stäbchen zu essen, die Geldbörse aus der Tasche zu holen, Geld auf den Tisch zu legen etc. Multimodale Beschreibungen solcher Abläufe können auch schematisch die Reihenfolge nacheinander stattfindender Geschehnisse durch eine listenförmige Reihe von in den Gestenraum oder auf eine Tischoberfläche gesetzter Beatgesten darlegen, von oben nach unten oder von der einen zur anderen Seite des Sprechers etc. (Calbris 2008, 2011).

⁸ Hinsichtlich Frames, Skripts, Szenarios oder Idealized Cognitive Models (Lakoff 1987) siehe auch Coulson 2001; Croft/Cruse 2004; Dancygier/Sweetser 2014; Ziem 2014.

3.2. Tendenzen im sprachlich-gestischen Evozieren von Frames

Bisher haben sich unsere Erörterungen auf Beispiele bezogen, die in der Literatur zur Frame-Semantik einen klassischen Status erlangt haben (Coulson 2001; Croft & Cruse 2004; Dancygier & Sweetser 2014; Ziem 2014). In diesem Abschnitt wird nun anhand eigener Daten dargelegt, inwiefern Frames allgemein und insbesondere im Kontext der hier relevanten Reiseplanungen durch sprachliche Elemente und gestische Handlungen evoziert werden und dabei Zugang zu enzyklopädischen Wissenskomplexen sowie Erfahrungen und Assoziationen gewähren können. Dabei liegt ein Hauptaugenmerk auf den motivierenden und pragmatischen Faktoren, die für framebasierte gestische Kommunikation von besonderem Belang zu sein scheinen (Mittelberg 2013a, erscheint; Mittelberg/Waugh 2014). Dabei ist wichtig, dass Frames als abstrakte konzeptuelle Strukturen zum gewissen Grad erfahrungsbedingt sind, d.h. „motivated by human experiences, social institutions, and cultural practices“ (Coulson 2001:18). Für Fillmore (1977, 1985) zählen außersprachliche, kontextuelle und ko-textuelle Faktoren zu den Kräften, die unser Verständnis von Begriffen und ihre Vernetzung konditionieren (Ziem 2014:24 u. 188 ff.). Wie wir zeigen möchten, reflektieren Gesten framebasierte Facetten, Elemente und Relationen. In den folgenden Ausführungen legen wir Mittelbergs (erscheint) Arbeiten zu verkörperten Frames, Metonymie und pragmatischer Inferenz zugrunde, und dabei insbesondere die Unterscheidung unterschiedlich komplexer, interagierender Framestrukturen, welche, so auch die hier vertretene Annahme, mit bestimmten Tendenzen gestischer Praktiken in der multimodalen Kommunikation korrelieren.

3.2.1. Basale Frames: Körperliche Handlungen und Gegenständliches

Basic physical action and object frames (Mittelberg im Druck) können durch nachahmende kommunikative Handlungen evoziert werden, die auf körperlichen Bewegungs- und Handlungsroutinen, wie zum Beispiel laufen oder schwimmen, oder der Manipulation von Objekten und Instrumenten beruhen, z. B. auf einer Klaviatur zu tippen, aus einer Tasse zu trinken oder eine Frucht mit dem Messer in Stücke zu schneiden.⁹ Gegenstandsbeschreibungen zählen auch hierzu. Sie sind erfahrungsbedingt in der direkten Interaktion mit der materiellen und sozialen Welt verankert. Ein Beispiel aus dem hier zugrunde liegenden MuSKA Korpus ist eine körperlich-gestische metonymische Andeutung von Schwimmbewegungen, die zusammen mit dem synchron produzierten verbalen Ausdruck „Schwimm-Wetter“ den entsprechenden Frame aufruft und illustriert bzw. inszeniert: Die Sprecherin auf der rechten Seite in Abb. 1 ahmt die für das Brustschwimmen typische Arm- und Handhaltung und die entsprechenden Bewegungen ikonisch-aktional nach.

⁹ Siehe auch Calbris (2003) bezüglich abstrakter Schneidehandlungen und Grandhi/Joue/Mittelberg (2011) für eine Anwendung dieser basalen transitiven Handlungsroutinen im Bereich gestenbasierter interaktiver Systeme.



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

Ein Blick auf den semiotischen Kontext dieser bimodalen Äußerung zeigt, dass ihr die größere Framestruktur WARMES WETTER zugrunde liegt und sie in eine längere Sequenz von einzelnen multimodalen Darstellungseinheiten eingebettet ist. Die Sprecherin beschreibt relativ ausführlich, was für sie Sommerwetter ausmacht und inwiefern diese Aspekte während eines zurückliegenden Aufenthalts in Budapest gegeben waren. Sie tut dies als Reaktion auf die Frage ihres Gesprächspartners GP_{L2}, wann die Budapest-Reise stattfand, und ob zum Beispiel Schnee lag.

(2) WARMES-WETTER Frame (01:28-01:39)

01GP_{L2}: [Lag Schnee?] ^{G1}
 02GP_{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}
 06GP_{L2}: %lacht [okay] ^{G6}

Seine Frage, ob Schnee lag, begleitet GP_{L1} mit einer beidhändig ausgeführten „Palm-up Open Hand“ Geste (PUOH, Müller 1998): mit zwei nach oben gedrehten offenen Handflächen (Abb. 2), die hier eine vorwiegend pragmatische (und nicht referentielle) Funktion haben. GP_{L2} erbittet, sozusagen mit leeren Händen, eine Antwort. Dies ist somit ein Beispiel für eine interaktive *seeking gesture* (Bavelas 1994; Bavelas et al. 1994 (Abschnitt 2.2); siehe auch Kendon 2004; Müller 2004). GP_{R1} kann sich zwar nicht an das genaue Datum erinnern, wohl aber an die sommerlichen Wetterverhältnisse. Zunächst eröffnet sie den WARM-WETTER Frame nicht nur verbal durch „war schon warm“, sondern dazu mit einer PUOH



Abb. 2: Seeking Geste (G1)

Geste (siehe Abb. 3b), die hier auch eine vorwiegend pragmatische, in diesem Falle modale, Funktion (Kendon 2004, Müller 2004) hat, indem sie unterstreicht, dass es für die Sprecherin zweifelsohne „warm“ war. Dann führt sie nacheinander Elemente und Handlungen auf, die für sie den aktivierten Frame konstituieren: „T-Shirt-Wetter“ begleitet sie durch beidhändige zirkulierende Handbewegungen

dicht an ihrem mit einem T-Shirt bekleideten Oberkörper (Abb. 3c), gefolgt von einem indexikalischen Verweis auf ihre Oberarme. Zeitgleich mit „Shorts-Wetter“ legt sie die Hände auf ihre Oberschenkel und markiert dabei, wo Hosenbeine von Shorts enden (Abb. 3c). Darauf folgt die eben beschriebene und in Abb. 1 bereits gezeigte Schwimmbewegung auf „Schwimm-Wetter“ (Abb. 3e). Ihr Gesprächspartner ist sichtlich amüsiert und ahmt sie mit einer ähnlichen Schwimmbewegung nach (Abb. 3f), welches als eine Instanz von gestischer Mimikry und *common ground* (Clark 1996; Holler/Wilkin 2011; McNeill 2005) zu sehen ist (siehe auch *body action icon* gemäß Mittelberg/Waugh 2014).

Es lässt sich festhalten, dass GP_{L2} die Beschreibung von GP_{R2} auf multimodale Weise rahmt: durch die eingangs gestellte Frage und zeitgleiche PUOH Geste einerseits und durch die abschließende ikonische Geste, durch welche er die Urlaubserlebnisse von GP_{R2} ein Stück nacherlebt, andererseits. Dabei wird der *Turn* von GP_{R2} strukturiert durch den größeren WARM-WETTER Frame, der auf einer Ebene stärkerer Schematizität und Abstraktion angesiedelt ist als die hintereinander illustrierten Teilespekte wie Schwimmen oder Shorts. Zusätzlich ist zu bedenken, dass Frames neben konventionellen auch kulturell und subjektiv geprägte Facetten aufweisen. Für Personen aus anderen Kulturreihen kann sich die Konzeption von warmem Wetter bedingt durch körperliche Erfahrung und sozio-kulturellen Normen anders gestalten und sich dementsprechend auch anders auf multimodale Darstellungen dieser Art auswirken. Die unten stehende Bildfolge zeigt Momentaufnahmen der eben beschriebenen multimodalen Darstellungen (siehe Transkript 2):



3a: Seeking Geste (G1)



3b: Modale Geste signalisiert Offensichtlichkeit (G2)



3c: Gestischer Bezug auf T-Shirt (G3)



3d: Illustration der Länge einer Shorts (G4)



3e: Ikonisch-aktionale Schwimm-Geste (G5)

Abb. 3a-f: Sequenz zum WARMES-WETTER Frame



3f: Nachahmung der Schwimm-Geste (G6)

In dieser Sequenz fällt weiterhin auf, dass nur die beiden Schwimm-Gesten (3e & 3f) prädominant ikonisch sind und ein basales, leicht erkennbares Bewegungsmuster pantomimisch ausüben bzw. einen Teil der Bewegung metonymisch exemplifizieren (Goodman 1968; siehe auch den Modus *action image icon* gemäß Mittelberg/Waugh 2014). Die anderen Gesten, welche Attribute von Kleidungsstücken metonymisch illustrieren, zeigen stärkere indexikalische Anteile, beispielsweise indem sie T-Shirt und Shorts am eigenen Körper dort verorten, wo sie typischerweise (und hier gerade von der Sprecherin) getragen werden (siehe *body part index* gemäß Mittelberg/Waugh). Somit fungiert der Körper der Sprecherin GP_{R1} hier als *material anchor* in Hutchins' (1995) Sinn (siehe Abschnitt 2; Sweetser 2012). Wir wenden uns nun weniger gegenständlichen Strukturen zu, den bereits eingeführten *imaginary material anchors* (Hutchins 2005) und damit verbundenen, anders gelagerten semiotischen Praktiken, die angesichts des dialogischen Aushandelns von Reiserouten eine wichtige Rolle spielen.

3.2.2. Komplexe, abstraktere Frame-Strukturen

Prozesse multimodaler Bedeutungskonstitution rekrutieren auch *complex, more abstract frame structures* (Mittelberg erscheint). Diese umfassen in erster Linie schematische, komplexere Frame-Strukturen, von denen angenommen wird, dass sie das konzeptuelle System, abstrakte Wissensbereiche, sprachliche Form, aber auch Turns und ganze Diskurse strukturieren (Coulson 2001; Ziem 2008, 2014). Unter anderem funktionieren sie als Ordnungsprinzipien von miteinander interagierenden verkörperten kognitiven Strukturen und Prozessen wie Bildschemata, Kräftedynamiken und mentale Räume (Dancygier & Sweetser, 2014). Dabei ist ein zentraler Aspekt, dass sich solche größeren schematischen Architekturen in der gestischen Modalität manifestieren können, jedoch nicht müssen. Als ein erstes Beispiel lässt sich der eben diskutierte WARM-WETTER-Frame anfügen, der die Sequenz multimodaler Einzeldarstellungen von Frameelementen von GP_{R1} unterfüttert. Einerseits sind die einzelnen Teile in der Erfahrung korreliert, andererseits fungiert er bereits auf einer schematischeren Ebene als die basalen Frames. Dass, oder wie die Teilespektre zusammenhängen, wurde selbst nicht thematisiert oder gestisch dargestellt, sondern ergibt sich aus der sequenziellen Abfolge (oder temporalen Kontiguität, Jakobson 1990) einerseits und aus erfahrungsbedingten Korrelationen von Elementen (z.B. wenn es heiss ist, trägt man T-Shirts, keine Pullover).

Komplexere und hochgradig abstrakte Framestrukturen operieren im Hintergrund von Kognition und Kommunikation (Coulson 2001). Sie können in

der gestischen Modalität eine kurzweilige semiotische Realität erlangen und so als dynamische, emergierende kognitiv-semiotische Strukturen verstanden werden, anhand, oder entlang derer sich Gesprächspartner in ihrer Vorstellung, im Nachdenken, Sprechen und Zuhören von einem Punkt zum anderen bewegen (Coulson 2001; Dancygier & Sweetser 2014; Fillmore 1977; Ziem 2014). Narrative Strukturen und genrespezifische Muster, die Diskurse, Texte oder Filme strukturieren und dabei Erwartungen seitens der Gesprächspartner kreieren, gehören zu diesen größeren Frames, welche reichhaltige Gewebe an bildlichen Vorstellungen, Teilelementen, Rollen etc. repräsentieren (Labov 1972; Ziem 2014). Zum Beispiel beruhen Episoden der Zeichentrickfilme *Canary Row* (welche extensiv als Stimuli für Gestenproduktion eingesetzt wurden, vgl. McNeill 1992, 2005) auf Skripts und Frames, welche die multimodalen Nacherzählungen der an Bewegungsereignissen (u.a. einzelnen basalen Handlungsframes) reichen, aufeinanderfolgenden Episoden strukturieren und vorantreiben. Gleichzeitig kann die (diagrammatische) Struktur von Diskursen, Konversationen und Turn Sequenzen auch gestisch in den Vordergrund gebracht werden, zum Beispiel durch den Einsatz rhythmischer *Beats*, *Cohesives* (McNeill 1992) oder anderer diskursstrukturierender Gesten (Müller 1998; Kendon 2004; Streeck 2009). Diagrammatische, auf komplexen Frames basierende Praktiken werden im folgenden Abschnitt detaillierter erläutert.

3.2.3. Frame-Evozierung bezüglich Orte und Länder in der Reiseplanung

Es ist generell anzunehmen, dass sich manche Frameelemente, Relationen zwischen Elementen, Dimensionen und Assoziationen intuitiver und unmittelbarer gestisch darstellen lassen als andere. Körperliche Handlungen oder Gegenstandsbeschreibungen, wie sie zu der ersten Gruppe von framebasierter gestischer Kommunikation zählen (siehe Abschnitt 3.2.1), lassen sich aufgrund der direkten erfahrungsbedingten Motiviertheit unmittelbarer, also ohne semiotische Umwege, darstellen als Orte oder Regionen. Da es in den dialogischen Reiseplanungen besonders um Städte- und Länderframes geht, sollen an dieser Stelle einige Beispiele der Evozierung solcher Frames gegeben werden. Im Kontext unserer Studie sind übergeordnete Frames zum Beispiel der REISE-Frame, der INTERRAIL-Frame und der EUROPA-Frame, welche beide einerseits aus erfahrungsbedingten Elementen bestehen und andererseits eine größere Komplexität und Vieldimensionalität aufweisen: Städte, Länder, Fortbewegung, Transportmittel, Übernachtungsmöglichkeiten, Essen, kulturelle und sportliche Aktivitäten, Wetter, etc. Dabei kommen allgemeine enzyklopädische Wissensbereiche und individuelle Gewohnheiten und Präferenzen zum Tragen (z.B. geprägt durch Familien- und Gruppenurlaube).

Als eine den detaillierten Sequenzanalysen vorausgehende Beobachtung verwundert es zunächst kaum, dass die Frames STOCKHOLM und PRAG (siehe Einleitung, Beispiel 1) nicht ikonisch oder durch gewisse Symbole metonymisch repräsentiert, sondern durch in den Raum gesetzte, verortende Zeigegesten und die synchron geäußerten Städtenamen aufgerufen und zueinander in Beziehung gesetzt werden. So setzt die in dieser Dialogkonstellation im Bild rechts situierte

Gesprächspartnerin (GP_{R2} ; Abb. 4) zunächst mit ausgestrecktem Zeigefinger von ihr aus gesehen links auf Schulterhöhe einen Punkt, der „Stockholm“ repräsentiert, in den Raum. Von dort aus zieht sie dann, ohne die Handfiguration zu ändern, eine leicht bogenförmige Linie runter bis in die Mitte ihres Gestenraums und verortet dort einen weiteren, verbal mit „Prag“ bezeichneten, Punkt.

(3) Von Stockholm nach Prag (03:20-03:23):

01 GP_{R2} : Pass auf... wir fahren [von Stockholm^{G7} nach Prag]^{G8}



4a: Verortender Index: „Stockholm“ (G7)

Abb. 4a-b: Verortung und In-Bezug-Setzen zweier Reiseziele



4b: Verortender Index: „Prag“ (G8)

Im Vergleich zu den eben besprochenen, präzise artikulierten, indexikalischen Gesten sind auch eher grobe Richtungsanweisungen im Raum und damit innerhalb des EUROPA-Frames zu beobachten, beispielsweise mit „gen Osten“ (in Richtung Kiew) oder „nach Süden“ (in Richtung Mittelmeer). Solche vagen Richtungsangaben wurden lautsprachlich mehrfach mit ungenauen adverbialen Bestimmungen wie „irgendwo“ oder „irgendwie“ begleitet. Ob dies auf undifferenziertes geographisches Wissen oder eine noch nicht weitgediehene Reiseplanung, d.h. noch unklare Vorstellungen, zurückzuführen ist, muss, wie im Abschnitt 4 dargelegt wird, stets anhand des Kontextes bestimmt werden (siehe auch Hassemer 2015; Wilkens 2003).

Eine weitere multimodale Praktik involviert den gestischen Ausdruck eines metaphorischen Raumverständnisses. Nachdem GP_{R1} angibt, dass sie nicht weiß, wo Vilnius liegt, erklärt GP_{L1} : „Vilnius ist die Hauptstadt von Litauen“. Er zieht dabei beide Hände seitlich hoch, wobei sich die Handinnenflächen im bedeutungstragenden Moment der Geste gegenüber stehenden. Durch diese Geste, welche die konzeptuelle Metapher COUNTRIES ARE CONTAINERS (CONTAINMENT; Lakoff/Johnson 1980) zu reflektieren scheint, konzeptualisiert er Litauen gestisch als einen dreidimensionalen geographischen Raum, d.h. einen Behälter, in dem sich Vilnius befindet.

(4) Vilnius ist Hauptstadt von Litauen (00:35-00:36):

01 GP_{L2} : ja [Vilnius ist die **Hauptstadt** von Litauen]^{G9}



Abb. 5a: Metaphorisches Raumverständnis von Litauen (G9)
Abb. 5a-b: CONTAINER Geste (G9)

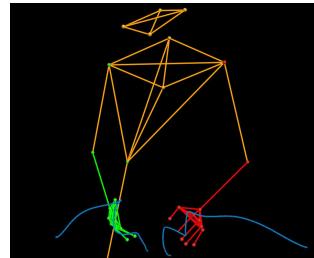


Abb. 5b: Motion-Capture Visualisierung (G9)

Welche Art von Frames besonders häufig gestisch instanziert und welche Aspekte dabei metonymisch profiliert werden (Mittelberg/Waugh 2014), muss durch größer angelegte empirische und sprachvergleichende Studien erst noch genauer untersucht werden.

3.3. Multimodal erwirkte Gestendiagramme in dialogischen Planungsprozessen

Diagramme gelten allgemein als ikonische Zeichenprozesse, *icons of relations* (Peirce 1955, 1960), die Beziehungen zwischen Elementen darstellen. In Peirces eigenen Worten: Ikone, “which represent the relations, mainly dyadic, (...) of the parts of one thing by analogous relations in their own parts, are diagrams” (Peirce 1960: 157; 2.277; siehe auch Pombo & Gerner 2010; Stjernfelt 2007). Straßen- und U-Bahnkarten sowie Familienstammbäume sind bekannte Beispiele. Hinsichtlich Gestendiagramme ist nun entscheidend, dass schematische Figurationen dieser Art oft durch den Einsatz unterschiedlich gestalteter indexikalischer Handformen und Bewegungen emergieren. Sprecher setzen zum Beispiel mit Zeigegesten Punkte in den Gestenraum, um sich 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 (Bressem 2012; Enfield 2003, 2011; Fricke 2007, 2012; Kendon 2009; Mittelberg 2006, 2008, 2010a, 2014; Roth, 2003; Tversky 2011). Dabei ist von zentraler Bedeutung, das 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. In unserem Korpus machen Gesprächspartner 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 den jeweiligen Sprecher einzelne Orte, Länder oder Aktivitäten räumlich und/oder zeitlich zueinander verhalten.

Wenden wir uns an dieser Stelle noch einmal Beispiel 3 (Abb. 3a/b) zu: „von Stockholm nach Prag“. Durch die Bewegung mit dem ausgestreckten Indexfinger

zeichnet die Sprecherin eine Verbindungslinie zwischen den beiden Punkten; zusammen stellen diese Punkte und Linie bereits ein einfaches Diagramm dar, dessen Verlauf in Abbildung 6 anhand von Motion-Capture-Daten visualisiert wird. Dieser Darstellung unterliegt das in der kognitiven Linguistik prominente Weg-Schema (Johnson 1987; Lakoff/Johnson 1999) mit einem Ausgangspunkt (SOURCE; Stockholm), einer Wegstrecke oder einem Vektor (PATH, nach) und einem Zielpunkt (GOAL; Prag). Unter Rückgriff auf numerische Motion-Capture-Daten sind wir in der Lage, die Positionierung von Punkten im drei-dimensionalen Raum und das allmähliche Entstehen von ansonsten unsichtbaren Verbindungslinien und größeren Diagrammstrukturen im von den Gesprächspartners geteilten Gesterraum Millimeter genau sichtbar zu machen.¹⁰

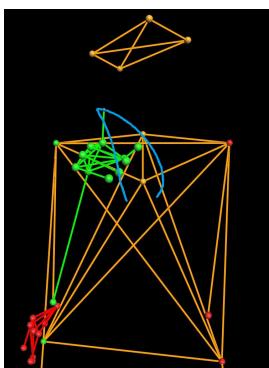


Abb. 6: Teil-Diagramm: „von Stockholm nach Prag“ (G7-8)

Wenn wir diesen Streckenentwurf weiter verfolgen, ergibt sich ein komplexeres Diagramm mit zusätzlichen Orten und Verbindungslinien bzw. Richtungen. Wie die nächste Veranschaulichung (Abb. 7) zeigt, ist diese diagrammatische Darstellung von Reiserouten im Gesterraum hochgradig subjektiv, relativ und approximativ. In der Einleitung haben wir anhand von Transkript 1 verdeutlicht, inwiefern die zahlreichen indexikalischen Funktionswörter (*closed-class items* nach Talmy 2000 und *shifters* nach Jakobson 1971) in der Sequenz ohne gestische Unterstützung nicht disambiguierbar werden können. Gleichzeitig wäre es schwierig, die Stationen und Wege dieses doch sehr skizzenartigen Diagramms ohne den vorhergehenden Diskurs und die synchrone sprachliche Nennung der Orte und Deiktika nachzuvollziehen („von Stockholm nach Prag und von Prag nach Kiew und von da nach unten irgendwo“). Auch hier kontextualisieren sich gestische und sprachliche Ressourcen in der dynamischen Bedeutungskonstitution gegenseitig (Jakobson 1956; Mittelberg/Waugh 2009).

Mit Rückblick auf die theoretischen Prämissen in Abschnitt 2 sind diese kognitiv-aktionalen Praktiken, angesichts der schwächeren oder nicht vorhandenen direkten erfahrungsbasierten Verhaftung (*grounding*; Barsalou 2010), körperliche und pragmatische Verankerungsmechanismen. Die Sprecherinnen imitieren dabei semiotische Handlungen wie das Zeichnen und Schreiben auf Papier oder Tafeln gestisch; sie tun so, als ob sie etwas schrieben oder aufmalten. Solche kognitiv-aktionalen Praktiken setzen ein gewisses Verständnis der zugrundeliegenden gegenständlichen oder abstrakten Strukturen

¹⁰ Dank an Bela Brenger und Ugur Töman für die Erstellung der MoCap-Visualisierungen.

voraus, in unserem Falle der Geographie Europas und bestimmter Länder (siehe *imaginary material anchors* nach Hutchins, Abschnitt 2.3).

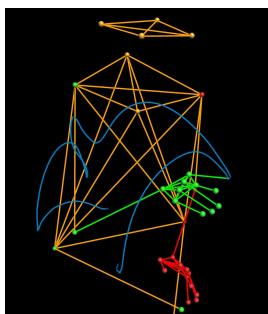


Abb. 7: Diagramm: Stockholm, Prag und Kiew

Diagramme scheinen hier die Überstruktur für punktuelle Frameaktivierung zu bilden. Sie fungieren als konzeptuell-semiotisches Gerüst für die einzelnen Punkte, an denen einzelne Frames aktiviert und somit semantische Netzwerke aufgerufen werden. Wir nehmen an, dass Sprecher mentaler Karten ausschnittartig im Kopf haben und in den Gestenraum projizieren; sie hangeln sich sozusagen gestisch und kognitiv-aktional von Ort zu Ort und Land zu Land.

4. Multimodale Bedeutungskonstitution: Eine qualitative Studie zum dialogischen Aushandeln von Reiserouten

Nachdem die vorhergehenden Abschnitte bereits dekontextualisierte Einblicke in das multimodale Evozieren von Frames und Diagrammen gegeben haben, ist dieser Abschnitt der detaillierten qualitativen Analyse eines Dialogs gewidmet.

4.1. Studiendesign, Korpus, Annotation

Die der vorliegenden Studie zugrunde liegenden Diskursdaten stammen aus dem MuSKA Korpus (Natural Media Lab; Brenger/Mittelberg 2015). Insgesamt wurden sechs Stunden synchronisierte Audio-, Video- und Motion-Capture Daten aufgenommen. Der Korpus ist unterteilt in zwei Gesprächsgenres (Dialog und Narration) sowie in zwei Sprachen (Deutsch und Amerikanisches Englisch). Für die vorliegende Studie wurden ausschließlich die deutschen Dialogdaten analysiert. Innerhalb dieses Sub-Korpus nahmen sechzehn Personen (acht Paare) an der Studie teil, fünf der Paare kannten sich bereits. Die Studienteilnehmer sind zwischen 20 und 35 Jahre alt. Im Folgenden präsentieren wir Ergebnisse qualitativer Analysen basierend auf ausgewählten Sequenzen aus einem der Dialoge von sich zuvor bekannten Gesprächspartnern (Abschnitt 3 beinhaltet Beispiele aus einem zweiten Dialog). Alle TeilnehmerInnen sind Rechtshänder.

Der Versuchsaufbau sieht folgenderweise aus: Jeweils zwei Gesprächspartner sitzen sich im Zentrum des Labors auf erhöhten Stühlen gegenüber. Die kommunikativen Bewegungen der Teilnehmer werden durch insgesamt 14 Infrarotkameras sowie eine HD-Videokamera erfasst. Zwecks Erfassung der Bewegungen mit den Infrarotkameras wurden bei jeder teilnehmenden Person 35

reflektierende Marker an den Händen, Armen, Schultern, Knie, Hüfte und Nacken befestigt.

Um das Aushandeln unterschiedlicher Reisepräferenzen, die Interaktion zwischen den Gesprächspartnern sowie das Aktivieren mentaler Karten und Frames zu fördern, hatten die Gesprächspartner zur Aufgabe, eine gemeinsame Reise zu planen. Vor dem Gespräch wurde ihnen mitgeteilt, 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).

Der lautsprachliche Diskurs wurde transkribiert und alle Gesten annotiert. Da es in dieser Studie in erster Linie um spontane redebegleitende Handgesten geht, wurden andere kommunikative Körperbewegungen wie beispielsweise Kopfgesten und Blickverhalten nicht detailliert transkribiert, sondern nur Auffälligkeiten notiert. Die verbalen Äußerungssequenzen wurden anhand einer Kombination der Diskurs-Transkriptionskonvention nach Du Bois et al. (1993) und Duncan (Duncan in McNeill 2005: 275ff.) in Intonationseinheiten unterteilt und transkribiert. Dabei wurden einige Details dem Erkenntnisinteresse der vorliegenden Studie angepasst. Des Weiteren wurden die Gesten mit Hilfe der Annotationssoftware Elan annotiert. Bei der Annotation der Gesten lag der Fokus auf den folgenden drei Aspekten: 1.) kinetische Eigenschaften, 2.) Beziehung zur gesprochenen Äußerung sowie 3.) die exakte Sprach-Gesten-Synchronie. So wurde der Verlauf jeder Gestenbewegung in eine typographische Repräsentation übersetzt und mit der Sprachtranskription korreliert. Jede Gesteneinheit wurde in Phasen von Ruheposition zu Ruheposition segmentiert. Gesteneinheiten werden in den folgenden Analysen, wie bereits in den Beispielen im ersten Teil des Artikels, durch eckige Klammer um die parallel hierzu stattfindenden verbalen Äußerungselemente markiert. Im Anschluss an jede Klammer wird jede Geste anhand eines hochgestellten G sowie ihrer jeweiligen Nummer gekennzeichnet (Beispiel: [...]^{G1}). Fokussiert werden solche Eigenschaften einer Geste, die am deutlichsten zu der Bedeutung der multimodalen Äußerung beitragen (*stroke* oder *stroke hold*, siehe Kita et al. 1998). Für die Bedeutung einer Geste können sowohl die Bewegungstrajektorie, die Handform, die Position zum Körper und im Gestenraum sowie die Orientierung der Handflächen ausschlaggebend sein (Stokoe 2005; siehe auch Bresssem 2013). Die fett markierten Elemente in den Sprachtranskripten repräsentieren Segmente, auf die der bedeutungstragende *stroke* fällt; zwei vertikale Linien orten einen *stroke hold*.

Im Detail beschrieben werden im Folgenden nur solche multimodalen Sequenzen, die für die inhaltlichen Schwerpunkte der vorliegenden Studie relevant sind. Diskursequenzen, die zwar gestisch begleitet werden, jedoch andere als die hier im Fokus stehende Funktionen ausüben, werden kurz zusammengefasst, um zu gewährleisten, dass dem Gesprächsverlauf gefolgt werden kann. Gesten, bei denen die Handform und/oder die Interaktion der Gesprächspartner im Vordergrund steht, werden anhand von Videostills gezeigt. Zusätzlich zu den bedeutungstragenden Momentaufnahmen aus den Videodaten, werden wie bereits in den vorigen Beispielen ausgewählte Gesten anhand von Motion-Capture-Bewegungsspuren dargestellt. Letztere eignen sich, um den

Verlauf der Gestenbewegung durch den drei-dimensionalen Gertenraum aufzuzeigen und dienen hier insbesondere zur Visualisierung von Gerten-Diagrammen.

4.2. Qualitative Analyse multimodaler Interaktion mit Fokus auf Reiseroutendiagramme

Im Folgenden fokussieren wir unseren Blick nicht nur auf das allmähliche Erstellen von Diagrammen, sondern auch auf die der Lautsprache und dem Körper eigenen indexikalischen Ausdrucks- und Hinweismittel (z.B. Zeigegesten, Präpositionen und Lokaladverbien), um ihre unterschiedlichen referentiellen und pragmatischen Funktionen zu eruieren (siehe auch Abschnitt 2).

4.2.1 Indexikalische Phase: Subjektive Äußerung und intersubjektive Anerkennung erster Wunschziele

(5) Indexikalische Gesten mit verortender und interaktiver Funktion (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 []^{GPL1 G17}
19 Osten
20 ist mir eigentlich egal
21 Kiew?
22 Kommen wir dahin?
23 GP_{L1}: ja
24 Osten können wir machen
25 GP_{R1}: ja ne



Abb. 8: Indexikalische *Citing* Geste signalisiert Zustimmung (G11)



Abb. 9: Modale Geste signalisiert Offensichtlichkeit (G14)



Abb. 10: *Shared Information Delivery* Geste (G15)



Abb. 11a: Indexikalische Geste markiert Sprecherrolle und Subjektivität (G16)

Abb. 11a-b: Interaktive indexikalische Gesten mit unterschiedlichen pragmatischen Funktionen



Abb. 11b: Nachahmung von G16 (GP_{L1}) affirms Sprecherrolle von GP_{R1} (G17)

Die Gesprächspartnerinnen GP_{L1} und GP_{R1} kennen sich von ihrem gemeinsamen Arbeitsplatz. Sie beginnen ihre Reiseplanung mit einer Auflistung von Wunschorten innerhalb Europas. Der Gesprächspartnerin (in den Abbildungen auf der linken Seite situiert (GP_{L1})) fällt direkt Barcelona als Wunschziel ein (01-04). Ihre Gesprächspartnerin (im Bild rechts (GP_{R1})) reagiert positiv auf diesen ersten Vorschlag (05); sie hebt ihre Stimmlage und zeigt einen fröhlichen Gesichtsausdruck. Auf „das“ zeigt sie mit ihrem ausgestreckten Zeigefinger auf ihre Gesprächspartnerin. Die Geste (Abb. 8; G11) weist eine indexikalische Beziehung zum Gesagten auf und signalisiert Zustimmung (vgl. Bavelas 1994). GP_{L1} nennt Lissabon und Stockholm als zwei weitere Ziele (06-08). GP_{R1} macht darauf aufmerksam, dass ihre Gesprächspartnerin bereits drei Ziele (von fünf möglichen) genannt hat (10-11). Auf „du“ öffnet sie kurz ihre rechte Handfläche und zeigt gleichzeitig mit den Fingerspitzen auf ihr Gegenüber (Abb. 9; G14) und unterstreicht somit, 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 (12-17). Auf den Ortsverweis „Stockholm“ führt sie mit der rechten Hand eine Zeigegeste mit nach oben gedrehter Handfläche und ausgestrecktem Zeigefinger aus (Abb.10; G15). Diese Geste hat zugleich referentielle und diskursive Funktion, indem sie auf einen von der Gesprächspartnerin zuvor genannten Ort anaphorisch verweist und ihn dabei mit einer gemeinsamen Reise

(„waren wir doch“) verknüpft (siehe auch *Shared Information Delivery* Geste nach Bavelas 1994 und Bavelas et al. 1995). Es wird ein neuer Frame, „Stockholm“, eingeführt; es folgt eine kurze Diskussion über diese Reise.

Sich auf die gestellte Aufgabe besinnend, überlegt sich GP_{R1} nun selbst ein Wunschziel (18-22). Auf „ich möchte“ (18) beugt sie ihren Oberkörper nach vorne, hebt ihren rechten Zeigefinger und zeigt damit auf ihre Gesprächspartnerin (Abb.11a; G16). Hierdurch kündigt sie das kurz bevorstehende Eröffnen eines neuen, ihrem Wunschziel entsprechenden Frames an, erweckt die Aufmerksamkeit ihrer Gesprächspartnerin und markiert gleichzeitig ihre Sprecherolle. Dann zieht sie ihren Zeigefinger an ihren spitz zugezogenen Mund, was Nachdenken und Zögern andeutet. Während GP_{R1} überlegt (18), ahmt GP_{L1} die vorhergehende Zeigepose ihrer Partnerin in leicht geänderter Ausführung nach. GP_{L1} zieht ihren rechten Zeigefinger auf Augenhöhe und zeigt auf ihre Gesprächspartnerin (Abb. 11b; G17). Mit dieser *Acknowledgement Citing* Geste (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 (23-24).

In dieser ersten Sequenz dienen die redebegleitenden indexikalischen Gesten dem Vorschlagen und Zustimmen hinsichtlich erster Reisezielvorschläge, dem Verweis auf gemeinsame Erfahrungen (*common ground*) sowie dem Signalisieren von Aufmerksamkeit. Somit kommt diesen Gesten eine primär interaktive Funktion zu. Auch wenn dabei erste Städteframes aktiviert werden, spielt die Verortung der Punkte im Raum und zueinander hier keine zentrale Rolle. In der folgenden Gesprächssequenz geht es um Verortung und Richtungen, zunächst jedoch noch auf approximative Arte und Weise.

4.2.2. Indexikalisch-ikonische Phase: Grobe Richtungen und erste Verbindungslien zwischen Städten und Regionen

(6) Anfang gemeinsamer Routenplanung mit groben Zielen und Richtungen
(01.16-01.42)

```

01GPL1: man könnte [auchG18
02      auch wenn man sich im Osten bewegtG19
03      sich von NordenG20 nach Süden
04      durch eh]G21
05      schlagen
06      und dann [zum Schluss]G22 noch so zwei drei Tage an Strand
07GPR1: das heißt wir machen quasi [erst son bisschen]G23
08GPL1: []G24 nach Kroatien
09GPR1: [erst son bisschen Richtung]G25 [eh]G26
10     Schweden
11     [dann rüber]G27
12     [nach]G28
13GPL1: [Estland Lettland]G29[]G30
14GPR1: so in die Richtung
15     dann könnten wir [runter]G31
16     das heißt wir brauchen noch [irgendwie]G32
17     Griechenland oder so was

```



Abb. 12a: Zögernder Ansatz einer Verortung (G18)

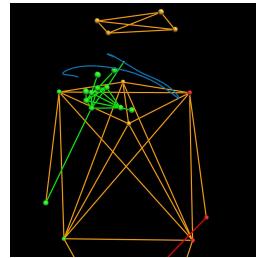


Abb. 12b: Gestischer Bezug „im Osten“ (G19)

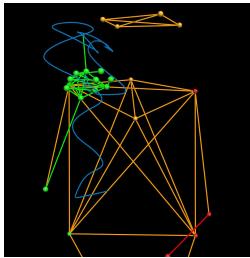


Abb. 12c: Diagramm: Von Norden nach Süden durchschlagen (G21)

Abb. 12a-d: Erster grober Routenvorschlag orientiert an Himmelsrichtungen



Abb. 12d: ENTSPANNUNG-Frame (G22)

GP_{L1} beginnt einen ersten groben, an den Himmelsrichtungen orientierten, Routenverlauf von Nord-Ost nach Süden (01-06), welcher durch fünf redebegleitende Handgesten allmählich im Gestenraum entsteht. Auf „auch“ (01) zieht sie ihren rechten Zeigefinger nach oben (Abb. 12a; G18), doch während sie verbal den möglichen Beginn der Reise im Osten anspricht (02), zeichnet sie mit ihrer rechten gekrümmten Hand zunächst eine leicht gebogene Linie nach rechts (Abb. 12b; G19). Diese Handform und die kurze angedeutete Linie dienen 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*; Dancygier/Sweetser 2012) rechts von ihrem Körper befindet. Dann führt sie ihre Beschreibung fort und äußert „sich von Norden nach Süden durch eh schlagen“ (03-05). Auf „Norden“ (3) setzt sie erneut ihren Zeigefinger ein wenig höher im Gestenraum an. Sie markiert somit einen imaginären Punkt, der auf „Norden“ referiert, und zieht dann auf „nach Süden durch eh schlagen“ (03-04) eine Schlangellinie bis in den unteren Teil des Gestenraums (Abb. 12c; G21). Die Schlangellinie drückt nicht nur die verbal erläuterte Richtung der Reise von Norden nach Süden aus, sondern impliziert zusätzlich durch den Schlangel-Charakter, dass der genaue Verlauf noch zu klären ist: Welche Städte- und Länderframes noch dazu kommen, muss erst noch ausgehandelt werden. Auf „zum Schluss“ (06) hält GP_{L1} beide Hände nebeneinander relativ niedrig im Gestenraum; die offenen gespreizten Handflächen zeigen nach unten (Abb. 12d; G22). Mit diesem Ausdruck eröffnet sie einen ENTSPANNUNG-Frame: Die Geste verkörpert den Prozess des Entspannens und einen Ruhezustand (siehe auch Rekittke/Wolf/Mittelberg 2015).

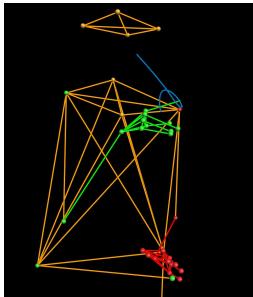


Abb. 13a: Schweden (G25)



Abb. 13b: Geste markiert Hesitation (G26)

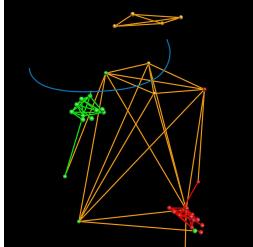


Abb. 13c: Von Schweden rüber nach Osten (G27)



Abb. 13d: GP_{R1} Geste markiert Zögern (G28)

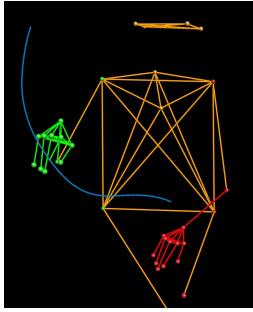


Abb. 13e: Gestische Bewegung Richtung „Estland Lettland“ (G29) (GP_{L1})

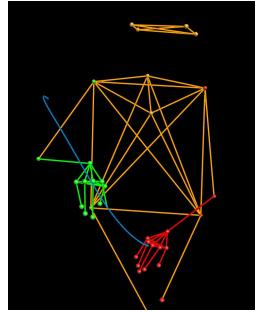


Abb. 13f: Erneut Bewegung nach unten (G30) (GP_{L1})

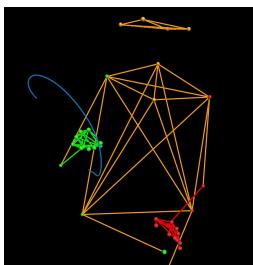


Abb. 13g: Gestische Bewegung „runter“ in Richtung „Griechenland“ (G31) (GP_{R1})

Abb. 13 a-h: Etappenweise gemeinsam entwickeltes Reiserouten-Diagramm



Abb. 13h: Hand bleibt in-situ auf „irgendwie“ (G32)

GP_{R1} greift daraufhin die vorgeschlagene, noch relativ grobe Route auf (07-17). Auf „erst so’n bisschen“ (07) zieht sie zunächst mit der rechten Hand eine Linie von links nach rechts mittig in ihren Gestenraum. Sie beginnt dann in ihrem oberen Gestenraum ein eigenes Diagramm aus ihrer Perspektive. Zunächst nimmt sie noch keine konkrete Lokalisierung vor. GP_{L1} wirft „nach Kroatien“ (08) ein; sie ist in Gedanken noch bei ihrem finalen Reiseziel und dem Wunsch nach Entspannung. Hierbei lässt sie ihren Zeigefinger grob gen rechts zeigen und zeichnet dann eine kurze Linie nach unten. Währenddessen führt GP_{R1} ihre

Reiseplanung weiter und wiederholt „erst son bisschen Richtung“ (09). Während dieser unpräzisen verbalen Äußerung zeichnet sie, ebenfalls mit der rechten Hand, eine Linie von links unten nach rechts oben in ihrem Gestenraum (Abb. 13a; G25). Obwohl sie ihr erstes Wunschziel noch nicht ausgesprochen hat, weist ihr Zeigefinger bereits nach oben und somit nach Norden. Sie überlegt kurz: „eh“ (09). Hierbei verharrt ihre rechte Hand in derselben Position, sie streckt die Finger nach oben und vollzieht mit ihnen eine wellenförmige Bewegung, welche eine kurze Denkpause suggeriert und gestisch füllt (Abb. 13b; G26). Mit der Routenplanung fortfahrend (10-11), greift sie das Konzept ihrer Gesprächspartnerin auf und konkretisiert daraufhin die grob genannten Himmelsrichtungen. Auf „dann rüber“ (11) zeichnet sie eine weitere Kurve, dieses mal von oben mittig nach rechts außen, ein wenig tiefer in ihrem Gestenraum (Abb. 13c); G27), und setzt dort einen nächsten Punkt, jedoch ohne ihn verbal zu bezeichnen. Sie markiert die ungenaue Zieldefinition durch eine Wiederholung der wellenförmigen Bewegung der Finger ihrer rechten Hand (Abb. 13d; G28). GP_{L1} geht auf dieses Diagramm ein und füllt die Lücke in der dynamisch-diagrammatischen Routenplanung, indem sie „Estland, Lettland“ (13) als mögliche Zwischenstopps im Osten benennt. Hierbei zieht sie mit der rechten Hand eine lange, aufsteigende Linie von links unten nach rechts oben (Abb. 13e; G29). Sie greift somit die Route ihrer Gesprächspartnerin auf und integriert diese zusätzlichen Informationen aus ihrer eigenen Perspektive. Sie zeichnet hierbei Estland und Lettland in Relation zu Deutschland ein. Objektiv betrachtet, befindet sich Deutschland süd-westlich von diesen baltischen Ländern. Auf ihre subjektiv ausgerichtete mentale Karte von Europa projiziert, befinden sich Estland und Lettland so auch rechts oben relativ zu Deutschland, was sie durch eine gerade Linie nach rechts oben illustriert (Abb. 13f; G30).

GP_{R1} hört ihrer Gesprächspartnerin aufmerksam zu, hält dabei ihre Hand (ähnlich wie in G28) nach oben und signalisiert GP_{L1} durch Kopfnicken auf „so in die Richtung“ (14) ihr Einverständnis. Sie fügt dann Griechenland als weitere potentielle Etappe im Süd-Osten von Europa hinzu (15-17). Dadurch eröffnet sich in Bezug auf diese größere geographische Region ein weiterer Länderframe. Während sie „runter“ (15) sagt, geht parallel ihre rechte Hand mit ausgestrecktem Zeigefinger nach links mittig unten (Abb. 13g; G31) und visualisiert so im Gestenraum einen Weg, der sich wahrscheinlich auf ihrer mentalen Karte von Estland oder Lettland in den Süden nach Griechenland erstreckt. Auf „irgendwie“ (16) wippen ihre Finger erneut hoch und runter (Abb. 13h; G32).

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.

4.2.3. Diagramatisch-ikonische Phase: Einigung auf komplette Reiseroute

Die folgende Sequenz ist uns bereits in Teilen von der Einleitung und Abschnitt 3 bekannt. Mit Fokus auf die kollaborativen Prozesse cross-modaler Bedeutungskonstitution präsentieren wir nun die komplette, multimodal

kontextualisierte Reiseroute, auf die sich die Gesprächspartnerinnen letztendlich einigen.

(7) Resümierendes Routendiagramm im peripheren Gestenraum (3.15-3.48)

01GP_{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}
 10GP_{L1}: ich weiß überhaupt nicht wo Kiew liegt [relativ zu Prag]^{G35}
 11GP_{R1}: []^{G36}
 12 ich auch nicht %lacht
 13GP_{L1}: %lacht aber gut angetäuscht^{G37}
 14GP_{R1}: []^{G38} ich denke
 15 [neee]^{G39} ich denke wir fahren dann so weiße [von
 16 da
 17 nach da
 18 rüber
 19 runter
 20 und dann nach da und dann wieder nach Hause]^{G40}.
 21 Nein?
 22GP_{L1}: doch



Abb. 14: Indexikalische Geste der Zustimmung (G33; *Citing Gesture*)

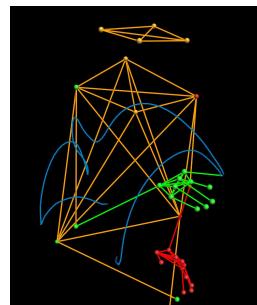


Abb. 15: GP_{R1} zeichnet Diagramm (G34; vgl. Abb. 7)



Abb. 16: Vages räumliches Inbezugsetzen (G35)



Abb. 17: Indexikalische Geste markiert Zögern (G36)



Abb. 18: *Acknowledgment Geste (G37) (GP_{L1})*

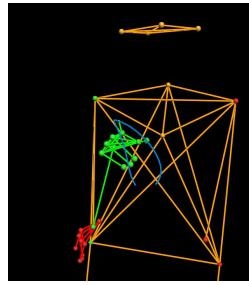


Abb. 19: GP_{R1} Geste nimmt folgende Beschreibung schon ungefähr vorweg (G38)

Nach fortgeschrittenem Gespräch sind sich die Gesprächspartnerinnen einig erst über das Baltikum zu reisen. GP_{L1} fällt Prag als potentieller Zwischenstopp ein. GP_{R1} teilt GP_{L1} daraufhin mit, dass sie bereits in Prag war (also bereits einen mit persönlichen Erfahrungen angereicherten Frame von der Stadt hat), und dass sich Prag sehr gut als Etappe in die Reiseplanung integrieren ließe (01-05). Auf „können wir auch machen“ (05) zeigt sie mit dem ausgestreckten Zeigefinger der rechten Hand auf ihre Gesprächspartnerin und signalisiert so ihre Zustimmung (Abb. 14; G33). Daraufhin präsentiert sie multimodal, wie sie sich die konkrete Reiseroute vorstellt (07-09). Dabei zeichnet sie zunächst aus ihrer eigenen Perspektive ein Diagramm in den geteilten Gesterraum zwischen ihr und GP_{L1} (Abb. 15; G34). Sie beginnt mit ausgestrecktem Zeigefinger links auf Schulterhöhe und setzt dort einen Punkt auf „Stockholm“ (07), zieht dann eine Kurve runter bis in die Mitte ihres Gesterraums und setzt dort einen weiteren Punkt auf „Prag“ (08). Sie positioniert den Finger noch einmal dort, zeitgleich mit „und von Prag“ (09), und zieht dann eine Linie weiter nach rechts, um dort (weiter im Osten) Kiew zu verorten. Auf „von da nach unten“ setzt sie noch einmal dort an und zieht die Linie weiter nach unten und dann mittig im Gesterraum. GP_{L1} gibt daraufhin zu bedenken, dass ihr die genaue Lage von Kiew nicht bekannt sei (10). Wir können somit annehmen, dass sie keinen differenzierten Frame von Kiew und der Region hat. In diesem Moment streckt sie ihren rechten Zeigefinger kurz nach oben und deutet dabei ohne einen Ort zu präzisieren in den Raum (Abb. 16; G35). GP_{R1} überlegt kurz, platziert ihren Zeigefinger noch einmal im Gesterraum und gibt dann zu, dass sie selbst nicht wisse, wo genau Kiew liegt (Abb. 17; G36). Die nicht konkret auf ein Ziel deutende Zeigegeste scheint dabei zu signalisieren, dass sie die geographische Lage der beiden Städte nicht gut kennt. Wir nehmen an, dass angesichts dieser ‚leeren Flecken‘ auf ihrer mentalen Karte, auch ihre gestischen Verweise vage und unentschlossen bleiben.



Abb. 20a: Beginn des Routendiagramms aus Frontalsicht für GP_{R1} (G39)
Abb. 20a-b: Perspektivwechsel und komplettes Routendiagramm

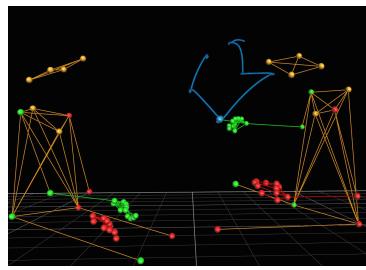


Abb. 20b: Geteilte Perspektive auf Gestendiagramm in der Peripherie (G40)

Als nun beiden deutlich wird, dass sie keine sehr detaillierten Vorstellungen (oder Frames) der thematisierten Länder bzw. Städte haben, fangen sie an zu lachen. GP_{L1} begegnet GP_{R1} mit „aber gut angetäuscht“ (13), zeigt mit ihrem ausgestreckten Zeigefinger auf sie und lacht weiter (Abb. 18; G37). GP_{R1} zieht mit ihren Zeigefinger eine kleine Kurve von einem Punkt im oberen Gestenraum nach rechts unten, dann wieder nach links und darauf wieder nach rechts (Abb. 19; G38). Mit diesem gestischen Verhalten zeigt sie ihrer Gesprächspartnerin, dass sie immer noch versucht auszutarieren, wo Kiew in Relation zu Prag liegt. Anschließend fasst sie den geplanten Routenverlauf zusammen (15-20). Auf „nee“ (Abb. 20a; G39) setzt sie kurz ihren Zeigefinger links oben in ihrem Gestenraum an, nimmt dann aber einen Perspektivwechsel vor, indem sie zunächst ihren Blick nach rechts wendet und auf „von“ (15) in der linken äußeren Peripherie ihren Zeigefinger etwas oberhalb ihrer Kopfhöhe neu platziert. Sie lädt so ihre Gesprächspartnerin ein, dieselbe Perspektive auf die folgenden multimodalen Erläuterungen einzunehmen, indem sie einen anders gelagerten geteilten Gestenraum in der Peripherie kreiert. Nun wendet auch GP_{L1} ihren Blick seitwärts und folgt der Zeigegeste von GP_{R1}. GP_{R1} zieht auf „nach da“ (17) eine Linie nach unten und auf „rüber“ (18) eine kurze, relativ gerade Linie nach rechts. Auf „runter“ (19) zeichnet sie entsprechend eine schräg nach links unten verlaufende Linie und auf „und dann nach da“ (20) bewegt sie ihren Zeigefinger weiter nach links und schräg nach oben (Abb. 20b; G40). 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 Zusammenfassung lediglich mit sprachlichen Deiktika auf die Punkte ihres Gestendiagramms hin, welche die bereits besprochenen Stationen der Reise repräsentieren. Auf „und dann wieder nach Hause“ (20) 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 (21). 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 Schlangellinien und skizzenhaften Teildiagrammen ist dieses abschließende Diagramm (Abb. 19c) von einer klaren Linienführung gekennzeichnet und lässt sich mit Blick auf die Geographie Europas durchaus nachvollziehen. Die mentalen Karten der beiden Gesprächspartnerinnen werden durch die Gestendiagramme – aus den jeweiligen Perspektiven – so beiden semiotisch habhaft und dienen der geteilten Aufmerksamkeit und Koordination von Vorschlägen und Präferenzen. Die Diagramme und aktivierten Frames scheinen dabei das Gespräch 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 klare Vorstellung der gemeinsamen Reise.

5. Diskussion: Indexikalische Bezüge orchestrieren dynamische multimodale Gesprächskontexturen

Aus unseren theoretischen Überlegungen und multimodalen Datenanalysen ergibt sich eine Reihe von Beobachtungen, von denen an dieser Stelle diejenigen, welche sich für das interaktive Diskursverhalten und die multimodale Bedeutungskonstitution als konstitutiv erwiesen haben, resümierend herausgestellt werden sollen. Die indexikalische Bedingtheit multimodaler Interaktion wurde bereits in Abschnitt 2 als ein den Gebrauch gestischer und sprachlicher Ressourcen stark motivierendes Prinzip erörtert, und auf ihre materiellen und kontextuellen Grundlagen zurückgeführt (Mittelberg 2013; Sweetser 2012). Unsere Analyseergebnisse bestätigen nun, dass der Gebrauch von Inhalts- und Funktionswörtern sowie von primär indexikalischen und primär ikonischen (sowie in hohem Grade konventionalisierten) Gesten prinzipiell perspektiviert (*viewpointed*), teilhaft (metonymisch) und situativ verankert (pragmatisch bedingt) ist. Mit Blick auf die in diesem Artikel analysierten Daten seien im Folgenden drei unterschiedlich wirkende, an der Orchestrierung von Diskursinhalten, Einigungsprozessen und Ausdrucksmodalitäten teilhabende, indexikalische Modi abschließend vor Augen geführt: semantischer Bezug (z.B. Referenzpunkte), interaktive Indizes, sowie intra- und intermodale Bezüge.

5.1. Verkörperte Referenzpunkte

Das mentale Aufspannen von Framestrukturen sowie das gestische Andeuten von Inhalten, Richtungen und Linien wird, so unsere im ersten Teil des Artikels vorgelegte und hier bekräftigte These, wenigstens zum Teil durch *verkörperte Referenzpunkte* (*embodied reference points*, Mittelberg erscheint) angestoßen, welche gleichzeitig sprachlich und situativ kontextualisiert werden. Unter verkörperten Referenzpunkten verstehen wir in Anlehnung an Langacker's Referenzpunkte (*reference-point constructions*; Langacker 1993) gestische Formen und Indizes, die zentrale Punkte in der Bedeutungskonstruktion darstellen und mentalen Zugang zu Zielbereichen gewähren (Mittelberg 2006; Talmy 2013).¹¹ Anhand der Gesten ist dieser Zugang nicht nur mental, sondern auch körperlich, mit anderen Worten: kognitiv-aktional. In unseren Daten dienen Gesten beispielweise als Eingangspunkte in mentale Karten und semantische Framestrukturen (z.B. ‚Europa‘, oder ‚Kiew‘). Durch das Markieren von Punkten und Regionen im Geste.raum, auf die sie teils recht präzise aber teils auch recht vage und approximativ verweisen (z.B. ‚irgendwie nach Süden‘), können indexikalische Gesten auch Ansatzpunkte von entstehenden semiotischen Strukturen (z.B. virtuellen Diagramme) darstellen. Mittels geringerer Grade von Indexikalität und nachempfundener Kontiguitätsbeziehungen können gestikulierende Hände zudem imaginäre Objekte, die sie zu halten oder zu

¹¹ 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".

manipulieren scheinen, suggerieren (Mittelberg 2010; Mittelberg/Waugh 2014; Müller 1998; Streeck 2009).

Bei diesen Praktiken der multimodalen Verortung und Andeutung von zu inferierenden semantischen (hier insb. geographischen) Zielbereichen kommen Prozesse der pragmatischen Inferenz und somit metonymische Modi zum Tragen (Barcelona 2009; Mittelberg 2013; Panther/Thornburg 2003). Diese ermöglichen es den Gesprächspartnern, weiterführende Assoziationen, z.B. Reiseideen, gedanklich zu verfolgen, multimodal weiterzuentwickeln und dabei zur gemeinsamen Betrachtung buchstäblich in den Raum zu stellen. Aktivierte Frames implizieren nicht nur größere Wissensnetzwerke, sondern evozieren auch subjektiv geprägte Elemente, die durch wiederkehrende, lebensweltliche wie kognitiv-semiotische Erfahrungen korreliert sind. Sie bieten Schnittstellen für neue ‚Brückenschläge‘ und damit für kreatives Denken und semiotisches Handeln während des Diskursgeschehens (Coulson 2001; Dancygier/Sweetser, 2014; Mittelberg erscheint; Ziem 2014). Unsere Analysen unterstützen gleichzeitig die Annahme, dass durch das Integrieren von und das indexikalische Bezugnehmen auf im Kontext oder in der Vorstellung existierende Strukturen und Gegenstände (s. *imaginary material anchors*; Hutchins 2005; Abschnitt 2.3) diese subjektiven und intersubjektiven Praktiken des diagrammatischen Denkens und Gestikulierens an dialogischen Prozessen der *distributed cognition* teilhaben (Hutchins/Saeko 2011; Streeck/Goodwin/LeBaron 2011).

5.2. Interaktive Indizes

Indexikalische Handformen und Handbewegungen sind, wie bereits mehrfach betont, die prädominanten gestischen Formen in den hier analysierten Daten. Ein zentrales Ergebnis dieser Studie ist, dass unterschiedlich ausgeführte indexikalische Gesten im Aushandeln von Reiseplänen ein breites Spektrum an interaktiven und diskursiven Funktionen aufweisen (Bavelas 1994; Bavelas et al. 1995; Kendon 2004; Müller 1998). Dies wird besonders deutlich im allmählichen Aushandeln von Kompromissen zwischen den Gesprächspartnern, indem die Dialogpartner zum Beispiel durch punktuelle, grob die Richtung anzeigende sowie zeichnerische Gesten verschiedene virtuelle Orte und Routenabschnitte im Gestenraum erstellen (vgl. Fricke 2007, in diesem Heft).

Zusammenfassend lässt sich sagen, dass interaktive, direkt auf die Gesprächspartnerin gerichtete Gesten besonders häufig in der ersten Phase des Gesprächs vorkommen, in dem es um 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}) geht. Dazu kommen die indexikalischen Gesten, die sich auf gemeinsame Erfahrungen und geteiltes Wissen, und damit vergleichbaren Frames („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 Zustimmen („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 *Turn* Gesten oder PUOH-Gesten, mit

denen Sprecher signalisieren, dass ihre Hände leer sind und sie eine Antwort vom Gesprächspartner erbitten (Abb. 2, G1, *Seeking Geste* nach Bavelas 1994; siehe auch Müller 2004). Hier schließen sich zahlreiche weiterführende Fragestellungen hinsichtlich interagierender Funktionen wie Einstellung (*stance*), Ironie, gegenseitige Herausforderung oder Besänftigung an. Was in den besprochenen Beispielen multimodaler Bedeutungskonstitution deutlich hervortritt, ist, dass es Gesprächsphasen gibt, in denen bestimmte indexikalische Praktiken stärker vertreten sind als andere: So ist zum Beispiel die erste Phase (4.2.1) prädominant indexikalisch hinsichtlich des multimodalen Ausdrucks subjektiver Wünsche und intersubjektiver Reaktionen.

Inwiefern die hier beobachteten, interaktiven Praktiken unterschiedliche Modi und Grade von Indexikalität aufweisen und dabei mit pragmatischen Faktoren und Grammatikalisierungsprozessen interagieren, sind interessante Aspekte für weiterführende Untersuchungen.

5.3. 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 sie weisen zudem sprachliche und gestische Wiederaufnahmen und Rückverweise auf. Dabei lassen sich *intramodal* bedeutungsstiftende von *intermodal* bedeutungsstiftenden Prozessen wie folgt unterscheiden.

Im sprachlichen Diskurs stellen anaphorische Verweise, etwa durch Personal- und Demonstrativpronomina, intramodal Rückbezüge her. In der gestischen Modalität beziehen sich Formen aufeinander, indem zum Beispiel Zeigegesten auf bereits gestisch gesetzte Positionen im Gestenraum erneut verweisen (Fricke 2007, dieses Heft). Auch das Wiederaufnehmen von ikonischen Gestenformen, die in derselben Gesprächssequenz bereits von demselben Sprecher oder vom Gegenüber in ähnlicher Weise erzeugt wurden, wirkt kohärenzstiftend (vgl. *catchments* oder *mimicry*, McNeill 2005). Die in Abbildung 3e-f wiederholte Schwimmgeste ist ein Beispiel für das ikonische Nachahmen einer von der Gesprächspartnerin ausgeübten Geste (*body action image icon*, Mittelberg/Waugh 2014).

Intermodale bedeutungsstiftende Prozesse (*cross-modal*) kommen besonders dann zum Tragen, wenn an sich vieldeutige gestische oder sprachliche Formen nur durch in der anderen Modalität synchron veräußerte Hinweise disambiguier und verstanden werden können. Sprachliche Deiktika wie ‚da‘, ‚darunter‘ und ‚von dort‘ beziehen ihre lokale Bedeutung aus dem sprachlichen und außersprachlichen Kontext, oder was Jakobson (1956) ‚semiotische Kontexturen‘ und Goodwin (2011) mit Bezug auf multimodale Interaktion ‚contextures of action‘ nennen. Wie wir gesehen haben, 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 (Kendon 2004). Sie können so zu den kontextbedingten Funktionselementen gezählt werden, die Jakobson (1971) 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. eine Kombination von Indizes und Ikonen nötig (Peirce 1960): Ohne einen Index kann ein Ikon keine lokale Bedeutung entfalten, was für bildliche wie sprachliche und gestische Zeichen gilt (Mittelberg 2006).

Nun konnten wir hier eine zusätzliche Variante bimodal erzeugter Indexikalität beobachten, und zwar bei der Ausführung des letzten, durch eine klare Linienführung hervorstechenden Reiseroutendiagramms. Das diagrammatisch-ikonische Gerüst, das gestisch von GP_{R1} in die Luft gezeichnet wurde, unterfütterte virtuell durch gedachte Linien und indexikalisch verankerte Punkte eine 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 und Regionen oder Ländern, sondern um Kombinationen, bestehend aus Präpositionen, Demonstrativpronomina und Lokaladverbien (darunter, rüber), und entweder primär deiktischen Gesten (Punkte, die durch das Verharren auf einer Position im Gestenraum entstehen), oder eben das Ziehen von Verbindungslien. Um die anaphorischen Prozesse in beiden Modalitäten – ohne Nennung von Eigennamen und Inhaltswörtern – nachzuvollziehen, muss die Gesprächspartnerin der Bedeutungskonstitution in beiden Modalitäten unter Berücksichtigung räumlicher Strukturen sehr genau gefolgt sein.

6. Abschließende Bemerkungen

Ein Hauptanliegen dieses Beitrages war es aufzuzeigen, inwiefern das Berücksichtigen von Frames, hinsichtlich der multimodalen Bedeutungskonstitution im Dialog, zu neuen Einsichten in die instrumentale Rolle von Gesten als semiotische Handlungen in diesen komplexen Prozessen geben können. Dabei war die indexikalische Verankerung und Bedingtheit multimodaler Interaktion eine Grundannahme, sowie, dass räumliche, materielle, soziale und interaktive Faktoren an verschiedene kognitive Prozesse, sensomotorische Muster und semiotische Prinzipien und Praktiken gekoppelt sind.

Angesichts der sprachlichen und gestischen Strategien, die Dialogpartner in der hier präsentierten Studie anwandten, um gemeinsam eine Europareise zu planen, scheinen Diagramme und Framestrukturen nicht nur das Entstehen von gestischen Darstellungen zu motivieren. Sie dienen dabei auch als kognitiv-aktionale Strukturen, welche das kollaborative, multimodale Entwickeln neuer Ideen und das gegenseitige Verstehen und Verständigen der Gesprächspartner unterstützen. Wie wir anhand der Motion-Capture Aufnahmen sichtbar machen konnten, erlangen Diagramme im dialogischen Aushandeln von Reiserouten eine gewisse, wenn auch flüchtige und virtuelle, semiotische Realität im Interaktionsraum.

Während Frames und Diagramme jeweils metonymisch organisierte Strukturen darstellen, sind gerade Frames mehr als hochabstrakte, konzeptuelle Strukturen; sie repräsentieren reichhaltige semantische Systeme, die individuelle lebensweltliche, soziale und semiotische Erfahrungen mit einschließen (Dancygier/Sweetser 2014; Ziem 2014). Die Analysen der untersuchten

Reiseplanungen geben Anhaltspunkte, inwiefern die Gesprächspartner über korrespondierende Frames oder ähnlich ‚gefüllte‘ bzw. ausdifferenzierte Framestrukturen in Hinsicht auf Städte, Länder und Kulturen verfügen. So konnten wir den Dialogen nicht nur unterschiedliche Wissens- und Erfahrungsbestände seitens der Gesprächsteilnehmer entnehmen, welche durch verschiedene Vorschläge sozusagen ‚abgeklopft‘ wurden, sondern auch Präferenzen, erneut in bekannte und bereits selbst bereiste Städte und Länder zu fahren (d.h. bezüglich derer Vorwissen und durch persönliche Erfahrungen angereicherte Frames bestehen), oder lieber Neuland zu erkunden. Letzteres würde das Integrieren zusätzlicher Wissens- und Erfahrungselemente sowie das Bilden neuer Verbindungen zwischen Frames ermöglichen.

Aus den im Laufe des Artikels sich herauskristallisierten Verfahren der dialogischen Bedeutungskonstitution erwiesen sich, neben ikonisch-diagrammatischen Praktiken, insbesondere verschiedene Modi und Grade von Indexikalität als die orchestrierenden Prinzipien in den analysierten dynamischen multimodalen Kontexturen: a) verkörperte Referenzpunkte als Einstiegspunkte in Framestrukturen, Diagramme und mentale Räume, b) interaktive Indizes des Sich-aufeinander-Beziehens, Sich-Einigens und des Turn-Managements, sowie c) intra- und intermodale (Rück-) Bezüge. Auf der Basis der hier dargelegten, auf qualitativen Analysen beruhenden Beobachtungen bietet es sich an, die in dieser Studie offen gelegten Tendenzen unter Zuhilfenahme quantitativer Methoden im gesamten MuSKA Datenkorpus, insbesondere in den numerischen Motion-Capture Daten, auf intramodale Muster und intermodale Clusterbildung hin zu untersuchen.

7. Literaturverzeichnis

- Andrén, Mats (2012): Children's gestures from 18 to 30 months. Lund: Lund University Press.
- Arnheim, Rudolf (1969): Visual Thinking. Berkley: University of California Press.
- Auer, Peter / Pfänder, Stefan (Hg.) (2012): Constructions. Emerging and Emergent. Berlin: Mouton de Gruyter.
- Barcelona, Antonio (2009): Motivation of construction meaning and form. The roles of metonymy and inference. In: Panther, Klaus-Uwe Panther / Thornburg, Linda / Barcelona, Antonio (Hg.), Metonymy and metaphor in grammar. Amsterdam, Philadelphia: John Benjamins, 363–401.
- Barsalou, Lawrence (1999): Perceptual symbol systems. In: Behavioral and Brain Sciences 22, 577–609.
- Barsalou, Lawrence (2010): Grounded Cognition. Past, Present and Future. In: Topics in Cognitive Science 2, 716–24.
- Bavelas, Janet (1994): Gestures as part of Speech. Methodological implications. In: Research on Language and Social Interaction 27, 201–221
- Bavelas, Janet / Chovil, Nicole / Lawrie, Douglas / Wade, Allan (1992): Interactive gestures. In: Discourse Processes 15, 469–489.

- Bavelas, Janet / Chovil, Nicole / Coates, Jan L. / Roe, Lori (1995): Gestures specialized for dialogue. In: *Personality and Social Psychology Bulletin* 21, 394–405.
- Bergen, Benjamin (2007): Experimental methods for simulation semantics. In: Coulson, Seana / Gonzalez-Marquez, Monica / Mittelberg, Irene / Spivey, Michael J. (Hg.), *Methods in Cognitive Linguistics*. Amsterdam: John Benjamins, 277–301
- Bergen, Benjamin (2012). Louder than words. The new science of how mind makes meaning. New York: Basic Books.
- Brenger, Bela / Mittelberg, Irene (2015): Shakes, nods and tilts. Motion-capture data profiles of speakers' and listeners' head gestures. In: Proceedings of the 3rd Gesture and Speech in Interaction (GESPIN) Conference, Sept. 2015, Nantes.
- Bredekkamp, Horst (2010): Theorie des Bildakts. Frankfurt: Suhrkamp.
- Bressem, Jana (2012): Repetitions in gesture: Structures, Functions, and Cognitive Aspects. Dissertation, European University Viadrina, Frankfurt (Oder).
- Bressem, Jana (2013): A linguistic perspective on the notation of form features in gestures. In: Müller, Cornelia / Cienki, Alan / Fricke, Ellen / Ladewig, Silvia H. / McNeill, David / Tessendorf, Sedinha (Hg.), *Body – Language – Communication. An international Handbook on Multimodality in Human Interaction*. Berlin, Boston: Mouton de Gruyter, 1079–1098.
- Bühler, Karl (1934/1982): Sprachtheorie. Die Darstellungsfunktion der Sprache. Stuttgart, New York: Fischer.
- Calbris, Geneviève (1990): *The Semiotics of French Gesture: Advances in Semiotics*. Bloomington: Indiana University Press.
- Calbris, Geneviève (2003): From cutting an object to a clear cut analysis: Gesture as the representation of a preconceptual schema linking concrete actions to abstract notions. *Gesture* 3, 1, 19–46.
- Calbris, Geneviève (2008): From left to right. Coverbal gestures and their symbolic use of space. In: Cienki, Alan / Müller, Cornelia (Hg.), *Metaphor and gesture*. Amsterdam, Philadelphia: John Benjamins, 27–53.
- Calbris, Geneviève (2011): Elements of meaning in gesture. Amsterdam, Philadelphia: John Benjamins.
- Cienki, Alan (1998): Metaphoric gestures and some of their relations to verbal metaphoric expressions. In: Koenig, John-Pierre (Hg.), *Discourse and cognition: Bridging the gap*. Stanford: CSLI Publications, 189–204.
- Cienki, Alan (2013). Cognitive Linguistics. Spoken language and gesture as expressions of conceptualization. In: Müller, Cornelia / Cienki, Alan / Fricke, Ellen / Ladewig, Silvia H. / McNeill, David / Tessendorf, Sedinha (Hg.), *Body – Language – Communication. An international Handbook on Multimodality in Human Interaction. Handbooks of Linguistics and Communication Science* (38.1). Berlin, Boston: Mouton de Gruyter, 182–201.
- Cienki, Alan (2015): Spoken language usage events. *Language and Cognition*, 7, 4, 499–514.
- Cienki, Alan / Müller, Cornelia (2008): *Metaphor and gesture*. Amsterdam, Philadelphia: John Benjamins.
- Clark, Herbert H. (1996): Using language. Cambridge: Cambridge University Press.

- Clark, Herbert H. (2003): Pointing and placing. In: Kita, Sotaro (Hg.), *Pointing. Where language, culture, and cognition meet*. Mahwah, London: Lawrence Erlbaum Associates, 243–268.
- Coulson, Seana (2001): Semantic leaps. Frame-shifting and conceptual blending in meaning construction. Cambridge: Cambridge University Press.
- Gonzalez-Marquez, Monica / Mittelberg, Irene / Coulson, Seana / Spivey, Michael J. (Hg.) (2007): *Methods in Cognitive Linguistics*. Amsterdam, Philadelphia: John Benjamins.
- Croft, William / Cruse, Alan A. (2004): *Cognitive linguistics*. Cambridge: Cambridge University Press.
- Dancygier, Barbara / Sweetser, Eve E. (2005): Mental spaces in grammar. Conditional constructions. Cambridge: Cambridge University Press.
- Dancygier, Barbara / Sweetser, Eve E. (2012): Viewpoint in language. A multimodal perspective. Cambridge: Cambridge University Press.
- Dancygier, Barbara / Sweetser, Eve E. (2014): *Figurative language*. Cambridge: Cambridge University Press.
- Deppermann, Arnulf (2011): How does 'cognition' matter to the analysis of talk-in-interaction? In: *Language Sciences* 34, 6, 746–767.
- Deppermann, Arnulf / Schmitt, Reinhold (2007): Koordination. Zur Begründung eines neuen Forschungsgegenstand. In: Schmitt, Reinhold (Hg.), *Koordination. Analysen zur Multimodalen Interaktion*. Tübingen: Mouton de Gruyter, 15–54.
- Dirven, Rene / Pörings, Ralf (2002): Metaphor and Metonymy in Comparison and Contrast. Berlin, New York: Mouton de Gruyter.
- Downs, Roger M. / Stea, David (1977): *Maps in Mind. Reflections in Cognitive Mapping*. New York: Harper & Row.
- Downs, Roger M. / Stea, David (1982): *Kognitive Karten. Die Welt in unseren Köpfen*. New York: dtv.
- Du Bois, John W./ Schuetze-Coburn, Stephan / Cumming, Susanna / Paolino, Danae (1993): Outline of discourse transcription. In *Talking data: Transcription and coding in discourse research*. In: Edwards, Jane A. / Lampert Martin D. (Hg.) Hillsdale , NJ : Erlbaum, . 45–89.
- Ehmer, Oliver (2011): *Imagination und Animation. Die Herstellung mentaler Räume durch animierte Rede*. Berlin/New York: Mouton de Gruyter.
- Enfield, Nick (2003): Producing and editing diagrams using co-speech gesture. Spatializing non-spatial relations in explanations of kinship in Laos. In: *Journal of Linguistic Anthropology* 13, 7–50.
- Enfield, Nick (2009): *The Anatomy of Meaning. Speech, Gestures and Composite Utterances*. Cambridge: Cambridge University Press.
- Fauconnier, Gilles (1985): *Mental Spaces: Aspects of meaning construction in natural language*. Cambridge: MIT Press.
- Fauconnier, Gilles (1994): *Backstage Cognition. Mental spaces. Aspects of meaning construction in natural language*. New York: Cambridge University Press.
- Fauconnier, Gilles / Sweetser, Eve E. (1996): *Spaces, Worlds and Grammar*. Chicago: University of Chicago Press.
- Fauconnier, Gilles / Turner, Mark (2002): *The way we think*. New York: Basic Books.

- Fillmore, Charles J. (1977): Scenes-and-frames semantics. In: Zampolli, Antonio (Hg.), *Linguistic structures processing* 4. Amsterdam, New York, Oxford: North Holland
- Fillmore, Charles J. (1982): Frame semantics. In: Linguistic Society of Korea (Hg.), *Linguistics in the morning calm*. Seoul: Hanshin, 111–137.
- Fillmore, Charles J. (1985): Frames and the semantics of understanding. In: *Quarderni di Semantica* 6, 222–254.
- Fricke, Ellen (2007): *Origo, Geste und Raum*. Berlin: Mouton de Gruyter.
- Fricke, Ellen (2010): Phonaestheme, Kinaestheme und multimodale Grammatik. In: *Sprache und Literatur* 41, 69–88.
- Fricke, Ellen (2012): Grammatik multimodal. Wie Wörter und Gesten zusammenwirken. Berlin, New York: Mouton de Gruyter.
- Gerner, Alexander / Pombo, Olga (2010): Studies in Diagrammatology and Diagrammatic Praxis. In: *Studies in Logic, Logic and Cognitive Systems* 24. London: College Publications Rosch.
- Gibbs, Raymond W. Jr. (1994): The poetics of mind. Figurative thought, language, and understanding. Cambridge: Cambridge University Press
- Gibbs, Raymond W. Jr. (2006): Embodiment and Cognitive Science. New York: Cambridge University Press.
- Gibson, James (1977): The Theory of Affordances. In: Robert Shaw / John Bransford (Hg.), *Perceiving, Acting, and Knowing*. Hillsdale: Erlbaum.
- Goldberg, Adele (1995): Constructions. A construction grammar approach to argument structure. Chicago: University of Chicago.
- Goodman, Nelson (1968): Languages of Art: An Approach to a Theory of Symbols. London: Oxford University Press.
- Goodwin, Charles (2003): Conversation and brain damage. Oxford, New York: Oxford University Press.
- Goodwin, Charles (2007): Environmentally coupled gestures. In: Duncan, Susan / Cassell, Justine / Levy, Elena (Hg.), *Gesture and the dynamic dimensions of language*. Amsterdam/Philadelphia: John Benjamins, 195–212.
- Goodwin, Charles (2011): Contextures of action. In: Streeck, Jürgen / Goodwin, Charles / LeBaron, Curtis (Hg.), *Embodied interaction. Language and body in the material world*. Cambridge: Cambridge University Press, 182–193.
- Goodwin, Charles / LeBaron, Curtis / Streeck, Jürgen (2011): Embodied interaction. Language and the body in the material world. Cambridge: Cambridge University Press.
- Grandhi, Sukeshini / Joue, Gina / Mittelberg, Irene (2012): To Move or to Remove? A Human-Centric Approach to Understanding of Gesture Interpretation. Proceedings of the 10th ACM conference on Designing Interactive Systems, Newcastle.
- Mittelberg, Irene / Schmitz, Thomas H. / Groninger, Hannah (im Druck): Operative Manufakte. Gesten als unmittelbare Skizzen in frühen Stadien des Entwurfsprozesses. In: Hinterwaldner, Inge / Ammon, Sabine (Hg.), *Bildlichkeit im Zeitalter der Modellierung. Operative Artefakte in Entwurfsprozessen der Architektur und des Ingenieurwesens* (Eikones Reihe). München: Wilhelm Fink.
- Haiman, John (1980): Iconicity in Syntax. Amsterdam, Philadelphia: John Benjamins.

- Hassemer, Julius (2015): Towards a theory of Gesture Form Analysis: Principles of gesture conceptualisation, with empirical support from motion-capture data. Dissertationsschrift, RWTH Aachen University.
- Haviland, John (2000): Pointing, gesture spaces and mental maps. In: McNeill, David (Hg.), *Language and Gesture*. Cambridge: Cambridge University Press, 13–46.
- Hutchins, Edwin (1995): *Cognition in the Wild*. Cambridge: MIT Press.
- Hutchins, Edwin (2005): Material anchors for conceptual blends. In: *Journal of Pragmatics* 37, 1555–1577.
- Hutchins, Edwin / Nomura, Saeko (2011): Collaborative construction of multimodal utterances. In: Goodwin, Charles / LeBaron, Curtis / Streeck, Jürgen (Hg.), *Multimodality and Human activity. Research on Human Behavior, Action and Communication*. Cambridge: Cambridge University Press, 9–43.
- Imo, Wolfgang (2013): *Sprache in Interaktion. Analysemethoden und Untersuchungsfelder*. Berlin, Boston: De Gruyter.
- Jäger, Ludwig / Linz, Erika (Hg.) (2004): *Medialität und Mentalität. Theoretische und empirische Studien zum Verhältnis von Sprache, Subjektivität und Kognition*. München: Wilhelm Fink.
- Jakobson, Roman (1956): Two aspects of language and two types of aphasic disturbances. In: Monville-Burston, Monique / Waugh, Linda R. (Hg.), *Roman Jakobson. On Language*. Cambridge: Harvard University Press, 115–133.
- Jakobson, Roman (1960): Linguistics and poetics. In: Pomorska, Krystyna / Rudy, Stephen (Hg.), *Roman Jakobson – Language in Literature*. Cambridge, London: Harvard University Press, 62–94.
- Jakobson, Roman (1966): Quest for the essence of language. In: Monville-Burston, Monique / Waugh, Linda R. (Hg.), *Roman Jakobson - On Language*. Cambridge, London: Harvard University Press, 407–421.
- Jakobson, Roman (1971): Shifters, verbal categories, and the Russian verb. In: Roman Jakobson (Hg.), *Selected Writings, Volume II: Words and Language*. The Hague, Paris: Mouton, 130–147.
- Jakobson, Roman / Waugh, Linda R. (1979): *The Sound Shape of Language*. Bloomington: Indiana University Press.
- Johnson, Mark (1987): *The body in the mind. The bodily basis of meaning, imagination, and reason*. Chicago: University of Chicago Press.
- Johnson, Mark (2007): *The meaning of the body. Aesthetics of human understanding*. Chicago: University of Chicago Press.
- Johnson, Mark / Lakoff, George (1999): *Philosophy in the flesh. The embodied mind and its challenge to Western thought*. New York: Basic Books.
- Kendon, Adam (1990): *Conducting interaction: Patterns of behavior in focused encounters*. Cambridge: Cambridge University Press.
- Kendon, Adam (2004): *Gesture. Visible action as utterance*. Cambridge: Cambridge University Press.
- Kendon, Adam (2009): Kinesic components of multimodal utterances. In: *Proceedings of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society.
- Kita, Sotaro / van Gijn, Ingeborg / van der Hulst, Harry (1998): Movement Phases in signs and co-speech gestures, and their transcription by human

- coders. In: Wachsmuth, I. / Fröhlich, M. (Hg.), Gesture and sign language in human-computer interaction, Lecture Notes in Artificial Intelligence 1371. Berlin: Springer, 23–35.
- Kita, Sotaro (2003): Pointing. Where Language, Culture, and Cognition Meet. Mahwah: Lawrence Erlbaum Associates.
- Kramsch, Claire (Hg.) (2002): Language Acquisition and Language Socialization. Ecological perspectives. London: Continuum.
- Krois, M., John /Rosengren, Mats / Steidele, Angela / Westerkamp, Dirk (Hg.) (2007): Embodiment in Cognition and Culture. Amsterdam: John Benjamins.
- Labov, William (1972): The transformation of reality in narrative syntax. In: Labov, William (Hg.), Language in the inner city. Philadelphia: University of Pennsylvania Press, 354–396.
- Ladewig, Silva H. (2010): Beschreiben, auffordern und suchen. Varianten einer rekurrenten Geste. In: Sprache und Literatur 41/1, 89–111.
- Lakoff, George (1987): Women, fire, and dangerous things. What categories reveal about the mind. Chicago: University of Chicago Press.
- Lakoff, George (1993): The contemporary theory of metaphor. In: Ortony, Andrew (Hg.) Metaphor and thought. Cambridge: Cambridge University Press, 202–251.
- Lakoff, George / Johnson, Mark (1980): Metaphors We Live By. Chicago: Chicago University Press.
- Langacker, Ronald W. (1987): Foundations of cognitive grammar: Theoretical prerequisites (Vol. 1). Stanford: Stanford University Press.
- Langacker, Ronald W. (1991): Foundations of Cognitive Grammar, Vol. 2. Descriptive Application. Stanford: Stanford University Press.
- Langacker, Ronald W. (1993): Reference-point constructions. In: Cognitive Linguistics 4, 1–38.
- Liddell, Scott (2003): Grammar, gesture and meaning in American Sign Language. Cambridge: Cambridge University Press.
- McNeill, David (1992): Hand and mind. What gestures reveal about thought. Chicago: Chicago University Press.
- McNeill, David / Cassell, Justine / Levy, Elena (1993). Abstract deixis. *Semiotica* 95 1/2, 5-19.
- McNeill, David (2005): Gesture and thought. Chicago: Chicago University Press.
- Mittelberg, Irene (2006): Metaphor and metonymy in language and gesture. Discourse evidence for multimodal models of grammar. Ann Arbor: UMI.
- Mittelberg, Irene (2008): Peircean semiotics meets conceptual metaphor. Iconic modes in gestural representations of grammar. In: Cienki, Alan / Müller, Cornelia (Hg.), Metaphor and gesture. Amsterdam/Philadelphia: John Benjamins, 115–154.
- Mittelberg, Irene (2010a): Geometric and image-schematic patterns in gesture space. In: Evans, Vyvyan / Chilton, Paul (Hg.), 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. In: Sprache und Literatur, 41(1), 112–143.
- Mittelberg, Irene (2013a): The exbodied mind. Cognitive-semiotic principles as motivating forces in gesture. In: Müller, Cornelia / Cienki, Alan / Fricke, Elan /

- Ladewig, Silvia H. / McNeill, David / Tessendorf, Sedinha (Hg.), *Body - Language - Communication: An International Handbook on Multimodality in Human Interaction. Handbooks of Linguistics and Communication Science* (38.1). Berlin, Boston: Mouton de Gruyter, 750–779.
- Mittelberg, Irene (2013b): Balancing acts. Image schemas and force dynamics as experiential essence in pictures by Paul Klee and their gestural enactments. In: Bokrent, Michael / Dancygier, Barbara / Hinnell, Jennifer (Hg.), *Language and the creative mind*. Stanford: CSLI, 325–346.
- Mittelberg, Irene (2014). Gestures and iconicity. In: Müller, Cornelia / Bressem, Jana / Cienki, Alan / Fricke, Ellen / Ladewig, Silvia H. / McNeill, David / Bressem, Jana (Hg.), *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction. Handbooks of Linguistics and Communication Science* (38.2). Berlin, Boston: Mouton de Gruyter, 1712–1732.
- Mittelberg, Irene (erscheint): Embodied frames, metonymy and pragmatic inferencing in gesture. In: *Gesture* (Sonderheft, Gesture Pragmatics (Sweetser, Eve / Wehling, Elisabeth (Hg.).
- Mittelberg, Irene / Schüller, Daniel (erscheint): Kulturwissenschaftliche Orientierung in der Gestenforschung. In: Jäger, Ludwig / Holly, Werner / Krapp, Peter / Weber, Samuel (Hg.), *Language – Culture – Communication. An International Handbook of Linguistics as Cultural Study*. Berlin, New York: Mouton de Gruyter.
- Mittelberg, Irene / Waugh, Linda R. (2009): Metonymy first, metaphor second. A cognitive-semiotic approach to multimodal figures of thought in co-speech gesture. In: Forceville, Charles / Urios-Aparisi, Eduardo (Hg.), *Multimodal metaphor*. Berlin, New York: Mouton de Gruyter, 329–356.
- Mittelberg, Irene / Waugh, Linda R. (2014): Gestures and metonymy. In: Müller, Cornelia / Cienki, Alan / Fricke, Ellen / Ladewig, Silvia / McNeill, David / Bressem, Jana (Hg.), *Body– language–communication. An International Handbook on Multimodality in Human Interaction. Handbooks of Linguistics and Communication Science* (38.2). Berlin, Boston: Mouton de Gruyter, 1747–1766.
- Mondada, Lorenza (2011): The management of knowledge discrepancies and of epistemic changes in institutional interactions. In: Stivers, Tanya / Mondada, Lorenza / Steensig, Jacob (Hg.), *Knowledge and Morality in Conversation. Rights, Responsibilities and Accountability*. Cambridge: Cambridge University Press, 27–57.
- Mondada, Lorenza (2013): Embodied and spatial resources for turn-taking in institutional multi-party interactions: The example of participatory democracy debates. In: *Journal of Pragmatics* 46, 39–68.
- Müller, Cornelia (1998): *Redebegleitende Gesten: Kulturgeschichte – Theorie – Sprachvergleich*. Berlin: Spitz Verlag.
- Müller, Cornelia (2010): Wie Gesten bedeuten. Eine kognitiv-linguistische und sequenzanalytische Perspektive. In: *Sprache und Literatur* 41, 37–68.
- Müller, Cornelia / Bohle, Ulrike (2007): Das Fundament fokussierter Interaktion. Zur Vorbereitung und Herstellung von Interaktionsräumen durch körperliche Koordination. In: Schmitt, Reinhold (Hg.), *Koordination. Analyse zur multimodalen Kommunikation*. Tübingen: Narr, 129–165.

- Müller, Cornelia / Posner, Roland (Hg.) (2004): *The Semantics and Pragmatics of Everyday Gestures*. The Berlin Conference. Berlin: Weidler Verlag.
- Müller, Cornelia / Tag, Susanne (2010): The dynamics of metaphor. Foregrounding and activating metaphoricity in conversational interaction. In: *Cognitive Semiotics* 10, 85–120.
- Panther, Klaus-Uwe / Thornburg, Linda L. (2003): *Metonymy and pragmatic inferencing*. Amsterdam: John Benjamins.
- Parrill, Fey / Sweetser, Eve E. (2004): What we mean by meaning. Conceptual integration in gesture analysis and transcription. In: *Gesture* 4, 197–219.
- Peirce, Charles S. (1955): Logic as semiotic: The theory of signs. In: Bucher, Justus (Hg.) *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 (Hg.), Cambridge: Harvard University Press.
- Pombo, Olga / Gerner, Alexander (2010): Studies in Diagrammatology and Diagrammatic Praxis. In: *Studies in Logic, Logic and Cognitive Systems* 24. London: College Publications Rosch.
- Rekittke, Linn-Marlen / Mittelberg, Irene / Wolf, Dhana (2015): How sound moves the body. Felt qualities of experience in gestural enactments of film sound. In: *Proceedings of the 3rd Gesture and Speech in Interaction (GESPIN) Conference*, Sept. 2015, Nantes.
- Roth, Wolff-Michael (2003): From epistemic (ergotic) actions to scientific discourse. The bridging function of gestures. In: *Pragmatics & Cognition* 11, 141–170.
- Schank, Roger C. / Abelson, Robert (1977): *Scripts, Goals, Plans, and Understanding*. Hillsdale: Erlbaum.
- Schmitt, Reinhold (2007): Koordination. Analysen zur multimodalen Interaktion. Tübingen: Narr.
- Schoonjans, Steven / Sambre, Paul / Brône, Geert / Feyaerts, Kurt (erscheint): Vers une analyse multimodale du sens. Perspectives constructionnelles sur la gestualité co-grammaticale. Langages.
- Stjernfelt, Frederik (2007): Diagrammatology. An Investigation on the Borderlines of Phenomenology, Ontology and Semiotics. Dordrecht: Springer.
- Stokoe, William C. (2005/1960): Sign Language Structure: An Online of the Visual Communication Systems of the American Deaf. In: *Journal of Deaf Studies and Deaf Education* 10/1, 3-37.
- Streeck, Jürgen (2009): *Gesturecraft. The Manu-facture of meaning*. Amsterdam, Philadelphia: John Benjamins.
- Streeck, Jürgen (2013): Praxeology of Gesture. In: Müller, Cornelia / Cienki, Alan / Fricke, Ellen / Ladewig, Silvia H. / McNeill, David / Bressem, Jana (Hg.), *Body – Language – Communication. An International Handbook of Multimodality in Human Interaction* (38.1). Berlin, Boston: Mouton de Gruyter, 674–688.
- Streeck, Jürgen / Goodwin, Charles / LeBaron, Curtis D. (Hg.) (2011): *Embodied Interaction: Language and Body in the Material World: Learning in Doing: Social, Cognitive and Computational Perspectives*. New York: Cambridge University Press.

- Sweetser, Eve E. (1998): Regular metaphoricity in gesture. Bodily-based models of speech interaction. In: *Actes du 16e Congrès International des Linguistes* (CD-ROM), Elsevier.
- Sweetser, Eve E. (2007): Looking at space to study mental spaces. Co-speech gesture as a crucial data source in cognitive linguistics. In: Gonzalez-Marquez, Monica / Mittelberg, Irene / Coulson, Seana / Spivey, Michael (Hg.), *Methods in cognitive linguistics*. Amsterdam, Philadelphia: John Benjamins, 201–224.
- Sweetser, Eve E. (2012): Viewpoint and perspective in language and gesture. From the ground down. In: Dancygier, Barbara / Sweetser, Eve E. (Hg.), *Viewpoint in language. A multimodal perspective*. Cambridge: Cambridge University Press, 1–24.
- Talmy, Leonard (1988): Force dynamics in language and cognition. In: *Cognitive Science* 12, 49–100.
- Talmy, Leonard (2000): Toward a cognitive semantics, Vol. 1: Concept structuring systems. Cambridge: MIT Press.
- Talmy, Leonard (2013): Gestures as cues to a target. Talk presented at the ICLC 12, University of Edmonton, Alberta, June 24, 2013.
- Tomasello, Michael (1995): Joint attention as social cognition. In: Moore, Chris / Dunham, Philip (Hg.), *Joint attention: Its Origins and Role in Development*. Hillsdale, NJ: Lawrence Erlbaum, 103–130.
- Tomasello, Michael / Carpenter, Malinda (2007): Shared intentionality. In: *Developmental Science* 10, 1, 121–125.
- Tversky, Barbara (2011): Visualizing thought. In: *Topics in Cognitive Science* 3, 499–535.
- Wilcox, Sherman (2015): Signed languages. In: Divjak, Dagmar / Dabrowska, Ewa (Hg.), *Handbook of Cognitive Linguistics*. Berlin: Mouton de Gruyter.
- Williams, Robert F. (2008): Gesture as a conceptual mapping tool. In: Cienki, Alan / Müller, Cornelia (Hg.), *Metaphor and Gesture*. Amsterdam, Philadelphia: John Benjamins, 55–92.
- Wilkens, David. Why pointing with the index finger is not universal (in sociocultural and semiotic terms). In: Kita, Sotaro (Hg.), *Pointing: Where Language, Culture and Cognition Meet*. Mahwah, New Jersey: Lawrence Erlbaum and Associates.
- Ziem, Alexander (2008): Frames und sprachliches Wissen. Kognitive Aspekte der semantischen Kompetenz. Berlin, New York: de Gruyter
- Ziem, Alexander (2014): *Frames of Understanding in Text and Discourse*. Amsterdam, Philadelphia: John Benjamins.
- Zima, Elisabeth (2014): Gibt es multimodale Konstruktionen? Eine Studie zu [V(motion) in circles] und [all the way from X PREP Y]. *Gesprächsforschung* 15, 1–48.
- Zima, Elisabeth / Brône, Geert (2015): Cognitive Linguistics and interactional discourse: time to enter into dialogue, *Language and Cognition* 7, 4, 485–498.