

**Rezension zu: Oliver Ehmer: Imagination und Animation.  
Die Herstellung mentaler Räume durch animierte Rede.  
Berlin / New York: De Gruyter 2011**

**Elisabeth Zima**

In seiner Studie zur Imagination und Animation in Gesprächssequenzen setzt sich Ehmer ein ebenso herausforderndes wie erstrebenswertes Ziel: eine Synergie zwischen gesprächsanalytischen und kognitiv-linguistischen Konzepten und Methoden. Sein empirischer Fokus liegt dabei auf dem Phänomen der animierten Rede im Gespräch. Anhand einer großen Fülle sowohl deutsch- als auch spanischsprachiger Daten argumentiert Ehmer für eine Auffassung von animierter Rede als koordinierter, interaktionaler Prozess der Imagination, das heißt des gemeinsamen Aufbaus und der gemeinsamen Ausgestaltung mentaler Räume (*mental spaces*, Fauconnier 1994; 1997). Diesen prozessorientierten Zugang zum Aufbau mentaler Räume im Gespräch sieht er als notwendige interaktionale Online-Erweiterung der *mental space theory*, in der mentale Räume bis dato zumeist lediglich a posteriori, als fertiges Produkt im Fokus standen. Wie diese Räume sequenziell, durch intersubjektiv geteilte Imagination Gestalt annehmen, wurde hingegen nur selten unter die Lupe genommen (siehe jedoch Ansätze bei Pascual 2006 und Hougaard 2005). Genau an diesem Punkt setzt Ehmers Studie zur animierten Rede an, denn der Aufbau mentaler Räume im Gespräch setzt für ihn per definitionem geteilte Kognition (*shared cognition*, Clark/Brennan 1991) voraus und bedingt einen gemeinsamen Ausverhandlungs- beziehungsweise Gestaltungsprozess zwischen Interaktanten sowie das gemeinsame Erleben der imaginierten Szene.

Charakteristisch für die von Ehmer analysierten Sequenzen ist, dass Sprecher nicht einfach die Rede Anderer wiedergeben (*reported speech*), sondern dass sie *eine Figur animieren*. Mit anderen Worten, ein oder mehrere Sprecher schlüpfen in einer Gesprächssequenz in die Rolle (am Gespräch teilnehmender oder nicht teilnehmender) Anderer beziehungsweise sie animieren ihr eigenes Handeln in einer imaginierten Situation, das heißt nicht im Hier und Jetzt des Gesprächs, sondern innerhalb eines fiktiven oder zeitversetzten mentalen Raums. Dabei übernimmt der Sprecher "die Perspektive der animierten Figur und versetzt sich mental in die imaginierte Szene" (S.1). Animierte Rede impliziert demnach Inszenierung. Vier zentrale Fragestellungen leiten die Analyse (S.1f.):

- Welche Formen aufgeführter mentaler Räume sind in Gesprächen zu finden?
- Auf welche Weise werden Aufführungen hergestellt?
- Wie kann methodisch der Aufbau gemeinsamer Vorstellungen im Gespräch nachgewiesen werden?
- Welche Funktionen erfüllt die animierte Rede im Besonderen?

Nach einer kurzen Einleitung in die Fragestellungen und einigen grundsätzlichen Beobachtungen zum Gegenstandsbereich "gesprochene Sprache" in Kapitel 1, steigt Kapitel 2 direkt ins Thema ein und skizziert Ehmers Zugang zur "Imagination als konversationelle Herstellung mentaler Räume". Als philosophische und sprachwissenschaftliche Wurzeln, aus denen Ehmer seinen Ansatz entwickelt, werden Bühlers Konzepte der Wahrnehmungssituation, des Phantasmas und der

Origoverschiebung angeführt und kurz skizziert. Darüber hinaus werden Chafes Studien zu Bewusstseinsmodi vorgestellt, sowie die Konzepte des Wirklichkeitsbereichs und der Interaktionsmodalität, und schließlich Zugänge zum Phänomen der Wahrnehmung und der Performance eingeführt. Dieser theoretische Überblick bietet einen kompakten, gelungenen Einstieg in die Thematik der Imagination und der, wie Ehmer es nennt, "mentalenen Versetzung" im Gespräch. Als kleiner Kritikpunkt kann hier allerdings angemerkt werden, dass manchmal im Unklaren bleibt, warum Ehmer manche Studien für seine Arbeit als besonders relevant und daher an dieser Stelle beschreibenswert ansieht und andere nicht behandelt werden. So basiert etwa der kurze Überblick zur Frage der Bewusstseinsmodi lediglich auf den Arbeiten Wallace Chafes (2001 (sic; sollte auf Nachfrage 1994 sein) und lässt dabei eine schier unüberschaubare Fülle anderer Studien und Ansätze zu diesem Thema außen vor. Dies ist kein Plädoyer für einen unnötig aufgeblasenen Literaturüberblick – vielmehr ist Ehmer für seinen Fokus auf das Wesentliche und einen die gesamte Arbeit kennzeichnenden klaren Formulierungsstil zu gratulieren – doch wäre es beispielsweise an dieser Stelle ergänzend interessant gewesen, wie das Verhältnis von Bewusstsein(modi) und "shared cognition" oder auch zu der von Fauconnier (1994; 1997) im Kontext der Blendingtheorie entwickelten und auch von Ehmer selbst an anderer Stelle (S.38) angeführten "backstage cognition" gesehen wird.

Das folgende Kapitel 2.2. geht genauer auf die Kognitive Semantik im Allgemeinen und die Theorie der mentalen Räume im Speziellen ein. Dabei gelingt Ehmer eine sehr ausgewogene Darstellung, die sowohl Stärken (Erfassen der kognitiven Prozesse der Konzeptualisierung, Annahme semantischer Unterdeterminiertheit, der Fokus auf Bedeutung jenseits der Wortebene, Bedeutungskonstitution als mentaler Prozess) als auch Schwächen der kognitiv-linguistischen Ansätze thematisiert (zu starrer, scheinbar transferorientierter, nicht konsequent dialogischer Zugang zur interaktionalen Bedeutungskonstitution, Analysefokus auf Produkt anstatt des interaktionalen Entstehungsprozesses von Bedeutung, Vernachlässigung von Gesprächsdaten). Kapitel 2.3. widmet sich schließlich einer der zentralen Fragestellungen der Arbeit aus theoretischer Sicht, nämlich welche Rolle mentale Räume in der Interaktion spielen und wie sie von Interaktanten aufgebaut und ausgestaltet werden. Hier dringt Ehmer in ein weitgehend unerforschtes Feld ein, das sich tatsächlich nur aus einer synergetischen Perspektive heraus, die kognitive und gesprächsanalytische Theorie und Methodik zu vereinen sucht, erschließen lässt. Dabei gilt es inhärente Schwächen sowohl der kognitiven als auch der interaktionalen Ansätze auszugleichen, denn während die *mental space theory* bis dato nur in einem geringen Ausmaß auf natürliche, interaktionale Gesprächsdaten angewandt wurde (Hougaard 2005; Pascual 2006; Oakley/Hougaard 2008) und auch hier der inkrementelle, intersubjektive Aufbau von mentalen Räumen nicht im Mittelpunkt stand, orientiert sich die Konversationsanalyse – und in geringerem Maße die Interaktionale Linguistik – an der Analyse sprachlicher Performanz, ohne jedoch daraus Rückschlüsse auf mentale Prozesse zu ziehen beziehungsweise ziehen zu wollen. Es besteht deshalb ein weit reichendes Theorie- und Methodikdefizit im Bereich der interaktional-orientierten Kognitiven Linguistik beziehungsweise der kognitiv-orientierten Gesprächsforschung (siehe auch Zima, im Druck). Ehmer (S.31) argumentiert jedoch ungeachtet der bestehenden Defizite für die Möglichkeit einer fruchtbaren Synergie, die sich für

ihn dadurch ergibt, "dass die Konversationsanalyse durch die konsequente Verortung gemeinsam hergestellter Bedeutung in sowohl für die Beteiligten als auch für den Analysierenden beobachtbarem Handeln auch für die Analyse mentaler Räume geeignet ist. Ausgangspunkt ist die theoretische Prämisse der Konversationsanalyse, dass Sprecher sich einander fortwährend ihr aktuelles Verständnis der Interaktion und des vorangegangenen Gesprächsverlaufs signalisieren" (*display*, Sacks/Schegloff et al. 1974). Mit anderen Worten scheint sich Ehmer der auch von Wolfgang Imo (2011) vertretenen Position anzuschließen, wonach mentale Prozesse an sich zwar nicht greifbar sind, jedoch im Sprachgebrauch für den Analytiker nachvollziehbare Spuren hinterlassen: "Cognitions are not observable – their consequences are". Das einführende, theoretisch ausgerichtete Kapitel zum Begriff der Imagination schließt mit einer kurzen Übersicht zu möglichen Formen mentaler Räume im Gespräch (generisch, gnomisch, typisiert, negiert, potentiell, hypothetisch, konditional, entworfen und fiktional).

In Kapitel 3 wird sodann das Phänomen der animierten Rede theoretisch aufgearbeitet. Im Unterschied zur bloßen Redewiedergabe, die explizit nicht im Fokus von Ehmers Ansatz steht, zeichnet sich animierte Rede dadurch aus, dass

- eine Einbettung einer Figur in die eigene Äußerung stattfindet,
- diese Einbettung mit der Übernahme der Perspektive dieser Figur einhergeht,
- über die Animation dieser Figur eine ihr zugeschriebene Handlung demonstriert wird.

Gerade diesem letzten Punkt kommt in Ehmers Ansatz zentrale Bedeutung zu. Er basiert auf dem von Clark und Gerrig (1990) postulierten Unterschied der Modi "Demonstrieren" und "Beschreiben" (Symbolisieren bei Ehmer). Demonstriert ein Sprecher eine Handlung, so übernimmt er die Rolle der die Handlung ausführenden Person. Figur und Handlung werden als direkt wahrnehmbar präsentiert (S.77). Mit anderen Worten, Sprecher inszenieren eine Handlung/Szene/Figur und machen diese somit für ihre Gesprächsteilnehmer erlebbar. Da die demonstrierten Handlungen nicht im Hier-und-Jetzt der Gesprächssituation situiert sind, ist Demonstrieren für Ehmer (S.77) gleichbedeutend mit dem Eröffnen und der Ausgestaltung eines mentalen Raums. Symbolisiert ein Sprecher, so *beschreibt* er hingegen eine Handlung. Sprecher versetzen sich nicht in diese beschriebene Handlung und nehmen nicht die Perspektive handelnder Personen ein, das heißt sie stellen diese nicht dar, sondern sie wahren eine perspektivische Distanz. Mit Langacker (1987) könnte man auf einem Metaniveau von objektiver Konstruierung im Falle des Symbolisierens und subjektiver Konstruierung bei Demonstrationen sprechen. Im Laufe der Fallstudien verdeutlicht Ehmer diesen Unterschied immer wieder anhand seiner Gesprächsdaten. Neben eindeutigen Fällen gibt es dabei durchaus – wie nicht zuletzt Ehmer selbst anführt (S.435) – einige Fälle, in denen die Grenze zwischen Symbolisieren und Demonstrieren schwierig festzumachen ist. Ein Faktor, der der Markierung und somit der besseren Erkennbarkeit von Demonstrationen dienlich ist, ist der Grad an Expressivität. Demzufolge werden Redesequenzen leichter als animiert und somit als demonstrierend wahrgenommen, wenn sie entweder verbal, prosodisch oder gestisch expressiv gestaltet sind. Es ist Ehmer in diesem Zusammenhang sehr hoch anzurechnen, dass er in seinen Datenanalysen folglich auch konsequent multimodal

vorgeht. So werden etliche Beispiele zu verschiedenen Teilaspekten (Entwerfen einer Handlung (4.2), Fiktionalisieren (4.3), Imagination von Identitäten und Positionieren (5.3)) auch in Hinblick auf das Zusammenspiel von verbaler Demonstration und gestischer beziehungsweise prosodischer Animation hin untersucht. Die theoretischen Grundlagen dieser multimodalen Analysen werden in Kapitel 3.3 erarbeitet. Im Mittelpunkt steht dabei der Unterschied zwischen *grounded mental spaces* (Lidell 1995), denen Sprecher über Gestik eine physische Realität zuschreiben können, und *ungrounded mental spaces*, die losgelöst von dem physischen Hier-und-Jetzt der Interaktanten (dem *Real Space* bei Lidell beziehungsweise dem *Reality Space* bei Ehmer) konstruiert werden. Ehmer beschreibt die Möglichkeiten von Gestik, einen mentalen Raum im Hier-und-Jetzt zu verankern (*grounding*) folgendermaßen (S.91):

Mentale Räume sind als kognitive Konstrukte nicht der direkten Wahrnehmung zugänglich. Durch das *grounding* jedoch werden Elemente oder Vorgänge in mentalen Räumen mit Körperteilen oder Körperbewegungen assoziiert (Lidell 1995:23). Hiermit wird ihnen eine physische Repräsentation zugeschrieben und sie werden sinnlich wahrnehmbar. So ist es z.B. möglich, durch körperliche Zeigege-  
sten auf Elemente zu referieren, die sich in einem mentalen Raum befinden. Aber auch durch ikonisch-pantomimische Gesten können Aspekte von Vorgängen in einem mentalen Raum und die Handlungen von Figuren verkörpert werden.

Eine entscheidende Rolle kommt deshalb im Bereich der animierten Rede der "Deixis am Phantasma" zu (Bühler 1934; Stukenbrock 2012). Demonstratum der Zeigege-  
ste und Referenzobjekt fallen in diesem Fall nicht zusammen, sondern "zwischen der Identifizierung des perzeptorischen Zeigeziels und der Referenzherstellung innerhalb eines mentalen Raums liegt ein interpretativer Prozess, bei dem der perzeptive Verweisraum auf einen mentalen Referenzraum projiziert werden muss" (S.97).

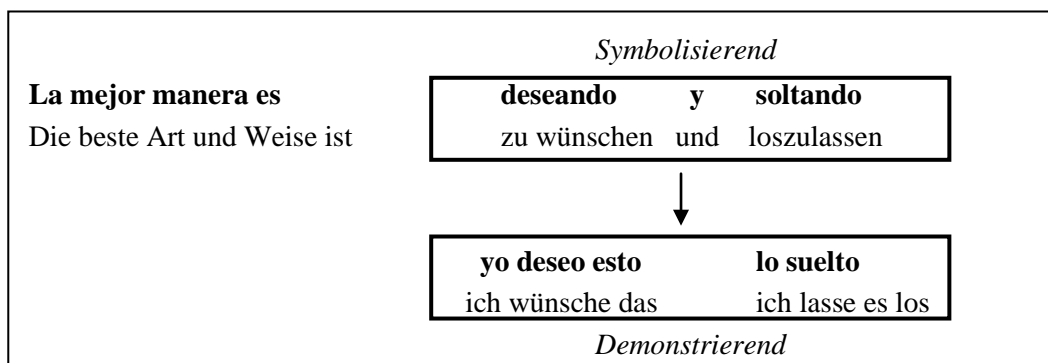
Nach dieser theoretischen Übersicht zu den für Ehmers Begriff der animierten Rede relevanten Konzepten und Unterscheidungen, wechselt Kapitel 3.4. zur Datenanalyse und illustriert den Unterschied zwischen Symbolisieren und Demonstrieren in Sequenzen, in denen lokal, unmittelbar aufeinander folgend, sprachliche Strukturen wiederholt werden und es zu einem Wechsel von einem Modus in den anderen kommt. Für die Kohärenz der Arbeit ist es dabei etwas abträglich, dass dieses Kapitel nicht direkt an die theoretischen Erläuterungen zu Clarks und Gerrigs Unterscheidung von Symbolisieren und Demonstrieren anschließt. Da jedoch in den Analysen in 3.4 auch multimodale Aspekte untersucht werden, ist der Einschub zur Rolle von körperlichen Mitteln zur Gestaltung mentaler Räume vor diesem datenbasierten Kapitel allerdings nachvollziehbar.

Anhand einiger spanischsprachiger Gesprächssequenzen arbeitet Ehmer heraus, wie Sprecher innerhalb kurzer Sequenzen, das heißt in unmittelbarer zeitlicher Abfolge, Handlungen zunächst symbolisierend beschreiben und danach demonstrierend inszenieren. (Auch die umgekehrte Reihenfolge, das heißt ein Wechsel vom Demonstrieren ins Symbolisieren ist möglich und wird belegt.) Eine entscheidende Rolle kommt dabei der Wiederholung von Konstruktionen zu. So zeigen Ehmers Sequenzen, dass Sprecher eine einmal etablierte sprachliche Struktur oft mehrmals wiederholen und im Zuge dessen Konstruktionen mal im symbolisierenden Modus instantiieren, mal mit animierter Rede kombinieren. Dabei bleibt der syntaktische Konstruktionsrahmen stabil und es werden teilweise auch

in animierten Redesequenzen Lexeme wiederverwendet, die davor im symbolisierenden Modus gebraucht wurden (S.122). Als Illustration dient das folgende Beispiel. Aus dem Transkript (S.104f.) werden lediglich jene Zeilen übernommen, in denen Konstruktionswiederholungen auftreten, über die der Wechsel vom symbolisierenden in den demonstrierenden Modus vollzogen wird.

- 05 G: que la mejOr maNEra (--) es solTANdo.  
dass die beste Art und Weise ist loszulassen (loslassend)
- 08 la (.) la mejor mane (--) es solTANdo  
die beste Weise ist loszulassen (loslassend)
- 10 <<all>> es diCIENdo>=  
ist zu sagen
- 11 =<<cv> y BUE::no;>  
na gut
- 12 y no SÉ  
und ich weiß nicht
- 13 me dA,  
es gibt mir
- 14 de-  
(von)
- 15 soltando pero nO sólo solTANdo.  
loslassend, aber nicht nur loslassend
- 16 deSEANdo,  
wünschend
- 17 y solTANdo.  
und loslassend
- 18 yo deseO ESTo.  
ich wünsche das
- 19 ((chasquea con los dedos))  
((schnalzt mit den Fingern))
- 20 lo SUELto.  
ich lasse es los

Anhand einer graphischen Darstellung in Tabellenform hebt Ehmer den Konstruktionsparallelismus hervor:



Ehmer argumentiert, dass der Wechsel vom Symbolisieren ins Demonstrieren sich hierbei dadurch vollzieht, dass die Sprecherin G die zuvor benannten Handlungen *zu wünschen und loszulassen* in *ich wünsche das (Fingerschnalzen) ich lasse es los* animiert. Sie animiert somit eine Figur, die eben diese Handlungen tatsächlich ausführt. Zusätzlich wird diese Animation auch über prosodische Mittel wie langsames Sprechtempo, Dehnung, geringe Sprechspannung etc. (vgl. S.106) inszeniert. Im Zuge der Sequenz werden außerdem immer wieder sprachliche Strukturen wiederaufgegriffen und variiert (*la mejor manera es soltando – la mejor mane es soltando – es diciendo – soltando pero no sólo soltando – deseando y soltando – yo deso esto – lo suelto*). An dieser Stelle hätte es sich aus meiner Sicht angeboten, Ehmers Beobachtungen (S.99-123) mit den Ergebnissen der Forschung zur Wiederholung im Gespräch (Tannen 1987; 1989; Norrick 1987 u.v.m.) und nicht zuletzt zur Resonanz und der Dialogischen Syntax (Du Bois 2001; 2007) zu vergleichen, denn der Wechsel von Ausdrucksmodi im Gespräch stellt eine Funktion der Wiederholung dar, die dort meines Wissens bisher noch nicht konstatiert wurde. Auch zum Konzept des *recycling with différence* und Per Linells (1998) Beobachtung, dass Wiederholungen einen Rekontextualisierungsprozess in Gang setzen, wobei sowohl Bedeutungen des ursprünglichen Kontext mitaktiviert und gleichzeitig aus diesem Kontext neue Bedeutungen geschaffen werden, ließen sich interessante Parallelen ziehen. Dass diese Einordnung an dieser Stelle nicht erfolgt, schmälert jedoch nicht den Wert der empirischen Beobachtung Ehmers, denn sie ist sowohl in Bezug auf die Frage der Art und Weise, wie mentale Räume im Gespräch entstehen, als auch für die Wiederholungs- beziehungsweise Resonanzforschung neu und relevant. Darüber hinaus erlaubt sie die Verbindung der traditionell eher funktional beziehungsweise konversationsanalytisch orientierten Wiederholungsforschung und dem kognitiven Ansatz der Mental Space Theory.

Das nächste Kapitel 4 bildet zusammen mit Kapitel 5 den empirischen Hauptteil der Arbeit. In Kapitel 4 stehen Prozesse des "Entwerfens" und des "Fiktionalisierens" im Mittelpunkt. Beide gelten nach Schütz/Luckmann (2003) als Spielarten des Phantasierens. Kapitel 5 beschäftigt sich mit dem "Evozieren mentaler Räume".

Kapitel 4.2. behandelt zunächst das Entwerfen von Handlungen. Charakteristisch für die unter diesem Aspekt analysierten Sequenzen ist, dass Sprecher eine *zukünftige* Handlung beziehungsweise Szene imaginieren. Ehmer zeigt diesbezüglich zunächst anhand eines Beispiels, wie sich etwa *mehrere Gesprächspartner* auf die Imagination einer Präsentation im Rahmen einer Prüfung einlassen und innerhalb des projizierten mentalen Raums (zukünftige Prüfungssituation), mögliche Szenarien durchspielen. Dabei wechseln Sprecher immer wieder zwischen symbolisierendem und animiertem Modus, wobei vor allem temporale Wechsel innerhalb des mentalen Raums symbolisiert, Handlungen dahingegen demonstriert werden (S.138). In einer folgenden Sequenz arbeitet Ehmer heraus, wie einem Akteur innerhalb des entworfenen mentalen Raums zwei real an der Gesprächssituation teilnehmende Personen zugewiesen werden. Dabei vollzieht sich der Wechsel vom an der Interaktion im Hier-und-Jetzt teilnehmenden Akteur zur Animation als handelnde Figur im mentalen Raum schrittweise. Ähnliches gilt auch für die nächste analysierte Sequenz, wobei hier spezifisch auch die Gestik animierender Sprecher in den Blick rückt und außerdem gezeigt wird, dass Spre-

cher zukünftige Handlungen gleichzeitig demonstrieren und negieren können und somit darlegen, wie sie eben nicht zu handeln gedenken.

Auch im anschließenden Kapitel zum "Fiktionalisieren" (4.3) steht das Negieren mentaler Räume im Vordergrund. Dabei werden Beispiele gezeigt, in denen Sprecher verdeutlichen, dass demonstrierte Handlungen nicht der Realität entsprechen (und dies auch nie werden). Zentrale Bedeutung kommt in den Analysen der Frame-Semantik (Fillmore 1977; 1982) und der Blending Theorie (Fauconnier/Turner 1996; 1998; 2003) zu. Es werden Sequenzen gezeigt, die sich dadurch auszeichnen, dass in interaktionaler Ausverhandlung fiktive mentale Räume aufgebaut werden, in denen inkompatible Frames miteinander überblendet werden. In einem ersten Schritt fokussiert Ehmer auf die Rolle des Zusammenspiels von Sprachstrukturen und phonetischer Realisierung beziehungsweise lautmalerischen Elementen im Kontext der Fiktionalisierung (4.3.2 und 4.3.3). In einem weiteren Schritt wird dieser Fokus auch auf Mimik und Gestik erweitert (4.3.4).

In 4.3.2 steht diesbezüglich zunächst im Vordergrund, wie Figuren in einen mentalen Raum eingeführt werden beziehungsweise inwiefern die bereits davor realisierte Etablierung eines Frames als geteiltes Hintergrundwissen der Interaktanten dazu dienen kann, ohne besondere Markierung (*space builder*, Fauconnier 1994) direkt vom symbolisierenden in den demonstrierenden Modus zu wechseln. Hierbei wird für eine animierte Äußerung vor dem Hintergrund zuvor etablierter Frames ein Äußerungskontext konstruiert (S.219). Ein Beispiel ist die deutschsprachige Sequenz "Fische" (S.187-191). Eine Gruppe von Bewohnern einer Wohngemeinschaft unterhält sich über die Luftverschmutzung in der ostdeutschen Stadt, in der sie leben, zu DDR-Zeiten. Einer der Bewohner ist Ostdeutscher, die anderen sind aus Westdeutschland. Dieses Wissen ist Teil des *common ground* (Clark 1996) der Gesprächsteilnehmer, der Frame DDR ist aktiviert als das Gespräch auf das Thema Wasserverschmutzung gebracht wird. Dem hier präsentierten Gesprächsausschnitt geht eine kurze Diskussion darüber voraus, dass vor einigen Jahrzehnten Flüsse sowohl in West- als auch in Ostdeutschland oft stark verschmutzt waren und kein wirkliches Umweltbewusstsein bestand. Die Sprecher T und H schlüpfen vor dem Hintergrund dieser etablierten Frames in die Rolle einer an einem Fluss stehenden und sich über Fische am Strand wundernden Person:

30 T: oh überall FISCHe?  
 31 (.)  
 32 S+T: ((lachen leise))  
 33 (.70)  
 34 T: schwimmen sie doch WEIter;  
 35 (.)  
 36 H+T ((lachen leise))  
 37 (.80)  
 38 H: dreh dich doch RUM;  
 39 alle ((lachen laut))  
 40 T: freiwillig am Ufer LOS-  
 41 T+H: ((lachen))  
 42 S: schwimm GLEICH-  
 43 alle ((lachen leiser))  
 44 T: wir ham Alles versucht.  
 45 H+C ((lachen sehr laut))  
 46 H: ham sie wieder ins wAsser geWORfen-  
 47 T: die (.) die wollten STRANden;  
 48 H+S: ((lachen leiser))  
 49 T: <<lachend> ch=weiß nicht waRUM;>

Während es sich in dieser Sequenz um eine nicht reale Person handelt, die animiert wird, zeigen andere Beispiele, wie konkrete Personen (vgl. die Sequenzen *stripper*, S.180-186, und *alumnos*, S.191-198) oder selbst real anwesende Personen animiert werden können. Im letzteren Fall geschieht dies zumeist auf einer frotzelnd-ironischen Ebene (siehe dazu auch den Begriff des *layering*, Clark 1996 und Brône 2008, der dieses ironische Layering mit Hilfe der Mental Space Theory modelliert). Ein weiteres Charakteristikum der hier analysierten Sequenzen (siehe auch die Sequenz "Fische") ist das Aufführen eines fiktiven Monologs oder Dialogs.

In 4.3.3 rückt die Blending Theorie in den Mittelpunkt. Das Kapitel gliedert sich nach der Art des Input Space 1. Zunächst (4.3.3.1) werden Fälle gezeigt, in denen dieser erste Input Space, der sodann mit einem weiteren Space überblendet wird, eine Erzählung beinhaltet. Das deiktische Zentrum dieses Input Spaces liegt folglich nicht im Hier-und-Jetzt der Gesprächssituation. Ehmer spricht in diesem Fall von *verschobenen Blends*. Danach wechselt der Fokus hin zu *Reality Blends* (4.3.3.2). In diesem Fall stellt das aktuelle Gespräch einen Ausgangspunkt für den Blend dar. Es wird mit einem teilweise inkompatiblen zweiten Input Space überblendet und somit fiktionalisiert. Real anwesende Gesprächsteilnehmer übernehmen entweder Rollen in diesem fiktionalen mentalen Raum oder eine zeitlich begrenzte Redesequenz wird auf ein nicht faktisches Niveau gehoben (vgl. auch hierzu wiederum Clarks (1996) Begriff des *Layering*). Der Effekt der Fiktionalisierung ist umso stärker, je inkompatibler die überblendeten Frames beziehungsweise mentalen Räume sind.

Ohne Zweifel gehören die in diesem empirischen Abschnitt präsentierten Analysen, die allesamt sehr detailliert einzelne Schritte im gemeinsamen Überblenden verschiedener Input Spaces (bis hin zur Entstehung von Mehrfachblends, vgl. das Prinzip des "running the blend" Fauconnier/Turner 1998) im sequenziellen Verlauf der Gesprächssequenzen unter die Lupe nehmen, zum Herzstück der Arbeit Ehmers. An dieser Stelle ist es unmöglich, die zahlreichen interessanten Detailbeobachtungen, die Ehmer dem Leser anhand seiner Beispiele offenbart, gebührend zu würdigen. Tatsächlich handelt es sich um minutiös vorangetriebene Analysen, die sich darum bemühen, keine post hoc Interpretation des Analytikers anzubieten, sondern sich stattdessen darauf zu konzentrieren, wie Interaktionsteilnehmer selbst Blends zu verstehen scheinen, das heißt, welche Frames sie aktivieren.

Bis zu einem gewissen Grad werden jedoch nicht zuletzt vermutlich einige Konversationsanalytiker der strikt nicht-kognitiven Ausprägung den Analysen Ehmers mit Skepsis begegnen. So ist anzumerken, dass es trotz der plausiblen Analysen Ehmers letztendlich keine Sicherheit geben kann, dass die beobachteten Sprecher tatsächlich die von Ehmer angeführten Frames aktivieren und die Blends auf die von ihm beschriebene Weise zustande kommen. Trotz der Verstehenssignale, die Sprecher geben (*display*), sei es verbal, sei es etwa durch Lachen, durch das Fortsetzen einer Frotzelaktivität etc., signalisieren Sprecher nicht explizit, wann Frames aktiviert werden, wie diese Frames im Wissensbestand der Sprecher strukturiert sind, und wie lange sie aktiviert bleiben (insbesondere in Bezug auf Sequenz 18 "zurückgemalt" (S.244-261) scheint mir Ehmers Analyse gar viele im Hintergrund aktivierte Frames zu enthalten). Aus meiner Sicht sind die Analysen – glücklicherweise! – hier deshalb auch nicht mehr gänzlich im Einklang mit der



Forderung der Konversationsanalyse, sich strikt an den Verständnissignalen, die Sprecher tatsächlich aussenden, zu orientieren, nur diese für die sequenzielle Interpretation heranzuziehen und somit lediglich objektiver Beobachter zu sein. Zwar formuliert Ehmer die "display"-Orientierung als eine der methodologischen Stärken der Konversationsanalyse (S.31, supra), die er in seinem Ansatz zur Anwendung bringen möchte, um kognitive Phänomene zu beleuchten, er geht aber in seinen Analysen darüber hinaus und gewährt uns einen Einblick in Gesprächssituationen, der sowohl Interaktanten- als auch Analytikerperspektive kombiniert. Dies ist auch notwendig, denn es geht hier um die Rekonstruktion von Bedeutungen und bedeutungskonstitutiven Prozessen. Die Annahme, diese würden sich dem beobachtenden Linguisten im und durch das Gespräch selbst offenbaren, ist unrealistisch und die Abwertung der interpretativen Leistung des Beobachters daher unangebracht (siehe auch Deppermanns (2000) Kritik am Primat des "displays" in der Konversationsanalyse).

Vertreter der Mental Space- und der Blending Theorie auf der anderen Seite sehen sich durch die Analysen Ehmers sicherlich bestätigt, denn auch wenn die Anwendung dieser Modelle und Theorien auf Gesprächsdaten bislang tatsächlich nur in einem sehr unzureichenden Maße erfolgt ist, bieten zum Beispiel die kognitiv-orientierte Humorforschung (Coulson 2000; 2005a/b; Brône/Feyaerts 2003; Brône 2008) ebenso wie die Kognitive Poetik (Brandt/Brandt 2005; Antonopoulou/Nikiforidou 2009; Semino 2009 u.v.m.) teilweise ähnliche Ansätze. Dabei wird jedoch nicht der inkrementelle Aufbau in einer längeren Diskurssequenz beleuchtet, sondern es werden lediglich kurze humoristische Sequenzen, Schlagzeilen, Cartoons beziehungsweise literarische Texte im Sinne des kreativen Blendings analysiert. (Eine allerdings nach dem Erscheinen Ehmers Buchs präsentierte Ausnahme ist Brône et al. 2012. Sie zeigen ebenfalls Blendingprozesse in einer längeren Diskurssequenz mit mehreren Teilnehmern.)

Besonders innovativ ist in diesem Zusammenhang auch das nächste Kapitel 4.3.4, denn es beschäftigt sich mit dem Einsatz multimodaler Mittel in Fiktionalisierungen und knüpft somit an zwei noch sehr junge Tendenzen innerhalb der Kognitiven Linguistik gleichermaßen an und verbindet diese: (1) den Trend zu einer kognitiv-linguistischen Gesprächsforschung und (2) zur multimodalen Studie von Bedeutung. Der Schwerpunkt Ehmers liegt diesbezüglich auf dem Zusammenspiel verbaler, prosodischer und körperlicher Mittel zur Animation von Figuren. Besondere Aufmerksamkeit gilt der Frage des Groundings eines fiktionalen mentalen Raums mittels körperlicher Ausdrucksmittel. Anhand zweier Ausschnitte aus einer deutschen Comedyshow ("Genial daneben") und einer weiteren Sequenz aus dem "Big Brother"-Korpus (der Lehrstühle Peter Auer und Susanne Günthner) werden zwei Arten des Einsatzes von Gestik herausgearbeitet (S.316):

Mit *constructed actions* bilden Sprecher ein Ereignis oder eine Handlung in einem mentalen Raum ab, indem sie Elemente des mentalen Raums mit ihrem Körper repräsentieren. Hierzu zählt insbesondere die pantomimische Übernahme der Rolle einer Figur. Zweitens wurden deiktische Gesten untersucht, mit denen Sprecher ausgehend von einer Figurenorigo auf mentale Objekte zeigen beziehungsweise referieren. Dabei besteht oft eine enge Verbindung von *constructed actions* und deiktischen Gesten, wenn Sprecher in der Rolle einer Figur agieren und auf Objekte zeigen.

Der zweite empirische Hauptteil der Arbeit (Kapitel 5) ist übertitelt mit "Evozierten mentaler Räume". Aus meiner Sicht ist dieser Titel allerdings nicht ganz optimal gewählt, denn es ist zunächst unklar, worin der Unterschied zu den in den Abschnitten davor gezeigten Analysen liegt beziehungsweise das "Evozieren" und das "Entwerfen" mentaler Räume erscheinen sich intuitiv wenig voneinander zu unterscheiden. Ehmers Unterscheidung beruht auf der Frage des Interaktionsgrades und des kollaborativen Charakters einer Imagination. Während in den Analysen in Kapitel 4 gemeinsame Imaginationen als Beispiel geteilter Kognition im Mittelpunkt standen, verschiebt sich in Kapitel 5 der Fokus hin zu Beispielen, in denen nur ein Sprecher mittels einer animierten Redesequenz einen mentalen Raum aufruft (evoziert). Ebenfalls in Kontrast zu Kapitel 4 wird dabei eine stärker grammatische Perspektive eingenommen und fokussiert, wie bestimmte Konstruktionen dazu eingesetzt werden, mentale Räume zu öffnen und zu strukturieren beziehungsweise wie konstruktive Leerstellen mit animierter Rede gefüllt werden. Kapitel 5.2. beschäftigt sich in diesem Zusammenhang mit der *X de Y-Konstruktion*, in der ein Nomen an der X-Stelle ein über eine animierte Redesequenz an der Y-Stelle näher bestimmt wird.

In Kapitel 5.3. stehen die Imagination von Identitäten und die Möglichkeiten zur Positionierung durch animierte Rede im Mittelpunkt. Es wird gezeigt, wie entweder ein sozialer Stereotyp beziehungsweise eine Personenkategorie oder eine reale Person in einer fiktionalen Szene animiert wird. Dabei illustriert Ehmer bezugnehmend auf das Konzept von Harvey Sacks (1972; 1979) des *membership categorization device*, wie Sprecher mittels animierter Rede Mitgliedschaftskategorien über das Demonstrieren von als typisch für eine bestimmte soziale Gruppe angenommene Handlungen konstruieren. Diese Technik ist besonders dann effektiv, wenn für das Verhalten eines sozialen Typus keine geeignete Kategorienbezeichnung besteht, das heißt, das Verhalten leichter demonstriert als benannt werden kann. Im Zuge der Demonstration positionieren sich Sprecher gegenüber dem demonstrierten Typus. In den von Ehmer gezeigten Beispielen handelt es sich dabei immer um distanzierende Positionierungen. Das heißt, die Sprecher konstruieren ihre eigene Identität im Kontrast zum demonstrierten sozialen Typus. Diese Art der Identitätskonstitution mittels Kontrastierens kann auch über negierte animierte Äußerungen vollzogen werden. In diesen Fall werden Personen bestimmte Handlungen beziehungsweise Handlungsdispositionen dezidiert abgesprochen. Besonders detailreich und vor allem auch aus einem multimodalen Blickwinkel heraus analysiert Ehmer das Beispiel eines Tangolehrers, der sich explizit nicht als allwissender und deshalb autoritärer Lehrer sieht, sondern einen gemeinschaftlichen Lernprozess zwischen Schülern und ihm anstrebt. Dabei argumentiert Ehmer, dass der Sprecher, das heißt der Tangolehrer, sowohl auf der verbalen als auch auf der phonetischen und körperlichen Ebene im Laufe der Sequenz eine zunehmend intensiviertere negierte Animation ausführt um sich von dem klassischen Bild des autoritären Tangolehrers abzusetzen. Während die Analyse des verbalen Kanals durchwegs einleuchtend erscheint, bleibt die körperliche und prosodische Analyse jedoch hier schwierig nachvollziehbar. Dies liegt nicht zuletzt an der Präsentation von Standbildern, bei denen jedoch nicht klar ist, in welchem Zeitabstand sie genommen wurden beziehungsweise ob sie jeweils den Höhepunkt der Geste abbilden. Zwar sind alle Standbilder mit Ausschnitten aus dem Transkript unterlegt, eine exakte zeitliche Zuordnung von Gestik, Mimik und Körperhaltung

und dem verbalen Kanal ist jedoch nicht möglich. Die Schwierigkeiten einer sinnvollen Transkription von körperlichen Ausdrucksmitteln betreffen jedoch selbstverständlich nicht nur diese Arbeit (für multimodale Transkriptionsverfahren siehe jedoch u.a. die Vorschläge in McNeill 1992; Kendon 2004; Stukenbrock 2009) und es kann Ehmer auch keinesfalls angelastet werden, nicht auch noch dieses komplexe Thema der Transkription von körperlichen Signalen gelöst zu haben. Eine Bereicherung der Arbeit wäre jedoch – wobei hier natürlich nicht zuletzt der Verlag entscheidet – die Integration der Beispiele inklusive Videosequenzen auf einer beiliegenden CD gewesen.

Zum Abschluss dieses Kapitels und damit der Studien Ehmers verschiebt sich der Fokus parallel zu Kapitel 5.2. noch einmal auf die konstruktionselle Ebene. Dabei werden zwei Konstruktionen in Bezug auf ihre Leistung im Bereich der animierten Rede untersucht, die bereits von den Arbeiten Karin Birkners (2006; 2008) bekannt sind: die *Ich bin ein Mensch, der X-Konstruktion* und die *Es gibt Leute, die X-Konstruktion*.

## Fazit

Oliver Ehmers Studie zum Phänomen der animierten Rede im Gespräch verknüpft mit der Gesprächsforschung beziehungsweise der Interaktionalen Linguistik einerseits und der Theorie der mentalen Räume andererseits zwei Forschungsbereiche, deren Berührungspunkte bis dato nur in sehr geringem Ausmaß erforscht wurden. Dabei beweist Ehmer in beiden Bereichen Trittfestigkeit, zieht eine ganze Fülle an theoretischen Konzepten heran und – dies ist das Entscheidende – es gelingt ihm, deren Relevanz für die Online-Bedeutungskonstitution anhand einer eindrucksvollen Menge an empirischen Belegen herauszuarbeiten. Gerade die Ausdauer, mit der Ehmer auf über 400 Seiten nicht weniger als 43 zumeist recht komplexe Diskurssequenzen analysiert und dabei regelrecht seziert, ist beeindruckend, auch wenn sie den Leser an mancher Stelle dazu verleiten könnten, die eine oder andere Abkürzung zu nehmen. Doch Ehmers Voraussicht sei Dank, werden auch die von Zeitnot und Aufmerksamkeitsdefiziten geplagten Leser stets auf den richtigen Weg zurückgeführt, denn mit zu den größten Stärken, die Ehmers Buch Struktur geben und nicht zuletzt zu seiner allgemein sehr guten Lesbarkeit beitragen, gehören die vielfachen Synthesen am Ende von Teilabschnitten und Kapiteln. In ihnen werden die wichtigsten Erkenntnisse und Ergebnisse noch einmal auf den Punkt gebracht und Verknüpfungen zu vorangegangenen und folgenden Teilstücken hergestellt. Es ist Ehmer hoch anzurechnen, dass er sich diese Mühe macht und dies auch konsequent bis zum Schluss durchhält, gehören gerade diese Synthesen bisweilen doch zu den unbeliebtesten Fingerübungen der meisten Autoren.

Ehmers Arbeit kann also insgesamt nur als rundum gelungen bezeichnet werden, denn sie ist ebenso innovativ wie methodologisch sorgfältig und darüber hinaus sehr lesbar. Mit einem gewissen Bedauern muss allerdings angemerkt werden, dass das Buch nicht auf Englisch verfasst wurde und somit leider nicht die internationale Aufmerksamkeit bekommen wird, die es ohne Zweifel verdienen würde. Dies ist nicht zuletzt deswegen sehr bedauerlich, da gerade der eher nicht deutschsprachigen Mainstream-Ausrichtung der Mental Spaces- und Blending Theorie ein stärker interaktionaler Online-Zugang sehr gut täte.

## Literatur

- Antonopoulou, Eleni / Nikiforidou, Kiki (2009): Deconstructing Verbal Humor with Construction Grammar. In: Brône, Geert / Vandaele, Jeroen (Hg.), *Cognitive Poetics*. Berlin: Mouton de Gruyter, 289-314.
- Birkner, Karin (2006): (Relativ-)Konstruktionen zur Personenattribuierung: ich bin n=mensch der... In: Günthner, Susanne / Imo, Wolfgang (Hg.), *Konstruktionen in der Interaktion*. Berlin: Mouton de Gruyter, 205-238.
- Birkner, Karin (2008): *Relativ(satz)konstruktionen im gesprochenen Deutsch. Syntaktische, prosodische, semantische und pragmatische Aspekte*. Berlin: Mouton de Gruyter.
- Brandt, Line / Brandt, Per Aage (2005): Making Sense of a Blend. In: *Annual Review of Cognitive Linguistics* 3, 216-249.
- Brône, Geert (2008): Hyper- and misunderstanding in interactional humor. In: *Journal of Pragmatics* 40, 2027-2061.
- Brône, Geert / Feyaerts, Kurt (2003): The cognitive linguistics of incongruity resolution: Marked reference-point structures in humor. KU Leuven: Department of Linguistics Preprint, 205.
- Brône, Geert / Oben, Bert / Sambre, Paul / Feayerts, Kurt (2012): Creative space-building in (inter)action. The case of joint fantasy. Präsentation bei CSDL 2012, Vancouver, 17.-20.5.2012.
- Bühler, Karl (1934): *Sprachtheorie: Die Darstellungsfunktion der Sprache*. Jena: Fischer.
- Chafe, Wallace (1994): *Discourse, consciousness, and time: the flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.
- Clark, Herbert H. (1996): *Using language*. Cambridge: Cambridge University Press.
- Clark, Herbert / Brennan, Susan (1991): Grounding in communication. In: Resnick, Lauren B. / Levine, John M. / Teasley, Stephanie D. (Hg.), *Perspectives on socially shared cognition*. Washington: APA Books.
- Clark, Herbert / Gerrig, Richard J. (1990): Quotations as demonstrations. In: *Language* 66, 764-805.
- Coulson, Seana (2000): *Semantic leaps. Frame-shifting and conceptual blending in meaning construction*. Cambridge: Cambridge University Press.
- Coulson, Seana (2005a): Extemporaneous blending: Conceptual integration in humorous discourse from talk radio. In: *Style* 39, 2, 107-122.
- Coulson, Seana (2005b): What's so funny? Cognitive semantics and jokes. In: *Cognitive Psychopathology/Psicopatologia cognitive* 2, 3, 67-78.
- Deppermann, Arnulf (2000): Ethnographische Gesprächsanalyse: Zu Nutzen und Notwendigkeit von Ethnographie für die Konversationsanalyse. In: *Gesprächsforschung Online – Zeitschrift zur verbalen Interaktion* 1, 96-124.  
URL: <http://www.gespraechsforschung-ozs.de/heft2000/ga-deppermann.pdf> [Letzter Zugriff 30.10.2012].
- Du Bois, John. W. (2001): *Towards a dialogic syntax*. University of Santa Barbara: Manuskript.

- Du Bois, John W. (2007): The stance triangle. In: Englebretson, Robert (Hg.), *Stancetaking in discourse: subjectivity, evaluation, interaction*. Benjamins: Amsterdam/Philadelphia, 139-182.
- Fauconnier, Gilles (1994): *Mental spaces: aspects of meaning construction in natural language*. Cambridge: Cambridge University Press.
- Fauconnier, Gilles (1997): *Mappings in thought and language*. Cambridge: Cambridge University Press.
- Fauconnier, Gilles / Turner, Mark (1996): Blending as a Central Process of Grammar. In: Goldberg, Adele E. (Hg.), *Conceptual Structure, Discourse, and Language*. Stanford: Center for the Study of Language and Information, 113-130.
- Fauconnier, Gilles / Turner, Mark (1998): Conceptual Integration Networks. In: *Cognitive Science* 22, 2, 133-187.
- Fauconnier, Gilles / Turner, Mark (2003): *The way we think. Conceptual blending and the mind's hidden complexities*. New York: Basic Books.
- Fillmore, Charles J. (1977): Scenes-and-frames semantics. In: Zampolli, Antonio (Hg.), *Linguistic Structures Processing*. Amsterdam: North-Holland, 55-81.
- Fillmore, Charles J. (1982): Frame semantics. In: *The Linguistic Society of Korea (Hg.), Linguistics in the Morning Calm*. Seoul: Hanshin, 111-37.
- Hougaard, Anders (2005): *How're we doing: An Interactional Approach to Cognitive Processes of Online Meaning Construction*. Doctoral dissertation. SDU, Odense, Denmark.
- Imo, Wolfgang (2011): Cognitions are not observable – but their consequences are: mögliche Aposiopese-Konstruktionen in der gesprochenen Alltagssprache. In: *Gesprächsforschung* 12, 265-300.  
URL: <http://www.gespraechsforschung-ozs.de/heft2011/ga-imo.pdf>  
[Letzter Zugriff: 19.09.2012].
- Kendon, Adam (2004): *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.
- Langacker, Ronald W. (1987): *Foundations of cognitive grammar: theoretical prerequisites*. Stanford: Stanford University Press.
- Lidell, Scott K. (1995): Real, Surrogate, and Token Space. In: Emmorey, Karen / Reilly, Judy S. (Hg.), *Language, Gesture, and Space*. Hillsdale: Erlbaum, 19-41.
- Linell, Per (1998): *Approaching dialogue: talk, interaction and contexts in dialogical perspectives*. Amsterdam;Philadelphia: Benjamins.
- McNeill, David (1992): *Hand and Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press.
- Norricks, Neal R. (1987): Functions of repetition in conversation. In: *Text* 7, 245-264.
- Oakley, Todd / Hougaard, Anders (2008): *Mental Spaces in Discourse and Interaction*. Amsterdam;Philadelphia: Benjamins.
- Pascual, Esther (2006): Fictive interaction within the sentence: A communicative type of fictivity in grammar. In: *Cognitive Linguistics* 17 (2), 245-267.
- Sacks, Harvey (1972): On the analyzability of stories by children. In: Gumperz, John / Hymes, Dell (Hg.), *Directions in sociolinguistics. The ethnography of speaking*. Basil Blackwell: New York, 325-345.

- Sacks, Harvey (1979): Hotrodder: a revolutionary category. In: Psathas, George (Hg.), *Everyday Languages: Studies in Ethnomethodology*. New York: Irvington, 7-14.
- Sacks, Harvey / Schegloff, Emanuel A. / Jefferson, Gail (1974): A simplest systematics for the organization of turn-taking for conversation. In: *Language*, 50, 696-735.
- Schütz, Alfred / Luckmann, Thomas (2003): *Strukturen der Lebenswelt*. Konstanz: UVK.
- Semino, Elena (2009): Text worlds. In: Brône, Geert / Vandaele, Jeroen (Hg.), *Cognitive Poetics*. Berlin: Mouton de Gruyter, 33-71.
- Stukenbrock, Anja (2009): Herausforderungen der multimodalen Transkription. Methodische und theoretische Überlegungen aus der wissenschaftlichen Praxis. In: Birkner, Karin / Stukenbrock, Anja (Hg.), *Die Arbeit mit Transkripten in Fortbildung, Lehre und Forschung*. Mannheim: Verlag für Gesprächsforschung, 144-170.
- Stukenbrock, Anja (2012): *Deixis in der Interaktion*. Habilitationsschrift, Albert-Ludwigs-Universität Freiburg.
- Tannen, Deborah (1987a): Repetition in discourse: towards a poetics of talk. In: *Language* 63, 574-605.
- Tannen, Deborah (1989 [2007]): *Talking voices: repetition, dialogue, and imagery in conversational discourse*. Cambridge: Cambridge University Press.
- Zima, Elisabeth (im Druck): *Kognition in der Interaktion. Eine kognitionslinguistische und gesprächsanalytische Untersuchung von Zwischenrufsequenzen in österreichischen Parlamentsdebatten*. Heidelberg: Universitätsverlag Winter.

Dr. Elisabeth Zima  
Albert-Ludwigs-Universität Freiburg  
Freiburg Institute for Advanced Studies (FRIAS)  
Albertstraße 19  
79104 Freiburg

[elisabeth.zima@frias.uni-freiburg.de](mailto:elisabeth.zima@frias.uni-freiburg.de)

Veröffentlicht am 21.2.2013

© Copyright by GESPRÄCHSFORSCHUNG. Alle Rechte vorbehalten.