

LANGUAGE, COGNITION, INTERACTION.
CONCEPTUAL BLENDING AS DISCURSIVE PRACTICE

Inaugural-Dissertation
zur
Erlangung des Doktorgrades
der Philosophie des Fachbereichs 05
der Justus-Liebig-Universität Gießen

vorgelegt von
Vera Stadelmann
aus Innsbruck

Gießen 2012

<u>Dekan:</u>	Prof. Dr. Magnus Huber
<u>1. Berichterstatter:</u>	Prof. Dr. Joybrato Mukherjee
<u>2. Berichterstatter:</u>	Prof. Dr. Helmuth Feilke

TABLE OF CONTENTS

Erklärung.....	iv
Acknowledgments	v
List of abbreviations and acronyms	vi
List of tables	vii
List of figures	viii
List of transcripts	x
Transcription conventions	xi
 1. LANGUAGE, COGNITION, INTERACTION. CONCEPTUAL BLENDING IN A SOCIAL-INTERACTIONAL COGNITIVE LINGUISTICS	 1
1.1. Introduction	1
1.2. Conceptual blending as discursive practice: Preview	5
 2. MENTAL SPACES AND CONCEPTUAL INTEGRATION THEORIES	8
2.1. Mental Spaces Theory	9
2.2. Conceptual Integration Theory	16
2.2.1. Constitutive principles	19
2.2.2. Governing principles	22
2.2.3. Blending typology	25
2.3. Critical discussion	28
2.3.1. Post-hoc vs. ad-hoc	29
2.3.2. The Generic Space	29
2.3.3. Generalisation, terminology and falsifiability	30
2.3.4. Delineating Mental Spaces	32
2.3.5. Whose Mental Spaces? Whose blends?	33
2.3.6. Loci and levels of mappings and blending	33
2.3.7. Homunculus entering through the stage door	35
2.3.8. Blending in context, blending in interaction	35
2.3.9. Blending typology	37
2.4. First conclusions	40
 3. PUTTING THEORY TO THE TEST IN INTERACTIONAL COGNITIVE SEMANTICS	 42
3.1. The State-of-the-Art of empirical and experimental research within MSCi .	42
3.1.1. Discursive approaches to MSCi	44
3.1.2. EMCA-inspired approaches to MSCi	46
3.1.3. Studies on blending taxonomies	48
3.2. Methods for a socially grounded MSCi: Conversation Analysis	49
3.2.1. CA: Historical development and philosophical underpinnings	51
3.2.2. CA: Key assumptions and concepts	55
3.3. A Social-Interactional Approach to MSCi: Preliminary hypotheses on conceptual reconfigurations	62
3.3.1. Cognition as centre-stage rather than back-stage	64
3.3.2. Common Ground and Mental Spaces	65
3.3.3. Coordination and Mental Spaces	67

3.3.4. Mental Spaces as interactionally relevant and phenomenologically evident	68
3.3.5. Current Discourse Space as a basis for communication	69
3.3.6. Blending analyses in a social-interactional Cognitive Semantics ...	71
3.4. Data collection and research design	73
3.4.1. Domain for study: Impersonation humour	73
3.4.1.1. Polyphonic narrative, double-voiced discourse and impersonations	74
3.4.1.2. Performance and theat(ic)ality in impersonations	77
3.4.1.3. Impersonations, play and humour	81
3.4.2. Impersonations as conceptual blends	86
3.4.3. Data source	90
3.4.3.1. Studying interaction-as-broadcasted in <i>Never Mind The Buzzcocks</i>	90
3.4.3.2. Transcription	95
3.4.4. Data Analysis: Process	98
3.4.4.1. Cognitive-semantic analysis: General principles	99
3.4.4.2. Interactional analysis: General principles	100
4. IMPERSONATION BLENDS: CASE STUDIES	103
4.1. The Sports Commentator: A Mirror Blend Impersonation	103
4.1.1. Interactional organisation guides conceptualisation processes	105
4.1.2. Signalling the blend: Changing voices	110
4.2. The Proprietor: A Single-Scope Impersonation Blend	111
4.2.1. Belief space and narrative space in impersonation prefaces	113
4.2.2. From preface to performance	115
4.2.3. Frame structuring in The Proprietor	117
4.3. The Metal Detector: A Double-Scope Impersonation	119
4.3.1. Interactional organisation: Negotiating viewing frames	120
4.3.2. 'Zooming in' on a narrative	121
4.3.3. Frame structuring in double-scope impersonations	123
4.4. The Way We Think, The Way We Act	125
5. INTERACTIONAL ANALYSIS OF IMPERSONATION BLENDS	127
5.1. Local occasioning and diegetic prefaces	128
5.1.1. Local occasioning and preface formats in mirror impersonations .	129
5.1.2. Local occasioning and preface formats in single-scope impersonations	136
5.1.3. Local occasioning and preface formats in double-scope impersonations	141
5.1.4. Discussion	143
5.2. Demarcation markers: Parsing the diegetic from the mimetic	145
5.2.1. Quotative systems	146
5.2.1.1. Quotatives in mirror impersonations	151
5.2.1.2. Quotatives in single-scope impersonations	155
5.2.1.3. Quotatives in double-scope impersonations	161
5.2.1.4. Discussion of findings	164
5.2.2. Embodied demarcation markers: Gaze and postural orientation ...	165
5.2.2.1. Gaze and postural positioning in mirror impersonations	168
5.2.2.2. Gaze and postural positioning in single-scope	

impersonations	172
5.2.2.3. Gaze and postural positioning in double-scope impersonations	177
5.2.2.4. Discussion	178
5.2.3. Conclusion	184
5.3. Designing performances: Floor management, chunking and listing	185
5.3.1. Floor management and performance design in mirror impersonations	188
5.3.2. Floor management and performance design in single-scope impersonations	196
5.3.3. Floor management and performance design in double-scope impersonations	200
5.3.4. Discussion	201
5.4. Ending and exiting impersonation sequences	202
5.4.1. Repeats	203
5.4.2. Punch line-mode	204
5.4.3. Coda	207
5.4.4. Laugh particle transitions	208
5.4.5. Practices for reengaging with topical talk	210
5.4.6. Discussion	213
5.5. An action-based typology of impersonation blends	215
 6. THE PRIMACY OF PRAGMATICS	 218
6.1. Conceptual blending in interaction: A discursively grounded model of in- situ meaning coordination	 219
6.2. Project summary	225
6.3. Prospects for further research	229
 References	 232
Appendix	254
A1 German summary	255

ERKLÄRUNG

Ich erkläre:

Ich habe die vorgelegte Dissertation selbständig und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht.

Vera Stadelmann

Gießen, den 7. März 2012

ACKNOWLEDGMENTS

This work took shape as a result of discussions with many highly inspirational people. First and foremost, I'm deeply indebted to the faculty, staff and students in the Department for English and the International Graduate Centre for the Study of Culture at Justus Liebig University Giessen. I owe a great debt of gratitude to my advisors Joybrato Mukherjee and Helmuth Feilke for their valuable feedback, encouragement and continuous support. Rosemary Bock, Christiane Brand, Thorsten Brato, Sandra Götz, Christopher Koch, Svetla Rogatcheva and Marco Schilk not only provided their linguistic expertise – thank you for giving me such a warm welcome to the team and helping me to navigate my way through the department!

At the GCSC, I would like to express my most profound gratitude to Ansgar Nünning for getting me to Giessen in the first place. Doris Bachmann-Medick, Julia Faisst, Alexander Friedrich, Gero Guttzeit, Beatrice Michaelis, Martin Zierold, and Ann Van de Veire, to name but a few, provided the most stimulating of environments one could hope for.

At Innsbruck and Sheffield, I'd like to thank Manfred Markus for inviting me to the linguistic train way back when, and Alexander Onysko, Martin Peart, Christian Quendler and Peter Thompson for their valuable comments, friendship and inspiration throughout the years. Grant Price was an expert proofreader.

For the opportunity to carry out several research stays and engage in spirited discussions, I'd like to thank the faculty and staff at Syddansk Universitet Odense and the 2009 Linguistic Institute at UC Berkeley. I owe a huge amount to Anders Hougaard, Gitte Rasmussen and the entire SoPraCon team, who were of immeasurable help in the early phase of the project. Mange tak! I'm also very grateful to Katharina König and Jörg Bücker for welcoming me to their colloquia at Münster's CeSI, and providing expert feedback on methodology and theoretical aspects.

Finally, I'm hugely indebted to my parents and friends, who accompanied the project with great patience and humour. Kai Sicks with his never-ceasing understanding and empathy took the lion's share, and in countless heated discussions helped me to further key ideas.

All infelicities remaining are, of course, no one's but my own.

LIST OF ABBREVIATIONS AND ACRONYMS

BBC	British Broadcasting Cooperation
CA	Conversation Analysis
CDS	Current Discourse Space
CEO	Chief Executive Officer
CIT	Conceptual Integration Theory
CL	Cognitive Linguistics
CMT	Conceptual Metaphor Theory
DM	Disjunct Marker
EM	Ethnomethodology
EMCA	Ethnomethodological Conversation Analysis
EMCL	Empirical Methods in Cognitive Linguistics
ERP	Event-related brain potentials
fMRI	Functional magnetic resonance imaging
FPP	First-Pair Part
FTA	Face-threatening Act
GAT	Gesprächsanalytisches Transkriptionssystem
ICE	International Corpus of English
ICM	Idealised Cognitive Model
IPA	International Phonetic Alphabet
IMP	Impersonation
MSCI	Mental Spaces and Conceptual Integration Theory
MST	Mental Spaces Theory
NMTB	Never Mind The Buzzcocks
NTL	Neural Theory of Language
PET	Positron emission tomography
RQ	Research Question
SPP	Second-Pair Part
SSPNET	Social Signal Processing Networks
TCU	Turn Constructional Unit
TRP	Transition Relevant Place

LIST OF TABLES

Table 1	Typology of Impersonation Blends	88
Table 2	Overview of preface formats across blending types	144
Table 3	Distribution of quotatives across blending types (N=109)	144
Table 4	Overview of quotative functions in mirror impersonations	155
Table 5	Overview of quotative functions in single-scope impersonations	160
Table 6	Overview of quotative functions in double-scope impersonations	164
Table 7	Overview of action built through demarcation markers	185

LIST OF FIGURES

Figure 1	Research Questions	4
Figure 2	Mental Space Lattice	10
Figure 3	The Naked Lie	13
Figure 4	The Naked Lie II	14
Figure 5	Webster's counterfactual	15
Figure 6	Schematic blending diagram	18
Figure 7	Compression of vital relations	21
Figure 8	Brandt & Brandt's (2005) discursively-grounded model	46
Figure 9	Brandt's (2010:321) blending typology	49
Figure 10	Langacker's (2001:145) Current Discourse Space	69
Figure 11	Participants in the Ground of televised multiparty interaction	70
Figure 12	Polyphonic blends	87
Figure 13	NMTB panel set-up	92
Figure 14	Data analysis: process	101
Figure 15	CDS for "The Sports Commentator I", FPP ₁ –SPP ₂	107
Figure 16	CDS for FPP ₃₍₁₋₂₎	108
Figure 17	Detailed visualisation of "The Sports Commentator I"	109
Figure 18	Belief Space set up in the preface to "The Proprietor"	114
Figure 19	"The Proprietor" impersonation blend	118
Figure 20	Changing viewing frames in "The Metal Detector" (lines 01-04) ...	121
Figure 21	From narrative to impersonation (lines 06-07)	122
Figure 22	"The Metal Detector" impersonation blend	124
Figure 23	Gaze and postural orientation in "Narcissistic Barrowman"	170
Figure 24	Gaze and postural orientation in "The Sports Commentator I"	171
Figure 25	Bodily orientation in "Jason Orange"	174
Figure 26	Postural orientation in "Northern Mac"	176
Figure 27	Postural orientation in "The Metal Detector"	177
Figure 28	Postural orientation in "Steam-Powered Dildo"	180
Figure 29	Postural orientation in "Alan Bennett Monologue"	182
Figure 30	Postural alignment in "Easy Robin"	184

Figure 31	Intonation	187
Figure 32	Pitch contour on quotative <i>be like</i> in “Bemused Look”	189
Figure 33	Pitch contour on quotative <i>go</i> in “Flirting”	189
Figure 34	Rising pitch contour on “Sports Commentator I” preface	190
Figure 35	Pitch contour on “Narcissistic Barrowman”	192
Figure 36	Pitch contour on Plus-1 in “Narcissistic Barrowman”	193
Figure 37	Pitch contour in 1+1 format (“Banana Outfit”)	194
Figure 38	Pitch contour in 1+1 format (“Travel Lodge”)	195
Figure 39	Pitch contour on single-scope diegetic preface (“Polish Band”)	196
Figure 40	Pitch contour on single-scope diegetic preface (“The Proprietor”) ...	196
Figure 41	Pitch contour on quotative <i>be like</i> in “Justin Bieber”	197
Figure 42	Pitch contour on quotative <i>go</i> in “Upset French”	197
Figure 43	TCU-final rising intonation in list construction (“The Proprietor”) .	198
Figure 44	Co-constructing intervals in “The Proprietor”	198
Figure 45	Intensity for “Northern Lass” punch line-attachment	205
Figure 46	Intensity for punch line in “The Proprietor”	206
Figure 47	Pitch contour for punch line in “The Proprietor”	206
Figure 48	Intensity for punch line in “The Metal Detector”	206
Figure 49	Typology of impersonations for assessments	215
Figure 50	Typology for narrative joke-first impersonations	216
Figure 51	A discursively grounded model of in situ meaning making	220

LIST OF TRANSCRIPTS

In alphabetical order:

“Alan Bennett Monologue”	181
“Babybel”	162
“Banana Outfit”	131
“Bemused Look”	129
“Brothers”	137
“Cooking with Villains”	156
“Displeased Horse”	163
“Easy Robin”	183
“Flirting”	152
“Frustration”	135
“Holiday Horse”	161
“Jason Orange”	172
“Justin Bieber”	160
“Kylie”	158
“Listening”	165
“Mexican Building Society”	139
“Narcissistic Barrowman”	168
“Northern Lass”	134
“Northern Mac”	175
“Polish Band”	156
“Smug Jonas”	130
“Stevenson’s Rabbit”	141
“The Metal Detector”	119
“The Proprietor”	112
“The Sports Commentator I”	104
“The Sports Commentator II”	132
“Travel Lodge”	152
“Umbrellas”	160
“Upset French”	157

TRANSCRIPTION CONVENTIONS

Transcription conventions largely follow GAT I & II (Selting et al.; 1998, 2009).

Sequential structure

[]	Overlappings, simultaneous talk
[]	
=	Turn directly latching on to previous turn

Pauses

(.)	Micropause
(-), (--), (---)	Short, medium, longer pauses of approx. 0.25-0.75 secs., up to one sec.
(1.5)	Measured pause

Further segmental conventions

:, ::, :::	Vowel lengthening according to duration
er, erm, um	So-called filled pauses

Laughter

he(h)llo(h)	Speech laugh particles
hehe huhu	Syllabic laughter
@@, @@, @@@	(Canned) audience laughter according to duration
((laughs))	Description of laughter

Acknowledgment tokens

x, xx, xxx	Audience clapping according to duration
xXXXXx	Increased intensity in audience clapping
@x@x	Audience clapping and laughing

Stress

accentTUAtion	Main stress
Accent!TUA!tion	Extra stress

Intonation contour towards end of TCU

?	High rise
,	Medium rise
-	Plateau
;	Medium fall
.	Fall

Noticeable changes in pitch

↑	Rise
↓	Fall

Intralineal notation of changes in pitch accent

`SO	Falling
´SO	Rising
^SO	Rising-falling
—SO	Level
˘SO	Falling-rising

Changes in loudness and tempo

<<f>	Forte, loud
<<ff>	Fortissimo, very loud
<<p>	Piano, quiet
<<pp>	Pianissimo, very quiet
<<all>	Allegro, fast

<<len> >	Lento, slow
<<cresc> >	Crescendo, increasing loudness
<<dim> >	Diminuendo, decreasing loudness
<<acc> >	Accelerando, increasing speed
<<rall> >	Rallentando, decreasing speed

Additional symbols

((coughs))	Para- and extralinguistic activities and events
<<coughing> >	Para- and extralinguistic activities and events accompanying speech; indicating duration
<<smiling> >	Interpretative comments indicating duration
()	Passage could not be transcribed
(such)	Assumed word
(such/much)	Possible alternatives
((...))	Omissions in transcript
<@	Onset of impersonation
@>	End of impersonation

CHAPTER 1: LANGUAGE, COGNITION, INTERACTION. CONCEPTUAL BLENDING IN A SOCIAL-INTERACTIONAL COGNITIVE LINGUISTICS

1.1. Introduction

Cognitive Linguistics (CL) as a research enterprise began to emerge in the late 1970s as a reaction to the then-prevailing formal approaches attempting to discern the internal structures of language. In perceiving “language as reflecting human experience” (Harder, 2010:14), CL can be regarded as a counter-movement to generative (Chomskyan) linguistics and logical Semantics with their focus on core language, syntax, and symbol manipulation. Rather than seeing language as a separate faculty of the human mind, cognitive linguists started to view language as reliant on the same general cognitive principles as other features of human cognition, such as perception and categorisation, and emphasised its experiential grounding.

With its variety of foci (from the multifarious construction grammars to cognitive semantic¹ approaches) that share a fundamental interest in meaning and the processes that underlie the linking of form to meaning, CL has thus evolved into an attractive alternative to generative linguistics. Defying the central principles of formal approaches to language and the mind, cognitive linguists place special emphasis on *embodiment*. By way of example, CL argues that the bodily nature of language becomes evident in pre-linguistic image schemas, such as the container schema underlying our conceptualisation of the prepositions ‘in’ and ‘out’ (cf., e.g., Lakoff, 1987). Pre-linguistic bodily experience therefore constitutes the basis for the structuring of language and reasoning.

Whilst Cognitive Linguistics and Semantics widen the perspective and acknowledge the bodily dimension of all language use, the research enterprise has been accused of neglecting the social nature of language. For this reason, scholars such as Croft (2009), Harder (2010) and Geeraerts (2005) argue that a ‘social turn’, or a second-generation Cognitive Linguistics, is required if the research enterprise is to overcome “the long-standing separation of the social and psychological dimensions in

¹ Within the greater CL enterprise, Cognitive Semantics inquires after “the nature of meaning” (Williams, 2004:2). Meaning is assumed to result from localised and dynamic processes (*conceptualisation*), which recruit mental simulation and are grounded in embodiment.

the study of human behavior” (Croft, 2009:395). Language is, after all, first and foremost a means of communication and interaction, and is thus inherently social. While a select number of authors, such as Chafe, Clark, and Givón, have tackled the issue prior to this ‘social turn’ and provided important suggestions regarding how the ‘pragmatic’ relates to the ‘semantic’, the vast majority of research into Cognitive Linguistics still avoids focusing on the predominantly social nature of language and, for this reason, remains somewhat “incomplete” (Croft, 2009:395). Yet CL’s guiding philosophy, namely *experientialism* (cf. Lakoff & Johnson, 1980), is conducive to integrating not only the bodily, but also the social grounding of meaning (Harder, 2010). Accordingly, questions as to the “anchoring of meaning in feedback from the environment, outside the individual’s body” (ibid:3) should be addressed in order to advance the research enterprise.

With the present study, I aim to further contribute to the development of a social-interactional Cognitive Linguistics and Semantics. As a case in point, I concentrate on the highly influential Mental Spaces and Conceptual Blending Theory (MSCI). This cognitive-semantic theory, which claims to lay bare the human “heart of imagination” (Fauconnier & Turner, 2002:89), has proven to be highly attractive for the humanities, and has been utilised in literary studies, linguistics, religious studies and anthropology to account for the ways in which meaning is constructed. The theory posits that a new meaningful structure, the blend, emerges from the integration of two distinct semantic scenarios (so-called ‘input spaces’) via a generic space that subsumes semantic structure shared by both inputs. In MSCI, compounds such as *land yacht* (denoting large sedans) and phenomena as diverse as the narrative structures of novels, the development of mathematics, performative rituals, metaphors and symbols are described as employing one and the same cognitive process.

As the great number of applications in varied fields show, the theory has been proven to be extremely productive and intuitively attractive. Nevertheless, it has been criticised for two reasons. This is first of all due to for the veritable dearth of convincing experimental and empirical evidence that support the theory’s claims; secondly, it has been accused of neglecting the social dimension of in-situ cognition and meaning-making. In most MSCI literature, ‘blends’ are discussed as decontextualised cognitive ‘products’ from a post-hoc perspective.

However, the question of when and where all these processes are supposed to occur, of whether the authors are discussing the phylogeny or ontogeny of the

expression, is left unanswered. In addition, the etymology and cultural historicity of such ‘products’ of conceptual integration are neglected in first-generation blending literature. How are novel constructions introduced into discourse? What motivates their construction? In other words: which processes are recruited in a) the introduction of novel blends into discourse, and b) in the parsing of such products of blending? How is conceptual blending coordinated in interaction?

This project endeavours to contribute to answering these questions by focusing on two interrelated aspects. The first research question (RQ1) concerns the ‘bigger’ issue, investigating the ways in which conceptual blending is coordinated in interaction. Rather than starting with the end result and working in reverse to trace the manner in which an ideal cogniser might have arrived at a given construal, my interest lies in the way in which processes of conceptual blending are negotiated in interaction. How are distinct Mental Spaces set up over time? How are the mappings understood to occur between the input spaces signalled? How do interactants arrive at shared construals, i.e. intersubjectively negotiated meaning?

My second research interest (RQ2) concerns the micro-level, and examines the degree to which conceptual blending becomes visible in interaction. To this end, I focus on a theory-internal aspect, namely blending typology, and test the extent to which interactants rely on discrete interactional methods when coordinating the different blending types suggested by Fauconnier & Turner (2002). This typology is arranged on a continuum from less complex to highly complex blends. The latter, so-called ‘double-scope blends’, bring together semantically highly incompatible frames and, in the blended space, integrate elements that are found in both input spaces; conversely, they also show structure found in none of the inputs. The authors suggest that computer desktops (ibid:131) constitute such double-scope blends, as the concept contains elements from both the ‘office’ input space (e.g. folders, recycle bin) and the ‘computer’ input (e.g. save and print commands) without leading to clashes, or rather irritations. It is argued that our ability to integrate such apparently incompatible semantic structures into new scenarios that appear intuitively appropriate lies at the very heart of all human cognising. The highly advanced double-scope blending is contrasted with less complex frame integration, such as verb inflection or colour integration. If double-scope blending is different from less advanced blending and underlies human cognising, then the question arises as to whether this crucial cognitive ability becomes discernable in our daily interactions and whether it

constitutes a discursive practice² in its own right. By analysing the interactional processes employed to coordinate blending-in-interaction (cf. RQ1), I examine the extent to which interactants recruit different methods when coordinating highly advanced double-scope blending than when negotiating less complex frame integration. Figure 1 provides a visualisation of the two interrelated research questions.

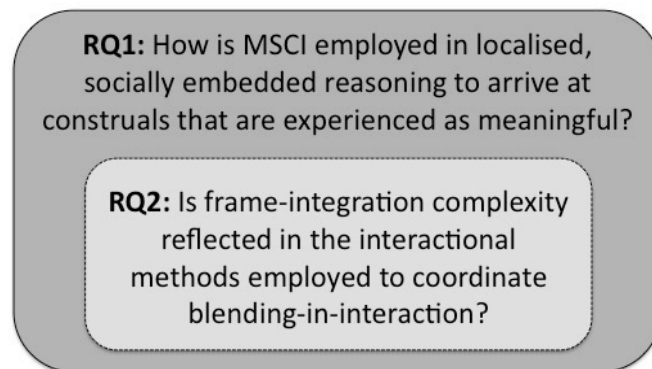


Figure 1: Research Questions

In my analysis, I focus on the discourse level, as this allows attention to be paid to the interactional coordination of creative, dissonant language usage, which involves processes of conceptual blending. This concentration on novel constructions rather than established expressions is important, as it first of all addresses the ‘ubiquity problem’ which MSCI has been accused of. Whether or not phenomena such as colour integration or verb inflection should be described in the powerful terminology provided by MSCI has been subject to debate. Constructions that are experienced as novel on the discourse level, on the other hand, more clearly exhibit the simultaneous maintenance of distinct inputs which MSCI suggests. Secondly, a focus on the discourse level allows for the interactional methods used to coordinate processes of interactional blending to be discerned. My emphasis lies on the ad-hoc ontogenesis of dissonant language use. With the latter, I refer to the production and processing of

² My use of the term ‘discursive practice’ differs from the one proposed in research projects conducted in the traditions of Foucault, Bourdieu and social-constructivism. Rather than analysing the construction of “social realities through actions that invoke identity, ideology, belief and power” (Young, 2008:1), I use the term to refer to interactional, discursive ‘doings’. These ‘doings’ are joint projects (such as negotiating the price of goods or joking), which require coordination, and are informed by abstractions of previous experiences. The term ‘practice’ therefore refers to the interactional negotiation of meaning and the notion that meaning-making is an interactional achievement. In doing so, the term emphasises the cooperative nature of language use, and reflects the fact that meaning is not something ingrained in words (cf. compositional semantics), but is jointly negotiated on the discourse level and thus woven into joint actions.

novel metaphors, similes or other forms of figurative speech that result in ‘irritations’ on the discourse level. Whereas fully conventionalised metaphors such as ‘table leg’ will not be experienced as exhibiting dissonance in most circumstances, in others they might be subject to semantic analysis and consequent ‘re-animation’ (Friedrich, 2012); an example would be an utterance such as ‘My table broke its upper leg’. Using fully conventionalised expressions in scenarios with which they are normally not associated constitutes another form of ‘re-animation’, i.e. dissonance on the discourse level. A case in point would be referring to a ‘blackbird’ as an ‘eagle’. Such incongruities require the cogniser to juggle the inputs (i.e. blackbird – eagle) and draw inferences as to the motivation – and hence meaning – of the use of non-conventional language.

For the analysis to yield comparable results, I examine impersonation humour, an interactional phenomenon that not only relies on discursive dissonance, but also exhibits frame integration of varying complexity. While mocking one’s interlocutor is, in the MSCI conception, less complex, embodying an entity such as a horse or a machine requires the integration of semantically distinct frames, and is hence more advanced. By employing the methodological standards set forth in Conversation Analysis (CA), I – in answer to RQ2 – compare the methods used in setting-up, performing and exiting these various impersonation blends in order to describe the extent to which highly advanced double-scope blending constitutes a discursive practice.

Based on the results, I propose a reconfigured model of MSCI that is discursively and interactionally grounded, and – in answer to RQ1 – therefore provides a means of analysing those social and contextual aspects governing in-situ online cognition which recruits conceptual blending.

1.2. Conceptual blending as discursive practice: Preview

This project aims to contribute to an advancement of both the methodological portfolio of MSCI and its theoretical conceptualisation within a social-interactional Cognitive Linguistics. To this end, a critical discussion of the theory will be presented prior to the suggestion of methodological and conceptual reconfigurations that will allow for the integration of social-interactional dimensions of meaning-making into

Cognitive Semantics. Building on these hypotheses, a detailed analysis of blending-in-interaction will not only serve to reassess the blending typology put forward by Fauconnier & Turner (2002), but will pave the way for a discursively grounded model of MSCi which transgresses the semantic-pragmatic distinction and provides suggestions regarding how socially shared knowledge and beliefs about intentions impact on meaning construction.

As a reflection of this overall structure, the study comprises six chapters in total. Chapter two provides an introduction to Mental Spaces and Conceptual Integration (MSCi) theory, and assesses the theory's potential and limitations with regard to a social-interactional Cognitive Linguistics. Summarising the debates that have surrounded MSCi for the past ten years, I go on to discuss the theory's 'ubiquity problem' and lack of solid empirical grounding. Particular emphasis is placed on the blending typology proposed by Fauconnier & Turner (2002), as the systematics is crucial to RQ2, and informs the analyses conducted in chapter five of the project.

Elaborating upon the results of this critical discussion, chapter three suggests a methodological and conceptual framework for testing MSCi in interaction. After providing a state-of-the-art account of empirical and experiential research in MSCi (with a special focus on discursive approaches) I propose to study the interactional coordination of meaning and the cognitive processes underlying said coordination with the help of the methodology presented in Conversation Analysis (CA). Having introduced the methodology, I provide hypotheses regarding the reconfigurations required when 'marrying' a cognitive-semantic theory to a methodology which considers meaning to be intrinsically linked to social action. By way of example, discussion therefore centres on the way in which notions such as Mental Spaces can be successfully aligned with interactional approaches through the consideration of common ground. Finally, this chapter introduces the domain studied, namely impersonation humour. By delineating the phenomenon and introducing the corpus, this section provides a transition to the social-interactional cognitive analysis that follow in chapters four and five.

Chapter four offers case studies that primarily shed light on both research questions. In addition to tracing how impersonation blends of varying frame integration complexity temporally unfold during interaction, the studies also provide initial evidence that it is not the blending typology suggested by Fauconnier & Turner (2002) that determines the varying interactional methods recruited, but the greater

social action that is negotiated in the performance. Furthermore, the case studies serve to provide an overview of the interactional analyses conducted in chapter five, as they present the individual methods identified in an integrated manner.

Chapter five closely examines the methods identified in the case studies. It provides detailed analyses of the interactional resources employed in coordination impersonation blends, such as prefacing sequences, quotatives, postural orientation and exiting methods, and compares them across blending types in order to determine the extent to which the frame integration complexity is oriented towards the interactional coordination of meaning. This chapter concludes with the assertion that double-scope blending is *not* a discursive practice in its own right, and emphasises the important role recursive beliefs on the social action currently coordinated have on the co-construction of meaning.

Chapter six utilises the findings gathered as a basis for further discussion, and integrates them into a discursively grounded model of in-situ meaning coordination which recruits MSCl. This model is capable of highlighting the effect that recursive beliefs on joint action and shared knowledge have on the emergence of meaning, and, in doing so, accounts for the issues raised in RQ1 and 2. The final sections provide a summary of the project and suggest points for further research.

CHAPTER 2: MENTAL SPACES AND CONCEPTUAL INTEGRATION THEORIES

*Visible language is only the tip of the iceberg
of invisible meaning construction
that goes on as we think and talk.*
Gilles Fauconnier (1997:1)

Where does meaning emerge? Other than formal, composite approaches in semantics, Cognitive Semantics argues that *conceptualisation* is central to meaning-making. In accordance with the encyclopaedic view of meaning representation, language is understood to provide clues that prompt processes of meaning construction in the form of cognitive representations, to purely provide fragmentary specifications for the construction of ephemeral cognitive representations that are fundamentally perceptual in nature, and can recruit mental simulations. As such, meaning does not result from the compositional content of individual sentences, nor is it a matter of relating linguistic entities to the external world; rather, it is essentially conceptual in nature, and results from dynamic mental processes of meaning construction, i.e. conceptualisations.

For purposes of meaning construction, cognisers are assumed to set up temporary conceptual domains in CL; these are so-called *Mental Spaces* which are connected to one another in various ways as discourse unfolds. Assuming mechanisms of conceptual projection also evident in Conceptual Metaphor Theory (Lakoff, 1986; Lakoff & Johnson, 1980) and schema induction (knowledge structures such as Idealised Cognitive Models, see Lakoff (1987)), Mental Spaces Theory as primarily developed by Gilles Fauconnier (1984, 1994, 1997) presents a detailed account of how language used in its social and cultural contexts provides instructions for local and dynamic meaning construction in discourse. As Mental Spaces Theory (MST) forms the basis upon which Conceptual Integration Theory (CIT) is built, the following will briefly illustrate the key assumptions of the first theory before discussing in detail the central ideas which constitute CIT.

2.1. Mental Spaces Theory

Central logical problems in the philosophy of language represented the starting point for the development of what was to become Mental Spaces Theory. In the late 1970s, Fauconnier proposed solutions for a number of longstanding puzzles in semantics, such as indirect reference, presupposition and counterfactuals, by showing that thought and ultimately language is dependent on our manipulating of “webs of mappings between Mental Spaces” (Fauconnier, 1997:149). In MST, two aspects essentially comprise meaning construction. Mental Spaces are set up first of all, while the second aspect is to establish mappings between these Mental Spaces. The mappings are situated and governed by the local context.¹ In more detail, Mental Spaces are, according to Fauconnier’s (1997:11) definition, “partial structures that proliferate when we think and talk, allowing a fine-grained partitioning of our discourse and knowledge structures.” These Mental Spaces contain elements which designate discourse entities, and are considered to be structured by Idealised Cognitive Models (ICMs) as proposed by Lakoff (1987). They therefore depend on the more stable domains from which they recruit information for ‘on-line’ meaning construction. ‘Meaning’ is consequently divided into discrete, temporary conceptual ‘packets’ that are built ‘on-line’ for purposes of local understanding of narratives, metaphors, speech acts and “general reasoning” (Fauconnier, 1997:5), and contain specific aspects of information that provide the “cognitive substrate” (ibid:34) for meaning-making. They are largely ephemeral, and rely on more stabilised forms of background knowledge and the discourse context, which includes local framing, genre and setting.

¹ Mental Spaces Theory refrains from formally separating the pragmatic from the semantic, and acknowledges (albeit theoretically) the power pragmatic aspects exercise upon on-line meaning construction processes. Other than research conducted in cognitive pragmatics (e.g. Carston, 2002; Schmid, 2012), however, MST does not place explicit emphasis on identifying those general cognitive principles and processes that form “pragmatic competence” and allow for context-dependent meanings to emerge.

As discourse unfolds, the different conceptual packets are built up and linked through *mappings*. *Space building*, the chunking of information into ‘objectively’ different types (locations, beliefs, time periods, hypothetical situations, counterfactuals, time periods etc), and/or the establishment of new spaces or the shifting of focus to spaces already built, is achieved through space builders. Space builders come in various shapes and forms, and impose mapping conditions upon spaces by constraining the number of possible interpretations. Grammatical expressions, such as prepositional phrases (“In London...”, “In 2002...”), adverbials (“In actuality...”, “Literally...”), subject-verb complexes (“Mary thinks...”) or conjunctions-cum-clauses (“Should the President agree...”), are one type of space builder that configure the number of possible interpretations available.

In Mental Spaces Theory, the spaces established as discourse unfolds are assembled in a Mental Spaces *lattice*, as illustrated in Figure 2.

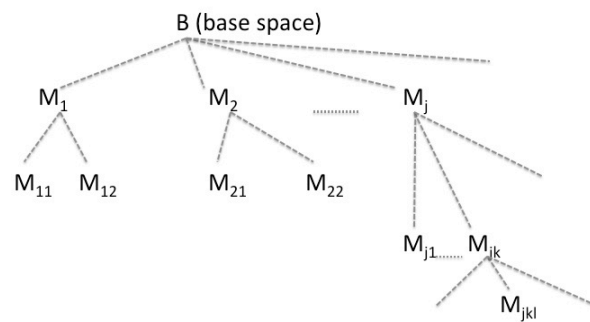


Figure 2: Mental Space Lattice (cf. Fauconnier (1997:39))

In the lattice, one space acts as the *base*, and another (or the same one in certain cases) is in *focus*. Focus Spaces are those spaces that are currently internally elaborated, whereas the Base Space is the starting point of a given construction and can always be returned to throughout the discourse. Further spaces set up will either relate to Base Space or Focus Space. In meaning construction, one of the spaces in the lattice acts as the *viewpoint*. Viewpoint Spaces are those spaces from which the others are accessed. Throughout the process of meaning construction, cognisers navigate the space lattice, shifting focus and viewpoint from space to space with the help of the language’s tense and mood system. The spaces set up by space builders and grouped in the lattice are linked to one another via *connectors* and other structures. Copula or ‘copulative’ verbs, such as “become”, can, for example, act as trans-spatial connectors. Of special importance here is the *Access Principle* (also called *Identification* or the *ID principle*), which

asserts that “if two elements a and b are linked by a connector F ($b = F(a)$), then element b can be identified by naming, describing, or pointing to its counterpart a ” (Fauconnier, 1997:41)². In less formal terms, the access principle maintains that expressions which identify or detail a given element in one mental space (so-called *triggers*) can be employed to access its match (so-called *targets*) in another mental space.

When building up the lattice, default structuring processes apply in order to yield structure that is not explicitly referred to in the utterances themselves. First of all, some structure of Mental Spaces is presupposed, and spreads by default through the lattice. This allows structure to be added to Mental Spaces without explicit reference to these aspects. This phenomenon is known as the *Presupposition Float*, and is formalised in Fauconnier (1997:61) as follows: “A presupposed structure Π in mental space M will propagate to the next higher space N , unless structure already in M or N is incompatible with Π , or entails Π .” In short, the Presupposition Float contends that structure will advance through higher spaces in the lattice until it encounters identical or opposing structure. Mental Spaces higher up in the lattice can consequently inherit structure from lower spaces; these are so-called ‘semantic entailments’. Consider the following example adapted from Fauconnier (1997:62):

Sarah wishes Anna’s children would visit her.

Here, the presupposition $\langle HAS\ CHILDREN \rangle$ will float from the newly built space M (set up through the space builder “Sarah wishes”) to the base space B . Consequently, the speaker is assuming that Anna has children and that Sarah is assuming that Anna has children. Naturally, the Presupposition Float can be cancelled in discourse when new information becomes available (e.g. that Anna actually does not have children). Secondly, *optimisation* is a mechanism that supplies structure not explicitly mentioned in the utterance itself from a ‘parent space’ to a ‘child space’. Aside from the Presupposition Float, then, optimisation is a form of downward spreading in the lattice.

² It becomes evident here that Fauconnier himself – despite insisting on a focus on ‘natural’ language – makes use of invented and construed sentences to discuss traditional problems in the philosophy of language. He also largely relies on traditional logics and mathematical notations that are reminiscent of formal approaches. A clear separation from formal semantics is therefore not fully evident, especially given that Fauconnier is rooted in formal discourse and prone to ‘mathematisation’ (at least in his early work (1994, 1997)).

The establishment of mappings between the Mental Spaces that are built up as discourse unfolds occurs in micro and macro-meaning-making processes. “A mapping, in the most general mathematical sense,” Fauconnier writes, “is a correspondence between two sets that assigns to each element in the first a counterpart in the second” (Fauconnier, 1997:1). Mappings, he goes on to explain, “are at the heart of the unique human cognitive faculty of producing, transferring, and processing meaning” (ibid). Mappings consequently enable meaning construction. Here, Fauconnier distinguishes between different sorts of mappings; wish to briefly introduce several key types here.

As the name suggests, *projection mappings* project parts of the structure of a source domain onto the structure of a target domain, as is the case in metaphors. These mappings are either culturally shared and, such as the English TIME AS SPACE conceptual metaphor, lexically entrenched (as in *Easter is approaching*) and hence opaque, or are set up in local contexts and are then experienced as novel and creative expressions (or “alive” metaphors, to use Ricoeur’s (1975) terminology). Fauconnier, unlike Ricoeur, refrains from differentiating these mappings formally, instead arguing that many of the expressions experienced as ‘creative’ are purely extensions of conventionalised (lexicalised and grammaticalised) mappings.

In a roughly similar fashion, *analogical mappings* induce schemas to produce analogies, such as the infamous *computer virus* that integrates abstract generic features of biological viruses with certain computer programmes. Analogical mappings might, in a second step, guide processes of categorisation when, for example, a number of computer programmes are categorised as viruses. The analogical mapping then becomes lexicalised and grammaticalised, and is readily available for further reasoning and elaboration.

Pragmatic function mappings, on the other hand, connect two locally established domains by virtue of two corresponding categories of objects mapped onto one another by a pragmatic function. This would, for example, be the case when matching patients with their illnesses, as in *The discectomy in four*.³ Consequently, metonymy and synecdoche in Fauconnier (1997) constitute pragmatic function mappings.

³ Fauconnier does not, however, elucidate further on how such pragmatic mappings occur. Why is it that in this example, the term ‘discectomy’ is interpreted as referring to the patient rather than, say, surgical instruments?

Furthermore, *schema mappings* are employed when schemas, frames or models structure a “situation in context” (ibid:11), such as is the case in grammatical and lexical constructions that prompt meaning schemas. This would, for example, apply to the *buying and selling*-frame described by Fillmore (1982), which identifies roles in a given frame. Many of these mappings are discussed in later works as involving frame-blending and emergent structure.

To illustrate the central ideas, concepts and notations of MST in further detail, I would like to briefly present Fauconnier’s (1997:120ff) *The Naked Lie* example, as this is one of the few examples he utilises that contains ‘actual’ interaction and, consequently, “negotiated space building” (Fauconnier, 1997:121). Furthermore, it shows the way in which frame blending effects can occur in the interaction of Mental Spaces. *The Naked Lie* is a film in which the following dialogue takes place:

Victoria: *What if it were your sister?*

Webster: *I don’t have a sister, but if I did, she wouldn’t be a hooker.*

In the first sentence, Victoria tries to prompt Webster to show sympathy for a murdered prostitute by proposing a counterfactual model. However, this is in turn answered by Webster with another counterfactual model that thwarts the inference intended by Victoria.

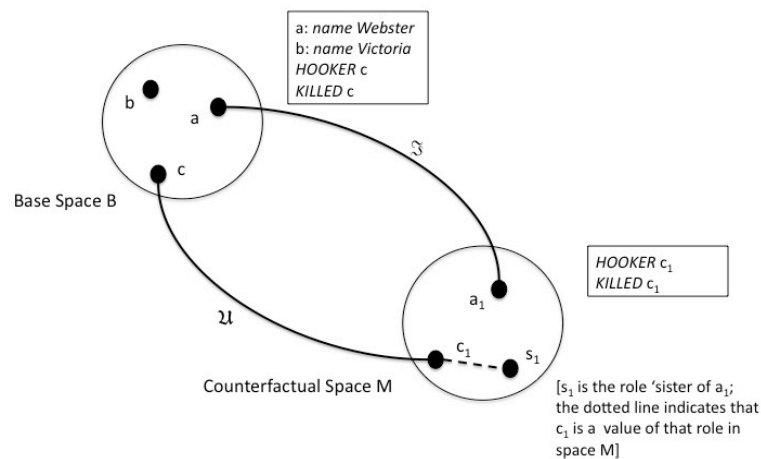


Figure 3: The Naked Lie (cf. Fauconnier, 1997:121)

Figure 3 illustrates the space configuration set up by Victoria in her first utterance: A counterfactual space M is constructed through the grammatical space builder “what if”. Elements c and c₁ are linked via the analogy connector \mathfrak{A} , as Victoria is not

proposing identity; by asking “What if it were your sister?” rather than “What if she was your sister?”, Victoria is setting up an analogy rather than construing a counterfactual that would maintain the ‘same’ girl in Space M whilst simultaneously being framed as Webster’s sister. Still, her use of “your sister” can be read as her presupposing that Webster had a sister. For the mental space configuration, this would mean that element c_1 in space M would be connected by identity connector \Im back to element c' (*Webster’s sisters*) in the base space. Element a , on the other hand, is linked to a_1 via the identity connector \Im , as Webster ‘remains’ the same in space M (see figure 4).

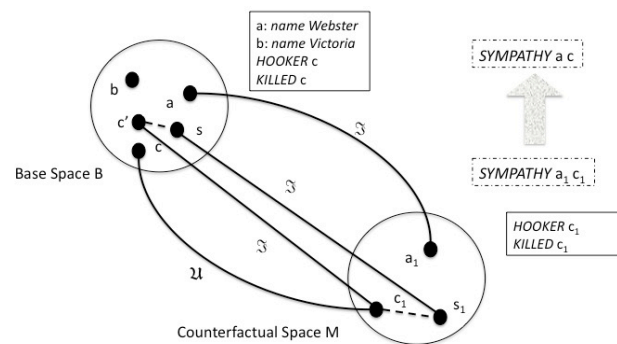


Figure 4: The Naked Lie II (cf. Fauconnier, 1997:123)

The counterfactual now allows for inferences to be generated that draw on cultural models such as sympathy for the death of family members. These inferences in turn float upwards to the base space, as illustrated in Figure 4. Furthermore, a frame blending effect is achieved, as element c_1 tallies with both domains: in the base space B, it fits c , the hooker, and in the counterfactual space M, it fits c' , Webster’s sister. As a result, pertinent qualities of c_1 can be derived. One example may be that Webster should feel sympathy for the murdered prostitute. Webster, however, rejects the inferences intended by the blend Victoria set up. The exchange at this point develops into a contest of attacking inferences and construing alternative inferences, as Webster intends to spoil Victoria’s construction in two ways. First of all, he states that he has no sister, consequently ‘removing’ c from the base space, which in turn leads to the role s (Webster’s sister) no longer having a value. This additionally results in the transfer of the ‘family frame’ (feeling close to one’s sister) to c_1 being prevented. Webster consequently has no need to feel sympathy for the murdered woman. In a second step, Webster blocks another potential construal in which he, in theory, could

have a sister working as a prostitute; he does this by preventing the condition *HOOKEE* from being linked to the role of his sister. This is specifically accomplished by Webster setting up his own counterfactual space, N, in which an imaginary sister *s* could, under no circumstances, be a prostitute, as illustrated in figure 5.

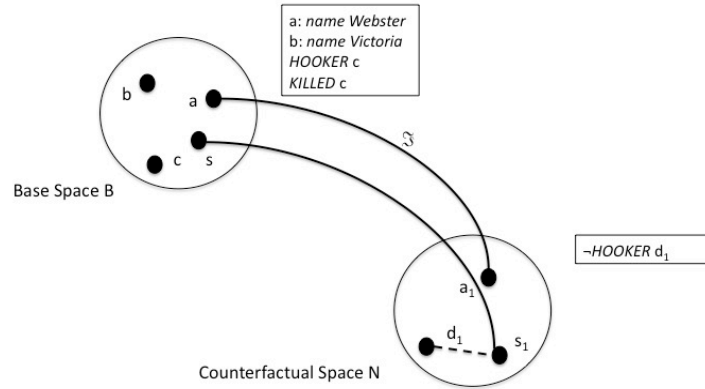


Figure 5: Webster's counterfactual (cf. Fauconnier, 1997:124)

In all, Fauconnier's analysis of *The Naked Lie*, illustrates the key elements of Mental Spaces Theory: Mental Spaces are built as discourse unfolds through pragmatic or grammatical space builders. The spaces themselves contain elements that can be connected to their counterparts in other spaces via connectors, such as the Identity connector. Spaces are furthermore structured by induced frames that can provide further roles and values. In the interaction of spaces, inferences might arise which can be attributed to frame blending effects. This will be discussed in further detail below.

With his model, Fauconnier manages to account for counterfactual reasoning and other long-unsolved puzzles in the philosophy of language. He shows that grammar only directs meaning construction to a certain degree, and further decisions as to the 'meaning' of utterances have to be made on a construction level, namely the detailing of the mappings, and on a pragmatic level, namely the detailing of the implications of the interpretation. Thought and reasoning are thus fundamentally creative processes, as essentially new connections and conceptualisations are set up at every moment of the unfolding discourse.

2.2. Conceptual Integration Theory

Conceptual blending is first introduced (Fauconnier, 1997; Fauconnier & Turner, 1994, 1996, 1998) as a cognitive operation integrating partial structures from distinct input domains which share generic structure and are connected analogically; however, they yield emergent properties that are found in neither of the input spaces in a fourth domain, the so-called blended space. Blending, the theory's founding fathers claim, is the basic and fundamental cognitive operation that underlies all forms of creativity and thought, be they linguistic (metaphor, narrative, grammatical constructions) or non-linguistic (art, mathematics, religion) cognitive phenomena. Consequently, blending is at work on micro and macro-levels of meaning-making processes, enabling grammatical constructions as well as complex mental explorations to occur. The success of Mental Spaces and Conceptual Integration Theory (MSCI) is largely due to its ability to elegantly unite disparaged approaches in metaphor theory, transcending traditional dichotomies, such as the literal-poetic or alive-dead metaphor binary oppositions, and being able to account for metaphors in a more flexible manner than Conceptual Metaphor Theory⁴ (Kok & Bublitz, 2011). Combining the mapping element that is central to Conceptual Metaphor Theory (Lakoff, 1987; Lakoff & Johnson, 1980) with ideas of emergent structure as proposed in interaction theories of metaphor (e.g. Black, 1962; Ricoeur, 1975) and the *tertium comparationis* structure evident in early metaphor theories (e.g. Aristotle; Lakoff & Turner, 1989) led to MSCI being enthusiastically received in linguistic and literary research on meaning construction in figurative language use (see Fludernik, Freeman, & Freeman, 1999 for an overview of the reception of MSCI).⁵

Nevertheless, emergence and blended structure are not entirely novel concepts. The notion behind 'emergence' has been in use in philosophy at least since

⁴ In a forthcoming publication, Fauconnier & Lakoff (fthc) argue that both Conceptual Metaphor Theory (CMT) and Mental Spaces and Conceptual Integration Theory (MSCI) are to be understood within the Neural Theory of Language (NTL) largely devised by Lakoff (2008) and Feldman (2006). However, as NTL remains somewhat contested within Cognitive Linguistic (cf. Harder, 2007) and – despite its name – thus far relies on computational modeling rather than investigating actual human cognising and its neurobiological grounding, I refrain from discussing MSCI in relation to NTL.

⁵ Fauconnier and Turner (2002) acknowledge Koestler's (1964) influence, yet do not refer to other theories which may have impacted on their model. Nevertheless, established ideas on metaphor and rhetoric are evident in their thinking.

the time of Aristotle, and in systems theory and science to describe the ways in which novel structures emerge from the interaction of established patterns. Aristotle's conception in particular closely resembles the notion advocated in MSCl. While not using the term emergence, Aristotle writes in part 6 of his *Metaphysics* (τὰ μετὰ τὰ φυσικά) that "In the case of all things which have several parts and in which the totality is not, as it were, a mere heap, but the whole is something beside the parts". As such, this already points to the characteristics ascribed to the emergent structure of a blended space in MSCl. The first person to use the term *emergence* to describe the concept outlined by Aristotle was physiologist George Henry Lewes (1875:413), who defined the term as follows:

"It is otherwise with emergents, when, instead of adding measurable motion to measurable motion, or things of one kind to other individuals of their kind, there is a co-operation of things of unlike kinds. The emergent is unlike its components insofar as these are incommensurable, and it cannot be reduced to their sum or their difference."

While Lewes refers to the question as to how consciousness may emerge from nerve activity, his treatment can nonetheless be read as analogous to concepts used in MSCl, namely the interaction of semantically incompatible frames and the emergent structure that arises from said interaction. The parallels in terminology are striking, too, as even at this early stage Lewes (1875:93) talks of 'blending' in relation to emergent structure when he writes that "the difference between an aggregate and a product is that in the first case the component parts are simply grouped together, added; in the second, the constituent elements are blended, multiplied into each other."

To illustrate how blending is intended to work in MSCl, I would like to focus on the following example which has been the subject of much discussion in blending literature (e.g. Fauconnier, 1997:155ff): in 1993, the catamaran *Great America II* sailed from San Francisco to Boston, a route also followed by the clipper *Northern Light* in 1853. Yachting magazines observing the catamaran's trip in the 1990s compared it to that of the clipper, using statements such as

As we went to press, Rich Wilson and Bill Biewenga, [the crew of the catamaran] were barely maintaining a 4.5 lead over the ghost of the clipper Northern Light [...].

In accounts such as this one, the two boats are portrayed as being in direct competition with one another. Both boats are understood to be sailing the same route in the same year, with the events of 1853 and 1993 being blended into a single event. The blended space in which the boats are racing against each other in the 1990s inherits partial structure from both input spaces. From 1993, the blended space takes the weather conditions, date and one vessel, while from 1853 it recruits the other vessel, but discounts any additional information. The emerging structure now facilitates a direct comparison of the two boats and an analysis thereof in regatta terms. This also adds an element of excitement despite the fact that almost nobody would believe that the 19th century clipper had made a magical reappearance, and that the catamaran was actually racing a ghost ship. Yet the mental “dynamic simulation” (Coulson & Oakley, 2000:176) of the scenario yields novel information. Through this emergent structure, inferences that reflect back onto the input spaces can arise. In the example, this would concern conclusions with regard to the speed of the two boats.

How is this ordinary cognitive operation achieved? In the most general and abstract terms, conceptual integration involves the following aspects: cross-space mapping between the input spaces, the setting up of a generic space, and the projection of partial structure from the input spaces into a fourth blended space that subsequently yields emergent structure. The schematic diagram in figure 6 illustrates this process.

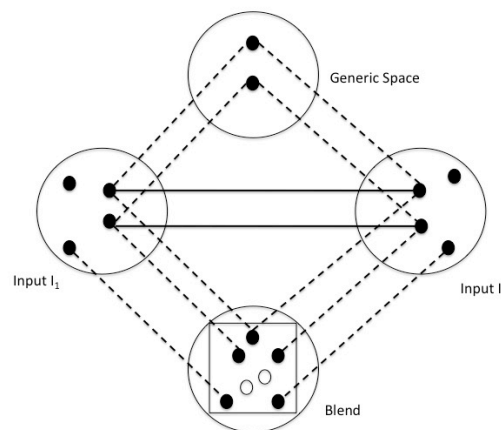


Figure 6: Schematic blending diagram (cf. Fauconnier, 1997:151)

As MSCI is a highly complex theory, the following is intended to shed more light onto the constitutive as well as the governing principles of conceptual blending before

summarising the blending taxonomy proposed by Fauconnier & Turner (2002) and offering a critical discussion of the theory.

2.2.1. Constitutive principles

As the example above has shown, conceptual blending relies on the setting up of Mental Spaces and the mappings that connect them. As in the original Mental Spaces theory devised by Fauconnier (1994), Mental Spaces in conceptual blending are largely ephemeral conceptual units that are set up for online meaning construction, but which rely on more established and stable domains of knowledge. Although frames structure Mental Spaces internally, they can be modified as discourse develops. External links to other Mental Spaces, namely *matchings*, are built up in meaning construction, during the course of which cognisers can also move freely around the space lattice. These matchings, also called *counterpart connections*, link roles and frames within and between input spaces. *Mappings* occur when matches across spaces can be identified, such as the captains of the two boats in the boat race example above (see *vital relations*). In neural terms, Fauconnier & Turner (2002:40) claim that Mental Spaces correspond to “sets of activated neuronal assemblies”. Mappings between spaces, they continue, are “coactivation-bindings of a certain kind” (ibid). As research into the neurosciences has thus far not been able to prove this neural theory of conceptual blending, at this point it can only be regarded as the substitution of the terminology of one theory with that of another without yielding any further insight into the mechanisms discussed. It remains to be seen whether future work in the neurosciences field will provide further evidence in favour of MSCl.

Three types of Mental Spaces are assumed in a blending network: input spaces, a generic space and a blended space. The *input spaces* submit knowledge relating to the domains introduced to the network, such as the scenarios of 1853 and 1993 in the example above. The *generic space* subsumes those features the input spaces have in common, and is thus reminiscent of the *tertium comparationis* structure found in earlier theories of metaphor (cf. Aristotle), but also in Lakoff and Turner (1989) where this structure was referred to as *generic-level*.

The crucial element of blending theory is, however, the blended space that denotes the emergent, novel structure arising from the interaction between the two input spaces. The emergent structure of figurative language is not new per se (see

interaction theories of metaphor such as Black (1962) or Richards (1936)), but Fauconnier & Turner describe it in unprecedented detail. In their account, blending yields emergent structure over the course of three interrelated steps. *Composition* brings together elements from the input spaces and allows for correspondences to surface that do not exist in the input spaces. In the regatta example above, the blend contains two ships racing against each other, while both inputs only contain one ship. Through *selective projection*, only select aspects are transferred into the blended space. In the regatta example, the weather conditions of the respective years are, for example, discounted. *Completion* adds further structure to the blend, such as the regatta-frame that is added to the processing of the example discussed above by virtue of the expression “maintaining a lead”. Pattern completion, which is understood to be the recruitment of additional frame features by virtue of having parts of an established frame available, is extremely common in blending. Completion leads to the blend being ‘integrated’, as it has become “an instance of a particular familiar frame” (Fauconnier & Turner, 2002:44), which a prerequisite for the final step in the blending process, *elaboration*. Elaboration is also referred to as a ‘running of the blend’, and denotes the imaginative simulations and creative modifications which are possible. Cognisers can now ‘live in the blend’ and further exploit its structure; one instance would be when cognisers profile other aspects of the regatta-frame induced in the example above, such as a tight finish. While ‘running the blend’, the input spaces remain active and present in ‘the backstage of our minds’, so that ‘backward projection’ – creating inferences relating to the input spaces – is possible. That which is ‘visible’ or ‘onstage’ to the cogniser, Fauconnier & Turner claim, is only that blended space, the mental simulation, the scenario that is run, while the mappings occurring between the input spaces and the generic space remain subconscious. In other words, we automatically run a mental simulation of the two ships racing each other, yet are unaware of the processes that enable this moment of epiphany, or “global insight” (Fauconnier & Turner, 2002:323).

Not all blends are novel, however, as some may become ‘entrenched’ through extensive usage. Fauconnier & Turner claim here, for example, that complex numbers are blends which now provide the cognitive scaffolding for a wider community. This also illustrates the fact that blends are employed for a variety of purposes: they may, for example, be used to yield emotional reactions in moral and political debates (cf. Coulson’s (2000) discussion of framing in abortion discourse). Blends are further used

in counterfactual reasoning to arrive at novel results in science by altering the conceptual scaffolding (such as the introduction of complex numbers), or to integrate actions which allow for novel processes to make sense (such as in the introduction of a desktop interface on computers and other processes of remediation as described by McLuhan (1964) and Bolter & Grusin (1999)).

Finally, blending allows for *compression* to take place. Compression as a central feature of human cognition packs the complexity and multidimensionality inherent in aspects of such *vital relations* (akin to the *connectors* in the original Mental Spaces Theory) as time, place and identity into smaller packets to reduce them to ‘human scale’; this encompasses “direct action and perception inside familiar frames, typically involving few participants and direct intentionality” (Fauconnier & Turner, 2002:322). This is evident, for example, in images reducing millions of years of evolution into three figures that ‘read’ from left to right, with each metonymically representing one step in the process. Here, outer-space relations (links connecting input spaces) of identity, time, space, cause-effect, change, representation and uniqueness are compressed into inner-space links in the emergent blend. Changes that occurred over a time span of millions of years are represented by individual animals. A variety of locations is compressed into one, as figure 7 illustrates.⁶

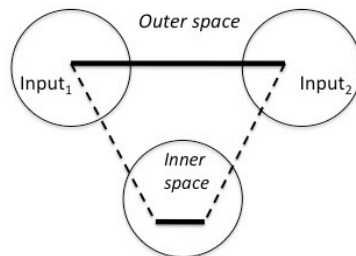


Figure 7: Compression of vital relations (cf. Fauconnier & Turner, 2002:94)

In the regatta example discussed above, the vital relation of *time* becomes compressed. A time span of 140 years separates the two journeys, yet the blended space compresses them into one scenario in which the boats compete directly against one other. Compression is what helps us to grasp complex scenarios more easily. Yet for compression to yield such insights, the blended space needs to be decompressed and

⁶ Hougaard (2004, 2005) discusses the ways in which conceptual blending not only facilitates a reduction to human scale, but also leads to an ‘expansion’ to human scale via conceptual disintegration.

disintegrated. The input spaces need to be simultaneously identified and kept in mind for backward projection to be possible. Returning once again to the regatta example, the cogniser will have to break down the blended scenario and identify the separate input spaces in order for the cognitive effect that the integrated blend hopes to achieve to surface.

On the constitutive layer, then, conceptual blending relies on the setting up of Mental Spaces and the matchings occurring between them by means of selective projection; these matchings yield novel, emergent insights that are not found in the respective inputs through selective projection via vital relations. Compression allows for global insight on a human scale to emerge in the blends, which often unite complex and semantically distant scenarios. Of course, blending is not only determined by selective projection and vital relations, but is also subject to constraints.

2.2.2. Governing principles

Does conceptual blending occur watonly? Is everything ‘blendable’? Early MSCI research was often accused of being too ‘unconstrained’, of advocating an ‘anything goes’ theory (cf. Gibbs, 2000), as it could not provide an adequate explanation for constraints on conceptual blending. Fauconnier & Turner’s 1998 and 2002 publications consequently saw the introduction of so-called *optimality constraints* or *governing principles* as a means of bringing order to the seemingly unpredictable and ubiquitous nature of blending. Conceptual integration, the authors claim, is on the one hand already heavily determined by the constitutive principles outlined above. Unfortunately, without further elucidation on the extent to which the constitutive principles of blending actually constrain conceptual integration, the authors suggest that certain governing principles additionally provide “strategies for optimizing emergent structure” (2002:311).

Governing principles primarily regard compression, but also concern “topology, pattern completion, integration, heightening of relations, maintenance of connections in networks, perspicuity of the blend, and relevance of structure in the blend for the entire network” (ibid:312). It quickly becomes evident that the majority of the authors’ governing principles are concerned with the ‘interior’ dimension of the blend rather than the ‘external’ world in which the blend is embedded.

For compression, the main blending constraint is to “achieve human scale” (ibid:312). Complex scenarios can thus be grasped in the blend, as it compresses such vital relations as time, space or identity. We are suddenly able to reduce evolution to a few individual figures and can make boats race against each other despite the fact that decades lie between their respective journeys. To maximise this ‘eureka’ effect of compression, blending for compression is subject to a number of governing principles. They first of all concern *borrowing for compression*. Here, inputs which are already tightly compressed to human scale are integrated with less compressed spaces. The compression of the first input is then projected into the blend, reducing the complexity of the second input. Fauconnier & Turner (2002) provide examples such as ‘digesting books’, in which the tightly integrated concept of digestion is integrated with the process of reading and interpreting a piece of literature.

Secondly, compressions are governed by the *scaling* of vital relations in the blend, “shortening the causal chain from many steps to few” (ibid:313), as well as by the *syncopation* of dispersed, ‘irrelevant’ structure. Compression further achieves its goal of global insight and the reduction of relations to human scale via *creation*, or the adding of a vital relation in the blend that is not present in the inputs (such as ‘intention’ in diagrams of evolutionary processes), and the *highlighting* of entire stories by means of the presentation of one central scenario, as with the Grim Reaper. The principles for compression are not absolute, as Fauconnier & Turner (2002) see them in competition with each other; as such, they cannot be witnessed to the same extent in every network.

Blending as a whole is also subject to constraints which Fauconnier & Turner (2002) present as follows:

- The *Topology Principle* aims at preserving the organisational structure of the inputs whilst at the same time enabling optimal compression. As human scale is not achievable when too much of the original frame structure of the inputs is transferred to the blend, the topology principle is at odds with the overall goal of compression. Balance between the two needs to be achieved for the blend to be successful.
- The *Pattern Completion Principle* asserts that elements in the blend should be completed by integrating established compressed frames whose relations are compatible with the inputs. The principle introduces familiar frame structure

as additional inputs to the blended space as a means of achieving global insight. This was, for example, evident in the 2009/2010 US Health Care Debate during which opponents recruited the ‘Obama’s death panel’ narrative, claiming that seniors would be asked to end their lives in order to spare the system. Here, the tightly integrated, familiar frame of ‘death panels’ was drawn on to create a scenario in which innocent elderly citizens are effectively murdered by the government for cost-saving purposes.

- The *Integration Principle* maintains that the overall goal of blending should be to achieve an ‘integrated blend’, in which potential clashes between inputs are resolved and a new structure that can itself be manipulated has been created. This then leads to structure that might potentially disturb the blend being omitted.
- The *Maximisation of Vital Relations Principle* contends that while blends develop internal vital relations, they should also reflect the outer-space vital relations to achieve the best-possible ‘eureka’ effect. Fauconnier & Turner (2002:330) briefly mention that “purpose” may play a role here, but otherwise do not touch upon ‘blending external’ aspects of meaning-making in further detail.
- In a similar vein, the *Intensification of Vital Relations Principle* hopes to intensify the vital relations that are already present in the input spaces rather than the ones that develop in the blend. This applies when cause-effect chains are simplified and profiled in the blend, as happens frequently in political discourse. Consider once more the debate surrounding US Health Care System Overhaul, where opponents also used the ‘killing granny’ narrative to frame their argument. A predecessor to the ‘death panel’ discussed above, the ‘killing granny’ narrative stated that the new Health Care system would encourage your grandmother to die. Leaving other aspects aside, this blend compresses hundreds of thousands of seniors into a single familiar person.
- The *Web Principle* subsumes the idea that “manipulating the blend as a unit” (Fauconnier & Turner, 2002:331) must be possible without the blend losing its connections to the input spaces. As such, all spaces in the network are always subconsciously present.
- The *Unpacking Principle* states that the blend alone should enable the reconstruction of all inputs contributing to the blend. This, the authors claim, is often due to incongruities. These incongruities in the product of blending

prompt ‘receivers’ to reconstruct the inputs (as would be the case in the many advertising examples discussed in the literature). Fauconnier & Turner (2002:333) also briefly allude to the role that the “discourse environment” plays in setting up inputs, but do not delve into the matter further.

- The *Relevance Principle* is the most pragmatic principle found among the governing principles, yet aside from its blending-‘external’ dimension – the relevance to communication – the authors focus predominantly on “Network Relevance”. This blending ‘internal’ aspect is satisfied when elements act as successful prompts for unpacking the blend. Thus, “[t]he Relevance Principle pressures networks to have relationships in the blend that are compressions of important outer-space relations between the inputs” (Fauconnier & Turner, 2002:334). As with the other principles, this principle also focuses largely on network-internal aspects rather than taking into account the discourse context in which all communicative signals are embedded.

The blending constraints proposed by Fauconnier and Turner (1998, 2002) were introduced to account for criticism voiced towards the theory’s apparent ‘randomness’. Although the constraints proposed indeed allow for a better understanding of the power of conceptual blending, they still fail to account for the ‘worldly’ dimension of cognition, as they neglect the impact which contextual factors have on the shaping of cognitive processes and the power they exceed on determining the ‘content’ and, in consequence, the meaning of an emerging blended scenario.

2.2.3. Blending typology

In canonical blending theory, four different kinds of network are proposed that account for “the variety and creativity in the way we think” (Fauconnier & Turner, 2002:121): Simplex networks, mirror networks, single-scope and double-scope networks. These blending operations are thought of as occurring on a continuum of increasing complexity, with double-scope networks being the most advanced.

Simplex networks are the most basic type, and denote integrations in which one input provides a familiar frame and the other the values. To consider an example, the culturally entrenched ‘family frame’ can be blended with unframed values such as ‘Peter’ or ‘Mary’. The blended space would then contain a family with the value of

‘Peter’ filling the role of ‘Mary’s husband’. The family frame now provides structure to the formerly unframed values. As such, the mapping occurring between the input spaces is a “Frame-to-values connection” (ibid:120). No clashes between inputs occur, as they are both compatible. Simplex blends consequently do not feel like ‘proper’ blends, as they are reminiscent of compositional and truth-functional theories of meaning-making. However, Fauconnier & Turner claim that simplex networks are a type of blending, as they integrate a frame that provides roles with a space containing elements lacking a frame. The emergent properties of the simplex blend give rise to new roles, such as that of ‘Mary’s husband’, thereby compressing the vital relation of role-value.

Next on the blending continuum from simplicity to complexity are so-called *mirror networks*. In this blending type, all spaces share the same organising structure, i.e. the frame that sets out the “nature of the relevant activity, events, and participants” (ibid:123). The organising structure, however, needs to be cognitively ‘representable’ (though no further information on the nature of this cognitive representation is given). The regatta example discussed above is a mirror network blend. Here, both input spaces as well as the generic and blended space share the structure of two boats sailing a course. The blended space, however, provides more information than its inputs. Its structure is richer, as we now not only have two yachts sailing a course, but in fact conceptualise them as racing on that course in a regatta. Mirror networks consequently share space topology on a frame-structuring level, which facilitates mappings. The only clashing that occur in these networks take place at the below-frame level. In the yacht example, clashing could be represented by different purposes of travel or different types of boats.

Other than simplex and mirror networks, *single-scope networks* integrate input spaces with conflicting frame structure. Only one of the organising frames is projected into the blended space, whereas the organising frame of the other input is neglected. To illustrate this network type, consider the boxing CEOs example discussed in Fauconnier & Turner (2002): it might well be possible to conceptualise two CEOs in a business competition as being engaged in a boxing match. The CEOs might receive blows, be knocked out, or recover in their day-to-day business lives. Understanding business in terms of boxing involves a single-scope network, as we are faced with clashing organising frames, namely that of business and that of boxing. Yet only one organising frame, the boxing, is transferred to the blend and provides its topology.

Single-scope blends are thus believed to lie at the root of many conventionalised conceptual metaphors (cf. Lakoff & Johnson, 1980). Here, the ‘source domain’ provides the topology in which the ‘target domain’ is understood in the emergent blended space.

Fauconnier & Turner (2002) differentiate further between two types of single-scope networks. The first type is not rooted in a larger historical discourse. In other words, there is no tradition in conceptualising CEOs as boxers, and no vital relations link the inputs directly. The other type of single-scope networks is embedded in a diachronic dimension, meaning that the inputs are directly *relevant* to one another and connected via vital relations. The emergent inferences which they yield are consequently much richer than the ones offered by the ‘ahistorical’ first type.

This differentiation notwithstanding, single-scope blends surmise asymmetrical frame clashes in which the frame structure of one input is used to ‘understand’ the other. Nevertheless, single-scope blends do not feel disruptive, as the analogies yielded seem to provide ‘eureka’ effects. This is due to the inner space relations in the organising frame already being tightly compressed and transferred onto the ‘messier’ ones of the other input space. In the boxing CEO example, the boxing frame is already tightly compressed. Causality, temporality, spatiality and identity are clear. Despite this, the CEO input is more diffuse and much harder to grasp. Blending it with the boxing space achieves a reduction to the human scale, as it facilitates apprehension of (certain aspects of) the business world.

At the heart of MSCI are *double-scope integration networks*. These, Fauconnier & Turner claim, are what make humans unique, and lie at the root of cultural achievements. Like single-scope networks, the inputs show clashing frames. Unlike single-scope networks, however, the blended space does not recruit the structure of only one of the inputs. Rather, it includes partial structure of both inputs, and yields emergent structure found in neither of the input spaces (*multi-scope blends* have more than two inputs). Double-scope blending is thus at the very heart of human creativity and imagination, making mathematics, arts and religion possible. To illustrate this point, Fauconnier & Turner suggest that “the history of mathematics shows that the concept of number has been repeatedly revised by creating blends in which we have two (or more) inputs – one with numbers of some kind, the other with elements of some kind” (2002:242). Following this basic idea, complex numbers, fractions or negative numbers can be discussed as double-scope blends. This idea is also pursued

by Lakoff & Núñez when they claim that “many of the most important ideas in mathematics are metaphorical conceptual blends” (2000:48). As blending is regarded as central to human imagination, rituals and magic can, in a similar manner to make-believe play, be described in blending terms. The Catholic Eucharistic rite of Transubstantiation, for example, blends the signifying (the bread) with the signified (Christ’s body). In the emerging structure of the double-scope Transubstantiation blend, the host becomes the Body of Christ, drawing on properties of both inputs: like bread, the host is broken and shared, but it is also Christ (cf. Sweetser, 2000).

To discuss double-scope blending in further detail, consider the expression *same-sex marriage* (Fauconnier & Turner, 2002:134). As inputs, we find here the concepts of traditional marriage in input 1 and homosexual domestic partnerships in input 2. Through selective projection, the elements of wedding ceremonies and tax benefits are, for example, transported to the blended space from the traditional marriage input. From the homosexual domestic partnership space, biological sexes are projected. However, the authors claim that this blend has emergent structure. A novel social structure will be defined through the emergent structure of the blend, awarding homosexual couples the social recognition that they had formerly been denied. Whether or not this blend is actually double-scope or more single-scope in nature remains open for discussion (after all, the organising frame of the blend is firmly structured by the heteronormative concept of marriage alone), as are many of the examples put forward by Fauconnier & Turner (2002).

Simplex networks, mirror networks, single-scope networks and double-scope networks constitute the types of blending identified by Fauconnier & Turner. These types range on a continuum from asymmetry in frame blending to both input frames contributing structure to the emergent blended space. MSCi thus provides a highly detailed and differentiated account of online meaning-making.

2.3. Critical discussion

Although MSCi has enjoyed tremendous success in disciplines as varied as linguistics, literary studies and ethnography, it has nevertheless been subject to a steady stream of criticism. Further to the problems concerning governing principles and blending

typology alluded to above, other aspects of the theory as championed by ‘first generation’ blending analyses have been extensively criticised. This section will summarise and reflect on the criticism levelled against MSCI. It is worth noting that although the section seems to take largely unfavourable stance against canonical blending theory, I would like to stress that the motivation for summarising the discourse is to ultimately further what I consider to be a highly promising approach to a theory of creative on-line meaning-making.

2.3.1. Post-hoc vs. ad-hoc

Like many other cognitive semantic theories – most prominently Conceptual Metaphor Theory (Lakoff, 1986; Lakoff & Johnson, 1980) – MSCI has been accused of delivering *post hoc* analyses only, meaning that it is ostensibly unable to account for actual online meaning construction (Gibbs, 2000; Harder, 2003; Hougaard, 2004, 2005). Starting with the ‘product’ of blending and then working backwards rather than following meaning construction *ad hoc* may lead to the data being tailored to fit the theory rather than the theory being derived from the data. It further leads to a failure in tracking the process of meaning-making as it unfolds in actual on-line cognising. However, advocates (e.g. Coulson & Oakley, 2000; Rohrer, 2005; Talmy, 2000) of the post-hoc approach argue that products constitute the only data currently available to researchers, and that it is impossible to track the psychological steps taken in any particularly accurate manner. Yet this might only be true for approaches focussing exclusively on the individual mind. When considering the interactional dimension of meaning-making, too, as interactional approaches to Cognitive Semantics have done (most notably Hougaard, 2004, 2005), evidence for step-by-step construction, essentially the processes of joint meaning-making over a number of turns, might be gathered (cf. chapter 3.3.).

2.3.2. The Generic Space

The question as to whether generic spaces are psychologically real and necessary for the faithful analysis of blends is suitable to be raised in conjunction with the first point addressed. In this regard, Hougaard (2004) points out that the generic structure proposed only adds abstracts from the input spaces to the network rather than further

semantics as well. Consequently, generic spaces are only required in the analyses of decontextualised, isolated examples that are not embedded in local contexts. In these cases, it is only the post-hoc constructed *tertium comparationis* structure that licenses blends, whereas in contextualised data it is local contexts that sanction, implement relevant blending operations, and guide the structure emerging from the blending process (cf. Brandt & Brandt, 2005). By way of example, imagine you and I are at a party where our friend Achilles, whom we both have not seen in a while, is sporting an impressive full blonde beard. If I now nod in the direction of Achilles and say to you “Oh wow, Achilles is a **li**on!” whilst perhaps gesturing towards my own face, you may not interpret this metaphor as relating to his bodily strength, but to his facial hair. The reason I devise this utterance is not ‘internal’ to the blend, but dependent on what is present in our shared perceptual field and discourse history. The features which both inputs have in common and that may now be depicted in the generic space (in this case facial hair) of this analogous metaphor are consequently dependent on the enunciative context.

Further to this, there are a considerable number of blending operations in which the identification of a shared, generic structure is virtually impossible. With impersonation, for example, humans blend with inanimate entities such as machines. Such impersonations are made relevant by the interactional context rather by what humans and machines have in common.

As such, generic space structure is only required to account for the motivation of conceptual integration in decontextualised analogies that are studied in isolation from a product and post-hoc perspective. However, when investigating non-analogous blends as they occur in their ‘natural’ contexts, it is evident that conceptual integration is motivated and made relevant by the interactional and cultural context in which it is embedded.

2.3.3. Generalisation, terminology and falsifiability

A key goal of Cognitive Linguistics in general and Cognitive Semantics in particular lies in identifying the general principles of human cognition that apply across a wide range of phenomena (cf. Fauconnier, 1999). It thus contrasts with approaches that assume separate facilities for different aspects of cognition, such as a “faculty of language” (cf. Hauser, Chomsky, & Fitch, 2002). This leads to the attempt by

Cognitive Linguistics to attain ‘powerful generalisations’, such as those provided by MSCI. After all, conceptual blending is supposed to capture The (general) Way We Think, encompassing such diverse phenomena as constructions, metaphors, art and mathematics. Although carrying out research as a means of arriving at general conclusions regarding human cognition via the collection of “converging evidence” from a variety of fields is in itself laudable, generalisations also generate numerous complicated predicaments. This includes becoming banal, or rather being unable to provide enlightening insights into specific phenomena and actual human behaviour that is always contextualised (Bache, 2005; Hougaard, 2004), and using terminology that is so broadly conceived that it runs the risk of becoming little more than empty phrases (Brandt, 2010).

With regard to the first predicament, the role of context in which the individual phenomenon is embedded is largely neglected due to MSCI’s subscription to central principles of Cognitive Linguistics; the historical, social and material dimension of cognition and its influence on meaning construction is similarly ignored (see also Harder, 2007; Sinha, 1999). Yet the ‘content’ of blending scenarios largely depends on the situation in which they are embedded, as has already been pointed out above. The emergent properties of a blend will subsequently differ depending on conversational salience, genre and situational relevance. Consider once again the example *Achilles is a lion* discussed above. The input spaces contain the Mental Spaces *Achilles* and *lion*, yet the features that are relevant and therefore guide meaning construction are highly dependent on discourse context. The properties of the emergent structure are consequently determined by interactional context. However, the importance of context in human behaviour is neglected in canonical analyses that hope to attain the most general results possible.

Continuing on from this, such ambitious generalisations as those propagated in canonical blending theory render it difficult to test the theory’s claim of describing the central components of human cognition. Falsifiability will only be possible if MSCI is assessed in its local rather than global dimension (cf. Gibbs, 2000). Instead of discussing eclectic and varied examples in a top-down approach, whereby the data may be adjusted to fit the theory, one type of human behaviour that is assumed to involve conceptual blending should be analysed *in situ* and taking into account local peculiarities in order to provide a better understanding of how blending may – if at all – actually work.

2.3.4. Delineating Mental Spaces

The notion of Mental Spaces is central to MSCI. However, this terminology leaves something to be desired. What exactly are Mental Spaces? What is it that they contain? Why are they ‘spatial’ in nature? Where are their boundaries? What is ‘mental’ about them?⁷ Here as in other areas, Fauconnier & Turner provide little information, and delineating and determining the content of Mental Spaces in a univocal manner is virtually impossible. Some canonical MSCI analyses consider adjective-noun constructions to involve two Mental Spaces, such as Sweetser’s (1999) treatment of ‘red ball’, whereby she contends that the external colour and the shape both reside in different Mental Spaces. Indeed, according to Sweetser, this means that Mental Spaces are primarily perceptual, a claim which leads P.A. and Line Brandt (2005; 2010:280) to enquire about the ‘mappability’ of such spaces; their resulting suggestion is that purely visual spaces containing only colour and shape lack the counterparts required for mappings. In addition, ‘unpacking’ such alleged blends is virtually impossible – how (and why) should we imagine a colourless ball? It thus seems as though issues of qualia are accepted as being on par with semantic structure, once again leaving open the question as to what Mental Spaces really ‘contain’, and which benefits arise from employing a single term for phenomena as varied as colour perception and complex number.

The same applies to simplex blends as proposed by Fauconnier & Turner (2002), which are thought of as bringing together unframed elements with framed spaces. How can elements lacking a frame constitute a fully-blown mental space? After all, other authors assume that Mental Spaces contain entire domains of knowledge, such as Dancygier’s (2008) story worlds. What is it, then, that all of these alleged Mental Spaces have in common? Is it the unframed elements of simplex blends, the colours, the grammatical constructions, the rich story worlds found in novels? All in all, it seems that first generation MSCI analysts largely follow their intuition alone when attempting to depict the content of Mental Spaces. Accordingly, the question as to what Mental Spaces contain and how they are delineated (their ‘spatialness’) should definitely be of central importance to a refinement of blending analyses.

⁷ See Hougaard (2004), P.A. Brandt (2005) and L. Brandt (2010:247ff) for a detailed discussion of this point.

2.3.5. Whose Mental Spaces? Whose blends?

The notion of Mental Spaces is not only problematic when it comes to defining their ‘boundaries’, but also when considering the cogniser. Whose Mental Spaces are blending analysts really depicting when offering a detailed diagram of meaning-making processes in idioms such as the “grim reaper”? Is it the Mental Spaces of the original producers or those of today’s recipients? Alternatively, are they rather modelling their own cognising processes when interpreting a given artefact? It therefore seems as if many blending analyses assume an ideal cogniser and identical processes for the production and construal of cognitive artefacts, such as poems, jokes or metaphors (cf. Rohrer, 2005). Yet it is highly unlikely that the same mappings and blendings should have led, for example to the production of a novel and the various readings it generates among the audience. It should therefore be made explicit as to exactly what it is the analyst is depicting, or rather whose Mental Spaces and mapping processes are being investigated.

2.3.6. Loci and levels of mappings and blending

In connection with these prior points, further questions have been raised as to whether blending is really the feature at work in the majority of the examples discussed by canonical blending analysts – these examples range from sexual fantasies, phantom limbs, complex numbers, cultural rituals and personal identity to compounds, syntax and visual perception. Is it really the same cognitive process of conceptual integration that works across all of these different phenomena?

First of all, ‘dead’ metaphors and idioms such as “to dig one’s own grave” or the “grim reaper” are discussed as being blends without any further explanation as to whether it is the original mappings that led to the establishment of the symbol or the still valid processes of meaning-making that are being discussed. In other words, the question is left open as to whether we unpack the “grim reaper” every time we encounter the expression. Should it really be assumed that cogniser unpack such ‘tightly integrated’ phrases every time they encounter them, the authors seem to neglect both etymology and parsimony by claiming that idioms are semantically analysed when interpreted. This would constitute a severe neglect of the psycholinguistic discussion surrounding the semantic analysability of idioms (cf.

Dobrovolskij, 2011 for a current treatment). Even if MSCI aimed only to account for the processes that led to these expressions becoming symbols, they failed to take into account the diachronic development which these expressions underwent in becoming fixed expressions or the cultural narratives on which they rely. In short, the question regarding whether the daily usage of such fixed expressions as the “grim reaper” requires the highly complex blending processes (cf. Gibbs, 2000) suggested by Fauconnier & Turner remains unclear.

Secondly, blending may not actually play a major role in meaning construction on the grammatical level (so-called grammatical blends), such as in Goldberg’s (1995, 1999) much discussed “She sneezed the foam off the cappuccino”. Here, (in the blending analyses carried out by other researchers, not in Goldberg’s own work) the syntactic structure, a cause-motion construction, is thought to be blended with semantic structure, rendering it a caused-motion event (cf. Mandelblit, 1997; Mandelblit & Fauconnier, 2000). However, research in language acquisition has shown that children up to the age of four are unable to simultaneously bear two different images in mind (Harder, 2003). Nevertheless, the juggling of distinct Mental Spaces lies at the heart of blending theory, and by the age of three children have generally mastered large portions of their mother tongue’s grammar. Why then, Harder (2003) convincingly asks, should we assume that conceptual blending, the entertaining and juggling of distinct Mental Spaces, is at work in the processing of linguistic constructions?⁸ He concludes that “space building [is] something sophisticated that occasionally interferes with the ordinary default assumption [...]” (2003:93). Contrary to Fauconnier & Turner’s (2002:89) claim that “fundamental cognitive operations are the same across radically different phenomena”, Harder (2003) considers conceptual blending to function only when distinct epistemic domains whose relationships are not conventionalised are entertained simultaneously. This point is associated with questions ruminating on the nature and content of Mental Spaces that are largely missing in canonical blending theory.

⁸ Cf. Brandt (2010) for a more theoretical discussion of problematic aspects in grammatical blends.

2.3.7. Homunculus entering through the stage door

MSCI hopes to shed light on invisible ‘backstage’ cognition through ‘network’ analyses. The theory thus draws heavily on established philosophical metaphors of stages and net(work)s and leaves open the question regarding what exactly it is that differentiates ‘onstage’ from ‘backstage’. What is it that selects a given integration network from the many that are “attempted and explored in an individual’s backstage cognition” (Fauconnier & Turner, 2002:309)? The authors deal with this question briefly, stating that “the nature of consciousness [is to give] us effects we can act on, and these effects are correlated with the unconscious processes” (ibid:56). In other words, “the moment of tangible, global understanding comes when a network has been elaborated in such a way that it contains a solution that is *delivered* to consciousness” (ibid:57; my emphasis, VS). But who/what is the agent delivering, and to whom is it being delivered? As consciousness and unconsciousness are not discussed in detail (only approximately a page is dedicated to the matter) and lack reference to philosophical or neurobiological discourse, one is left to wonder whether the notion of ‘backstage’ cognition in MSCI might after all involve a Homunculus translating the ‘backstage’ to the ‘onstage’, “selecting” Mental Spaces and blends from the “bubble chamber of Mental Spaces” (ibid:321) that is our brain. The question of what ‘offstage’ and ‘onstage’ and the notion of the brain as a ‘bubble chamber’ really mean should consequently be dealt with explicitly in blending research in order to avoid the pitfalls posed by implying that there is a Homunculus acting as a ‘stage manager’.

2.3.8. Blending in context, blending in interaction

Ever since Wittgenstein (1953), it has been widely acknowledged in non-formal linguistics that all human cognising is situated in social, historical and material contexts. Meaning emerges in actual usage. If blending is “the way we think”, then it should be able to successfully account for or at least acknowledge situated cognition. In this respect, Fauconnier (1999:97) openly states that “introspection and intuition are woefully insufficient to tell us about general operations of meaning construction. [...] It is only in rich contexts that we see the full force of creative on-line meaning construction.” Yet despite Fauconnier’s own assessment and the ‘social turn’ in Cognitive Linguistics called for by Sinha (1999), Croft (2009) and Harder (2010),

canonical blending theory still largely neglects the interactional grounding and context-sensitive dimension of meaning-making, and instead largely focuses on seemingly isolated individuals experiencing seemingly isolated and eclectic artefacts, such as images, poetic narrative texts, or advertisements. These singular tokens are, in addition, analysed from a still somewhat subjectivist, atomistic and mentalist ‘armchair’ perspective, relying predominantly on introspection alone rather than psycholinguistic or empirical (including phenomenological) approaches to linguistics. Canonical blending theory is therefore rooted in the methodology of analytic philosophy, focusing more on competence than on performance (cf. Brandt, 2010; Brandt, 2005). As such, blending analysts provide highly detailed descriptions of the mappings and projections occurring in conceptual integration processes, avoiding what Harder (2007:1257) calls “the question of what the actual processes are whereby complex linguistic utterances are assigned meaning in actual interaction.” In consequence, the issue regarding how the various operations described produce the different meanings that may arise from one and the same expression (as is evident, for example, in misunderstandings or varying interpretations of novel compounds, metaphors or jokes) is at best left implicit in their discussion of decontextualised examples. However, all cognising is situated and meaning construction based on shared epistemological and normative aspects of the discourse community.⁹

Consider the Achilles example discussed above. Here, I would need to ensure that you knew what I was referring to by conceptualising our friend Achilles as a lion. I would consequently need to ensure that his beard is (perhaps perceptually) salient to both of us, and would probably also have to consider our respective attitudes towards his newly grown facial hair. I would then probably nod in his direction or point towards his or my face in order to further clarify what I was referring to, and, depending on my feelings on the beard, assume an ironic or impressed intonation pattern. In this fashion, I would provide cues, or *ostentive stimuli* (cf. Sperber & Wilson, 1986), for you to make the relevant mappings, decompress the blend and ultimately generate pragmatic inferences that are based on your assumptions about my assumptions. As even Fauconnier (1997:160) *en passant* points out, this goes to show that blending in interaction first of all demands for *coordination* to facilitate

⁹ Cf. the problem of Other Minds (‘Hume’s problem’) that topicalises the question of how ‘words as ideas’ can be intersubjectively shared (Sinha, 1999); it was originally devised to address problems evident in Locke’s ideational theory. MSCI faces a similar problem, as it has so far failed to account for the ways in which Mental Spaces and blends can be shared intersubjectively.

intersubjectively shared meanings. Secondly, the emergent meanings are not only determined by mappings alone, but also by our conceptualisation of the action underway, our joint beliefs and values. As Per Aage Brandt (2005:1584) succinctly puts it, in blending theory “[...] if a joke is analyzed, its mappings and blends can be modelled, but not the reason why it is funny.” Blending in interaction and the meanings emerging from blending processes are thus subject to negotiation and the consideration of shared values, knowledge and beliefs to arrive at roughly similar constructions; the mappings occurring in blending are an interactional achievement.

The interactional business required to coordinate meaning construction and the role played by conceptualisations and co-constructions of context, is, however, not the focus of most canonical blending research, which in this regard seems to still be rooted in more formal approaches and tends to place almost exclusive focus on grammatical space builders blending-internal governing principles. Eve Sweetser (2007) has recently turned to the question regarding how gestures could indicate mental space configuration; however, she too has neglected the question as to how such signalling and the situation in which signals are embedded may impede on joint meaning construction. However, researchers such as Anders Hougaard (2004, 2005, 2008), Esther Pascual (2002, 2009), Robert E. Williams (2004, 2008), Todd Oakley (1998; Oakley & Coulson, 2008; Oakley & Hougaard, 2008), Ed Hutchins (2005), Per Aage and Line Brandt (Brandt, 2010; Brandt & Brandt, 2005; Brandt, 2005) are hoping to shed new light on the interface of cognition and interaction, and provide promising insights into potential conceptual reconfigurations of MSCI.

2.3.9. Blending typology

Fauconnier & Turner (2002) differentiate between four types of blending, ranging on a continuum from simple to highly complex frame blendings. Yet the blending typology put forward by the authors is perplexing in a number of ways. Simplex networks are supposed to be the most basic type of blending, linking unframed roles to framed values in another frame. But what is it that input spaces contain if not frames? Why is it that a mental space can only contain unframed elements and the other nothing but frame structure? This is especially peculiar given that Fauconnier & Turner (2002:102) define Mental Spaces as “partial assemblies containing elements, *structured by frames* and cognitive models” (my emphasis, VS). How, then, can simplex

blends involve two Mental Spaces if one input does not show frame structure and cannot be regarded a mental space in the definition proposed by the authors themselves?¹⁰

At the heart of blending theory are double-scope blends, which reside on the most complex end of the continuum from single- to double-scope frame blending. However, single and double-scope blending remains paradoxical in a number of ways too. Consider the computer desktop example that Fauconnier & Turner (2002:131) present as their first (and, as I therefore consider it, prototypical) illustration of double-scope blending. Here, the familiar frame of office work with its folders, bins and files is blended with computer commands such as ‘save’ or ‘print’. The authors here claim that both frames contribute to the emergent structure of the desktop metaphor. I would, however, claim that the desktop metaphor is not a prototypical double-scope blend, as strong asymmetries can be witnessed.

The desktop metaphor was first used in the 1960s and 1970s when the graphical user interface was developed at Xerox PARC. It was later popularised by Apple Macintosh in the 1980s as a means of facilitating the use of computers for the layperson. Rather than typing in complex and abstract commands, users could now perform routines they were already familiar with – filing forms in respective folders, disposing of documents by putting them in the bin, copying documents by printing etc. Consequently, the established office space structure exerts much more influence on the emergent properties of the blended ‘desktop’ space than the computer input. It is an analogy through which the complexity of computers is reduced to more familiar practices. In line with theories of remediation (Bolter & Grusin, 1999), the computer desktop interface pays homage to established work routines in order to facilitate understanding, and is, in my view, therefore more of a single-scope blend, in that a familiar and highly compressed frame is used to shed light on more complex and less familiar routines. With the introduction of the desktop metaphor, working with computers was made possible. We understand computers in terms of established work routines¹¹, and the emergent space in this blend is consequently predominantly structured by the office input. It yields a sense of immediacy (cf. Bolter & Grusin,

¹⁰ Cf. Brandt (2010:253ff) for an exhaustive discussion of this aspect.

¹¹ Once the basic desktop metaphor became established, it was able to be used to develop novel characteristics, such as having numerous windows open simultaneously. However, this is a step-by-step process (the evolutionary dimension of meaning construction) that Fauconnier & Turner (2002) do not discuss in their analysis.

1999), and backgrounds the media it uses, the computer, to such an extent that it almost becomes invisible. The same is true for the same-sex marriage example discussed in detail above. This blend is also structured mainly by one of the two inputs. As such, it remains questionable as to whether the desktop interface and the same-sex marriage blends can thus faithfully be described as a fully-blown double-scope network. It should be noted that Fauconnier & Turner (2002) point to a category of highly asymmetric double-scope networks, yet seem to fail to mention examples of non-asymmetric double-scope blends.

Furthermore, Brandt (2010) points out that it is somewhat “perplexing” (251) that metaphors in Fauconnier & Turner are discussed within a variety of blending types: from XYZ-metaphors (such as the much discussed “vanity is the quicksand of reason”), which are thought of as simplex networks, to single-scope networks as epitomised by certain conceptual metaphors, and on again to double-scope networks such as the grave-digging idiom. Why, Brandt convincingly asks, should conceptual metaphors result in less advanced emergent structure than idioms such as “to dig one’s own grave” or the “grip reaper”, which are analysed as double-scope blends? The problem seems to me to be one of frame structure level. Depending on the level of detail and the aspects highlighted in the analysis, networks can either be classified as single or double-scope. When approaching this issue using a more interactionist theory of metaphor (Black, 1962; Ricoeur, 1975) instead of Conceptual Metaphor Theory (CMT), for example, metaphors can easily be interpreted as double-scope rather than single-scope in Fauconnier & Turner’s framework. Interactionist metaphor theories emphasise the pragmatic dimension of figurative language, and discusses metaphors in relation to the discourse they are set in, thereby focussing on the ever-contingent nature of meaning-making. The cognitive effort required to consciously make sense of non-conventional metaphors (i.e. language usage that poses some kind of dissonance or a violation of expectations in the current discourse setting) requires input spaces to be played with simultaneously, for one input space to be foregrounded and then another. Focussing on evolving processes of sense-making in non-conventional metaphors, then, foregrounds the double-scope nature of such expressions. CMT, on the other hand, involves a one-way mapping, as it regards metaphor as understanding one domain in terms of another. This theory is naturally prone to understanding metaphor as involving single-scope blending. The blending typology proposed is consequently top-down rather than bottom-up, and the

examples can easily be construed to fit either category of the classification. What is required, then, is a methodology that is able to empirically account for typological classifications.

2.4. First conclusions

MSCI has proven to be a highly attractive theory with which to model human cognising. Yet it suffers from a number of weaknesses, most notably ones of a methodological nature, including a certain disregard for philosophical grounding and the lack of consideration it shows for the contextual, social, historical, material and sometimes even psychological dimension of meaning-making. The following thus sets out to tackle these issues by discussing two main research questions:

- 1) Is the blending typology presented by Fauconnier & Turner (2002) cognitively real? Can a distinction really be drawn along the lines of asymmetry and the amount of structure projected from the input spaces? If we blend at all, do we orient ourselves towards the continuum suggested by the founding fathers? Is there actual evidence for double-scope blending (i.e. the aspect determining our very humanity according to the founding fathers) in *in-situ* cognition?
- 2) How do we, as both an *animal socialis* and *animal cogitans*, employ conceptual blending in our everyday interactions, in our localised, socially embedded reasoning, to arrive at construals that are experienced as meaningful?

To address these two aspects, a revised approach that allows for the explicit treatment of interactional and contextual dimensions of meaning through conceptual blending will be necessary. This interactional approach will first of all require terminological clarifications in order to equip MSCI with the tools necessary to deal with ‘real-life’ phenomena. It will secondly necessitate clearly defined and socially contextualised data that is analysed according to rigorous methodological standards in order to allow the interactional *processes* of meaning-making to be focussed upon. To this end, only novel ad-hoc blends (i.e. those resulting in discursive dissonance) will be investigated in environments in which two clearly distinct scenarios, semiotically speaking, that

contain rich semantic structure are *deliberately* and hence *consciously* recruited. For this reason, I will concentrate on non-routinised and explicitly incongruous conceptual blends involving interactionally clearly signalled representational, encyclopaedic theatrical ‘worlds’ which are epitomised by impersonations. In impersonations, a speaker takes on the voice of another character. For this voice-change to be successful, the two scenarios, that of the impersonator and the impersonated, are explicitly indicated by the performer to be separate. In the performance, the ‘worlds’ are maintained simultaneously, yet are brought to a point of interaction in the integrated performance. When decompressing the performance, emergent properties arise, the nature of which is interactionally negotiated.

With this in mind, the following will first of all propose a methodology which will allow the negotiation of conceptual blending in interaction to be studied accurately. It will secondly discuss the theoretical and terminological reconfigurations required when ‘marrying’ a largely ideational theory such as MSCI to a usage-based theory of meaning as proposed by most discourse analysts and followers of ordinary language philosophy.

CHAPTER 3: PUTTING THEORY TO THE TEST: TOWARDS A SOCIAL-INTERACTIONAL COGNITIVE SEMANTICS

3.1. The State of the art of empirical and experimental research within MSCI

Research in Cognitive Linguistics and Cognitive Semantics is traditionally based on introspection accompanied by theoretical analysis (Talmy, 2007). However, this approach has come under attack for various reasons, most notably for primarily attending to the mind of the trained researcher, for being too “solipsistic” and “too much ‘inside the head’” and therefore remaining incomplete (Croft, 2009:396), and for lacking “solid empirical evidence” (Kok & Bublitz, 2011:291). Language is, after all, first and foremost a means of social interaction and not an exclusively cognitive ability that exists in isolated minds.¹² Consequently, a great number of workshops (such as the Empirical Methods in Cognitive Linguistics (EMCL) series) and publications (e.g. Gonzalez-Marquez et al., 2006; Oakley & Hougaard, 2008) have recently attempted to further empirical research within Cognitive Semantics and find methodological alternatives to introspection, such corpus-based research, computational and algorithmic methods, audio and videographic analysis, and neurological or experimental studies in psycholinguistics; this has certainly enriched the portfolio of empirical methods in Cognitive Linguistics.

Of the approaches proposed and largely on the basis of Hougaard (2004, 2005, 2008), this project will investigate conceptual blending from the perspective of ethnomethodological Conversation Analysis, which allows for processes of ‘meaning-making’ in naturally occurring talk-in-interaction to be systematically studied. This approach is favoured over other approaches for a number of reasons that I wish to elucidate on briefly in the following; in doing so, I will draw on the results of previous research.

¹² Cf. Wittgenstein’s (1953, §243-271) discussion of the impossibility of a “private language” that would lead to its signs becoming meaningless for both the originator and other humans. Language is social. There is no meaning without embedding signs in interaction.

First of all, while a quantitative study based on extensive corpora would to a certain degree be able to address the subjectivity problem encountered by introspective methods, major corpora, such as the *International Corpus of English* (ICE), are available in written, transcribed form only, and are hence not able to fully cover the multimodal expressive and emotive aspects of interaction. As a result, they are neither appropriate for providing insight into the multidimensional aspects of talk-in-interaction nor its timing nor epistemic dimensions, elements which are of crucial importance when investigating processes of meaning construction in interaction. Furthermore, as I am not particularly interested in the frequency of a certain phenomenon, but more in the fine-grained mechanisms of interaction, a qualitative approach is in required.

Computational approaches have been used in previous attempts to validate MSCI (e.g. Veale & O'Donoghue, 2000), testing the computational validity of blending theory with the help of algorithmic models. Although such approaches can show that blending is a “computationally tractable process” (ibid:254), they fail to take account the ‘human component’ in blending. If we wish to find out if and how social beings ‘blend’ in their everyday lives and interactions, our interest should lie on humans.

In this regard, neurological studies employing ERPs (event-related brain potentials), PET (positron emission tomography), or fMRI (functional magnetic resonance imaging) scans and similar technology initially seem to be the obvious choice for empirically validating blending theory, especially as blending has been linked to neural binding processes.¹³ Here, changes in activity are monitored in different brain regions when carrying out certain tasks, such as manipulating linguistic representations; this therefore makes it possible to draw conclusions as to brain cell activity during specific tasks. Studies utilising ERPs have been used in metaphor and MSCI research to measure processing times (Coulson, 2007). Although this approach is certainly fascinating and can yield interesting results, it suffers from the fallacy of providing subjects with decontextualised, limited stimulus. It is therefore unable to account for the intricate interweaving at play in meaning construction processes in

¹³ Neural binding theory (or: ‘the problem of neural binding’) refers to the mechanisms by which we are able to construct a single, coherent object from distinct neural representations (such as colour and shape of an object). These perceptual integration mechanisms, the theory assumes, are possible because of the “synchronized firing of distinct neural regions” (Grady, 2000:337). Although no consensus yet exists as to the particulars of neural binding, it seems uncontested that a certain type of integration process must be at work.

actual interaction. Furthermore, fMRI technology currently seems unable to provide images of the brain which are fine-grained enough to support the idea of blending as a process of neural binding.¹⁴

In an approximately similar vein, experimental methods favoured by psycholinguistics and cognitive psychology are promoted in Cognitive Semantics. This approach relies on confronting subjects with pre-designed stimuli devoted to a specific cognitive factor. By monitoring responses, these studies hope to draw inferences as to the workings of this factor (Talmy, 2007). Although controlled experiments allow fine-grained mechanisms to be focussed upon, isolated and extracted, these mechanisms are only rarely found in actually occurring, non-experimental settings. In addition to this, the extraction of decontextualised factors might lead to the inability to perceive the ways in which these factors interact with others. While psycholinguistic approaches garner fascinating insights into, for example, the conceptualisation of fictive motion, it is not suitable for investigating the much ‘messier’ (A. Hougaard, 2008; Rohrer, 2005) contextualised and interactional sense-making processes found in naturalistic data.¹⁵

When investigating blending-in-interaction, then, the most useful methodology will be the systematic study of audiovisual material, as it allows for the crucial, interwoven factors of talk-in-interaction, such as vocal dynamics, body language and gaze in meaning-making processes, to be focussed upon. Examining these factors to the maximum is only possible when interactants engage in actual, uncontrolled speech events (Talmy, 2007). In order to determine whether people in actual interaction orient to the blending typology advocated by Fauconnier and Turner (2002) and to successfully address crucial points of the criticism levelled against it, it is therefore necessary to employ a methodology that allows the structures of naturally occurring interaction to be studied whilst abstracting away from idiosyncratic behaviour.

3.1.1. Discursive approaches to MSCI

Discourse Analysis with its various sub-disciplines constitutes an approach which is able to account for the structures behind actual interaction. In the past, distinctly

¹⁴ It should also be noted that the ‘neural’ approach recently favoured by Fauconnier and Lakoff (fthc.) and Lakoff (2008) is not related to neurolinguistic research, but based solely on computer modelling.

¹⁵ See also Turner (2010) for ideas on experimental studies in MSCI.

discursive approaches have been employed by a small number of scholars conducting research within an MSCI framework. An excellent collection of papers employing this approach can be found in Oakley & Hougaard (2008). As well as providing insights into the ways that MSCI can successfully be applied to draw attention to the ways in which interactants understand discourse, this volume also highlights the revised theoretical configurations arising from the fine-grained analyses of situated interaction.

These discursive approaches to MSCI have provided a number of important discussions concerning the theoretical reconfigurations of MSCI theory. Most prominently, they contribute the notion of a *semiotic base space* that serves as a basis for blending processes as proposed by Brandt & Brandt (2005). Brandt & Brandt's influential base space is, however, not to be confused with Fauconnier & Turner's (2002) notion of a base space, which denotes that the space serves as a basis for other spaces which are set up as 'discourse' unfolds. This original base space contains assumptions about 'reality', and all other spaces are understood in relation to this base space. In addressing this, Brandt & Brandt (2005) claim that Fauconnier & Turner's base space is *ontological* while their base space is *semiotic*. Brandt & Brandt's (2005:224) discourse base space consequently subsumes "the space in which utterances are uttered and come to mean whatever it is they are supposed to mean." Akin to Langacker's (2001) Current Discourse Space, it therefore subsumes the cogniser's representation of the situation in which utterances occur, and consequently introduces issues of context into blending. Brandt & Brandt (2005) assume three spheres of the discourse base space: the *expressive act* (the actual utterance, such as "Achilles is a lion"), the *relevant circumstances* in which the utterance occurs and the way it is framed by the interactants (such as a party setting in which our friend Achilles wears a beard), and the *human phenomenological life-world* (the world that surrounds us as it is available to, and might become part of, human cognising). All spaces comprising a mental space or blending network are constructed on the basis of the discourse base space. Figure 8 provides an illustration of Brandt & Brandt's (2005:239) model.

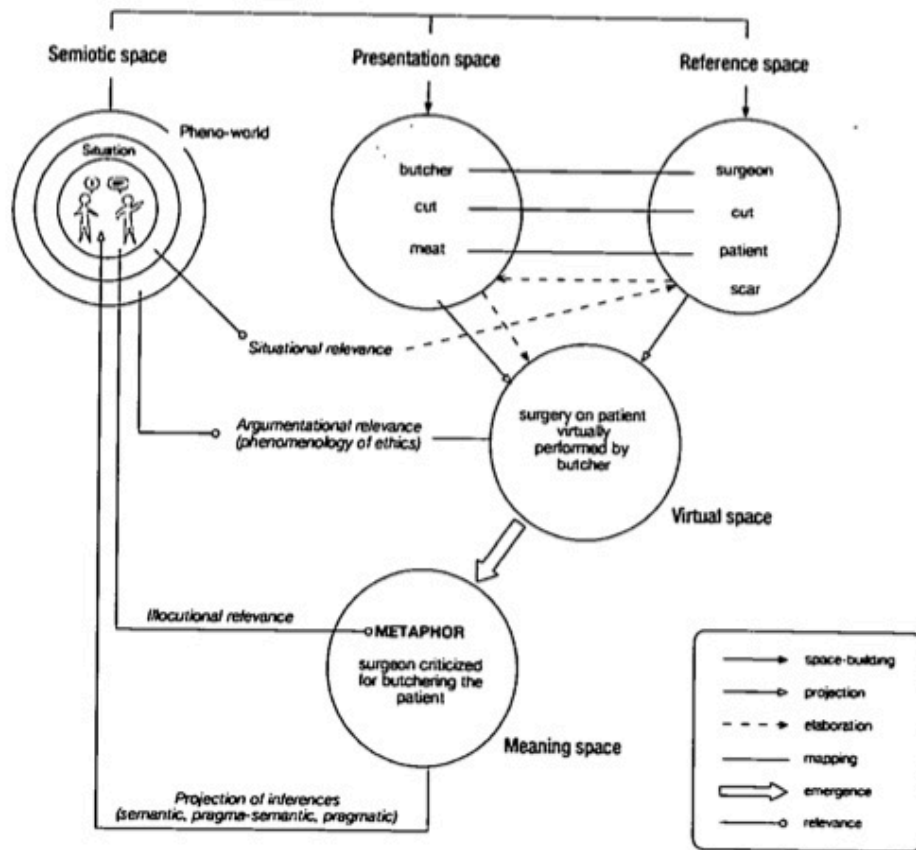


Figure 8: Brandt & Brandt's (2005:239) discursively-grounded model

Nevertheless, despite providing relevant conceptual reconfigurations, discursive approaches to MSCl have so far failed to provide solid empirical evidence that blending is really 'what people do' (Hougaard, personal communication). Accordingly, when wishing to determine the extent interactants to which display an orientation to blending processes in interaction, a research programme is required that focuses rigorously on the structures of interaction and the endogenous understandings displayed by interactants. Ethnomethodological Conversation Analysis (EMCA) is one such approach that allows for the detailed analysis of social interaction.

3.1.2. EMCA-inspired approaches to MSCl

An EMCA approach to studying processes of conceptual integration was previously most notably championed by scholars such as Anders Hougaard (2004, 2005, 2008) and, to a lesser extent, Robert Williams (2004, 2008). Williams deals with the

“relations between material structures and conceptual operations in the everyday activity of time-telling” (2004:xiii). Building on Hutchins’ (1995) concept of “distributed cognition”¹⁶ and using CA-inspired methodology and video analysis, he investigates how meaning is constructed in elementary school time-telling lessons through “guided conceptualisations” within the framework of MSCI. Williams (2004) here identifies “material anchors”, such as gestures, for conceptual blending operations (cf. also Hutchins, 2005) and develops a model of meaning construction in situated activities that combines the concept of distributed cognition with MSCI theory. Although he then provides interesting and relevant suggestions for the advancement of MSCI, Williams is not primarily concerned with the validity of MSCI theory or the question as to whether blending operations can be traced in actual interaction. Still, he proposes that cognitive ethnography could also be used to “investigate whether the conceptual mappings hypothesised by cognitive linguists are apparent in real-life instances of meaning construction” (2004:26). Anders Hougaard’s work is not primarily interested in the validity of the theory, either. Hougaard (2004) investigates compression in actual interaction, proposing an interactional approach to conceptual blending, but does not tackle issues in blending taxonomy or place explicit focus on optimality principles in interaction. Rather, his results indicate that there are discrete cognitive operations within conceptual blending in interaction, such as specific subtypes of conceptual disintegration, ‘splitting’ and ‘partitioning selection’ (see also Hougaard, 2005). As a result, his interactional approach offers interesting and highly relevant suggestions (such as the concept of ‘onstage cognition’ and the development of an interactional approach to conceptual integration), but is not concerned with the validity of blending taxonomies in interaction.

¹⁶ Distributed cognition refers to the notion that complex human processes of reasoning and learning are to a certain degree dependent on their social and material contexts. Cognition is, in short, “distributed among people and their artifacts” (Giere & Moffatt, 2003:303). The analysis hence goes beyond the individual mind and instead takes into account “entire cognitive functional systems” (Williams, 2004:14) as a means of determining how individuals achieve a cognitive task via interaction with artifacts and other functional systems. Cognitive processes are understood to be distributed in three ways: as distributed across various representational media (such as calculators or maps), as across social groups (with experts in varying, interrelated fields), and diachronically (with older cognitive processes impacting on present ones).

3.1.3. Studies on blending taxonomy

In comparison to the above scholars, an interest in blending taxonomies is of interest to Bache's (2005) and L. Brandt's (2010) work. Bache (2005) proposes a revised blending typology that he regards as complementary to Fauconnier and Turner's (2002) model. The typology consists of three main blending types, which Bache refers to as first-order, second-order and third-order blends (Bache, 2005:1619). These blending types are organised according to "complexity and sophistication". For Bache, first-order blends are, from a psychological perspective, binding: "instances of basic mental compression and unification of complex perceptual experience" (ibid). Second-order blending in Bache's model subsumes "the integration of basic abstract stories with abstract grammatical structures to produce actual grammatical constructions" (ibid:1622). The most complex blending type in Bache's taxonomy are third-order blends that "integrate more clearly separate and independent Mental Spaces" (ibid:1625). As I refrain, for reasons outlined above, from using the powerful terminology offered by blending theory as a means of account for issues of neural binding and grammatical constructions, it is in my view Bache's third-order blends alone which constitute 'proper' blends involving fully framed Mental Spaces; this is because only these frames require the simultaneous juggling of separate mental scenarios. Bache consequently subsumes those concepts which Fauconnier and Turner classify as mirror, single-scope and double-scope blends under the umbrella term Third-Order blends, and does so without further distinguishing between various levels of frame integration in this rather large and, to my mind, central group. For these reasons, I will refrain from taking Bache's (2005) terminology into consideration any further.

Blending typology is also at the heart of Line Brandt's (2010) model, in which she proposes a conclusive revision of blending typology within a cognitive-semiotic framework. She makes a strong case for the importance of situational context in meaning-making, and bases her elaborate taxonomy on enunciative aspects. In a similar manner to Bache (2005), she differentiates between Mental Spaces Blends that employ fully-fledged scenarios, and intraspace integration and conceptual fusion, which do not contain across-space mappings and would correspond to Bache's first- and second-order blends. In contrast to Bache, however, she identifies a considerable variety of Mental Spaces Blends (Bache's third-order blends); she groups these Mental

Spaces Blends in minute detail according to enuciative factors. Figure 9 visualises her typology.

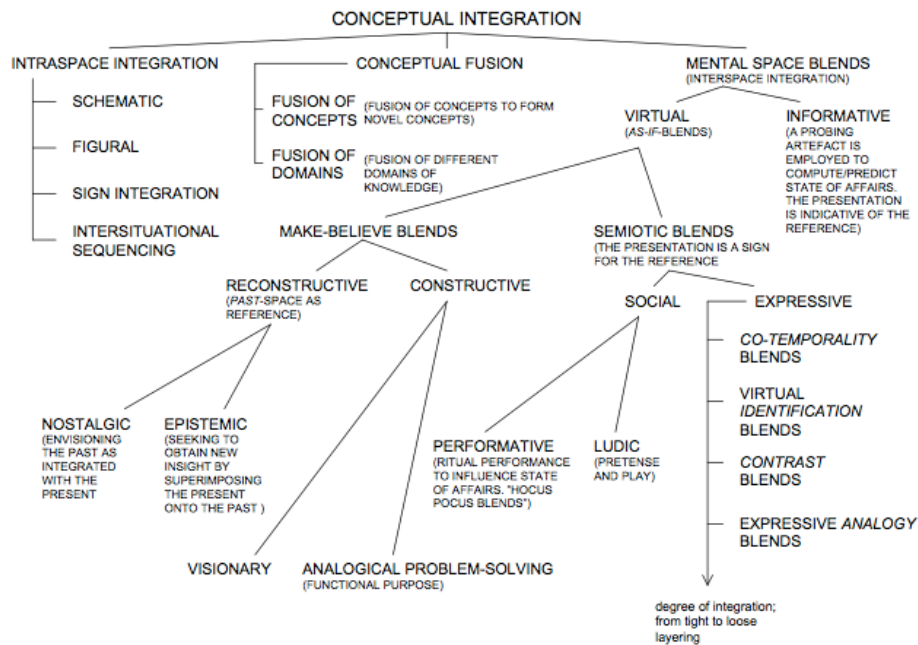


Figure 9: Brandt's (2010:321) blending typology

The issue with Brandt's highly detailed study is that it mainly relies on philosophical arguments rather than empirical data. The question as to whether interactants might after all *empirically* orient to the differentiation into mirror, single-scope and double-scope blends as put forward by canonical blending analyses is consequently left unanswered.

The following will now serve as a brief introduction to Ethnomethodological Conversation Analysis (EMCA) by first of all discussing its philosophical antecedents and then considering the consequences caused by the 'marriage' of Cognitive Semantics to this phenomenological approach. I will also introduce the reader to the data set and the methodology utilised.

3.2. Methods for a socially grounded MSCl: Conversation Analysis

Conversation Analysis (CA) is a systematic and rigorous methods for studying the orders of spontaneous, naturally occurring talk as well as the production and

understanding of social (inter)action in general, and stresses the essentially social role of language use (Drew, 2006).¹⁷ Adopting the basic assumption that ‘meanings’ are interactional, practical and social achievements (Hutchby & Woffitt, 1998) that are based on shared normative procedures (Ten Have, 1999), CA is now widely used in various disciplines within the social and cognitive sciences, such as social psychology, anthropology, discourse analysis, pragmatics and, of course, its ‘home discipline’ of sociology.

One of CA’s main tenets is to study and explicate those features of language use in the broadest sense that are demonstrably relevant to meaning production among participants in social action. Only those aspects of interaction that participants themselves evidently orient to in their interaction are considered to be valid and interactionally ‘real’. In Ethnomethodology (EM), from which Conversation Analysis (CA) derived and which provides CA’s main programmatic background, implies that construction cannot be separated from action. Meaning, and therefore meaningful events, does not originate in people’s heads, but, in the (rather polemic) words of Harold Garfinkel (1963:190), “are entirely and exclusively events in a person’s behavioural environment [...]. Hence there is no reason to look under the skull, for nothing of interest is to be found there but brains.” As a result, meaning allocation is not seen as an “inner private act of consciousness”, but rather as a “social, ‘public’ event” (Bergmann, 2004b:75). Contrary to trends in present-day Cognitive Linguistics (and consequently in MSCI, EM and CA), do not attempt to reconstruct “backstage” (Fauconnier, 1997) cognitive processes that lead to ‘internal’ understanding; rather, they aim to explicate the structural processes which lead to shared understanding as manifested in social action. Accordingly, CA may constitute an interesting starting point for investigating the relationship between individual and social cognition.¹⁸

Before delving into the analysis itself, it would be sensible to introduce Conversation Analysis in further detail. The following will outline the key assumptions

¹⁷ CA systematically studies everyday talk as situated in human interaction. In order to encompass other, non-verbal, yet relevant aspects that impact on the organisation of human social (inter)action (such as gaze or body alignment), the object of study is thus commonly referred to as *talk-in-interaction* rather than conversation (Hutchby & Woffitt, 1998). In this vein, linguistic units are understood to be objects that serve to methodically accomplish an action, such as apologising or complaining, rather than semantic units.

¹⁸ Although both cognitive linguists and conversation analysts often fall prey to certain prejudices with regard to the respective other ‘camp’ (see Hougaard & Oakley (1997:16-17) for an overview), rewarding links to cognitive sciences that perceive cognitive processes as ‘distributed cognition’ (see, e.g. Coulson & Matlock, 2005; Hutchins, 2005) have already emerged.

subscribed to by Conversation Analysis as a research programme by first of all briefly introducing its development and philosophical underpinnings and, in a second step, summarising central findings that are relevant to the project at hand.

3.2.1. CA: development and philosophical underpinnings

Conversation Analysis first emerged in the late 1950s and early 1960s, when Harvey Sacks and Emanuel Schegloff began to develop a new method of ‘doing’ sociology.¹⁹ Sacks and Schegloff established a methodology for analysing social interaction *in situ* and *in vivo*. This approach was to be methodologically sound and capable of attaining replicable results as a means of making sociology a “naturalistic, observational science” (Hutchby & Woffitt, 1998:24). In developing what was to become CA, Sacks, Schegloff and their fellow first generation conversation analysts (such as Gail Jefferson) were influenced by a number of researchers, schools and philosophies. As such, Erving Goffman’s and Harold Garfinkel’s work in particular constitutes the philosophical backbone of CA.

Starting in the 1950s, Erving Goffman began to develop a sociological framework that focused on what he later called the *Interaction Order* (1983) of naturally occurring, everyday, mundane human interaction. Goffman’s approach was essentially a qualitative one based on eclectic observations he had made and which he then used in order to arrive at and demonstrate a conceptual system. His conceptual studies drew from a variety of disciplines, such as drama studies and game studies, and showed the great sociological relevance that can be found in everyday face-to-face interaction. Such naturally occurring face-to-face interaction, Goffman maintained, was a locus where social order could be uncovered with the help of structural sociological analysis. His interest in the mundane, in non-experimental and naturally occurring interaction and his sequential analysis of the ‘interaction order’, with which he tried to recreate the sequence of events of an interaction, is reflected in CA research.

¹⁹ It should be noted that their interests lay in sociological problems and not so much in studying language. However, from its beginnings CA has been regarded as conducive to the study of language usage and human communication, and provides an interdisciplinary interface between sociology and linguistics. Following linguistics, CA understands language to be a structured system, and – in accordance with linguistic pragmatics – a means for communicative interaction (Hutchby & Woffitt, 1998; Ten Have, 1999). As in sociology, CA perceives interaction as a social process.

Harold Garfinkel's Ethnomethodology (1967) research programme (EM) may be regarded as the second major influence on the development of what was to become CA. Taking as a departure point the Hobbesian Problem of how social order can be remotely possible if man seeks to attain egoistic goals, Garfinkel, like Goffman, focuses on everyday social action. Despite this, Garfinkel's focus was a different one. Strongly opposing the then pre-eminent functionalism and the Parsonian approach²⁰, Garfinkel in his programmatic Ethnomethodology argued that members of a society are to be regarded as knowledgeable beings that are aware of and understand their own actions. Social actions in EM are practical actions that should be studied as emerging, locally produced practical achievements, as all facets of shared understandings are dependent on a variety of tacit, socially shared methods of reasoning (Heritage, 1995). Consequently, in Garfinkel's view, the element to be researched is the situated, commonsense knowledge people have and exhibit, the methods relied on by interactants when achieving said action, or rather when producing and interpreting the "mundane world and commonsense understandings of it" (Schegloff, 1992:xxx).²¹ Additionally, Garfinkel emphasised the contextuality and "multiplex relevancies" (Heritage, 1995:392) of all signs, arguing that they are embedded in actions and will thus be interpreted as part of these action. Garfinkel consequently challenges the pre-Wittgensteinian notion of a direct link between sign and referent. Evidently, Garfinkel's Ethnomethodology was heavily influenced by Wittgenstein's language philosophy and phenomenology in the vein of Husserl and Schütz given that he takes a clearly hermeneutic stance and thus opposes the model of objectivity argued for in the positivist strand of sociology (Schegloff, 1992). Consequently, Ethnomethodology's central belief that "the structural uncertainty of meaning in everyday interactive

²⁰ Garfinkel (1967) goes so far as to (once again rather polemically) state that functionalism treated members of a society as "cultural dopes", as the Parsonian approach saw members of a society as acting upon internalized norms and rules and therefore considered social norms as the drivers of social action (Heritage, 1995). Garfinkel's insistence on perceiving members of a society not as passive subjects that act on internalized norms and social pressure within a 'container' context, but rather as active and reality-transforming agents went hand in hand with the spirit of the 1960s – a fact that Bergmann (2004b) suggests might have further added to the rapidly increasing popularity of EM.

²¹ Hence the name of Garfinkel's research programme, Ethnomethodology. It refers to the tacit knowledge, the taken-for-granted *methods* and practices which members (*ethnos*) of a given culture have at their disposal when interacting with each other, when rendering their surroundings meaningful (Flick, Kardorff, & Steinke, 2004; Bergmann, 2004c). In Garfinkel's (1967:11) own words, Ethnomethodology refers "to the investigation of the rational properties of indexical expressions and other practical actions as contingent ongoing accomplishments of organized artful practices of everyday life".

events is a constitutive condition for certainty of meaning” (Bergmann, 2004b:76) recalls Wittgenstein’s (1953:63) conclusion that “where there is meaning there must be complete order. [...] The most perfect order must also be hidden, therefore, in the vaguest sentence.” When considering phenomenology, on the other hand, Ethnomethodology and Conversation Analysis share an interest in intersubjectivity. Moreover, Schütz’s work is visible in Garfinkel’s writings when he employs Schütz’s (Schütz & Luckmann, 1979) ideas concerning the typified, dynamic and revisable character of commonsense knowledge and the categories in which it is organised (Heritage, 1995; Heritage & Goodwin, 1990).

Although many of Ethnomethodology’s central principles (such as the focus on social action, the investigation into commonsense knowledge and the belief in knowledgeable agents) made their way into Conversation Analysis²², the two research programmes differ. Sacks firmly states that he aims to arrive at an observational, natural science²³ (Sacks 1992; Schegloff, 1992) that could account for the generic ‘machinery’ (Sacks, 1992) of conversation, a machinery that is external to individual speakers, but can be utilised to accomplish actions.

Apart from these two key thinkers in sociology, technical advancement – namely the invention of the tape-recorder – also exerted considerable influence on the development of the radically new research programme that Conversation Analysis was to become. Researchers were now able to tape conversations. These recordings made it possible for the analysts to repeatedly listen to individual segments, helped them to avoid the methodological pitfalls posed by other methods, such as participant observation (Hutchby & Woffitt, 1998), and allowed them to focus on the local

²² Conversation Analysis is often referred to as Ethnomethodology’s “research programme” (Flick et al., 2004:10), hence the equally common abbreviation EMCA (Ethno-Methodological Conversation Analysis). It should, however, be noted that tension between EM and CA can be observed with regard to the formalisation and ‘linguisticisation’ in CA, which, followers of ‘pure’ EM argue, leads to a neglect of the ‘haecceitas’, of the Heideggerian ‘here-and-now’ nature, of the “singularity of the real” (Bergmann, 2004b:79).

²³ A concise summary of Sacks’ argument for sociology to become a natural observational discipline can be found in Schegloff (1992:xxxi):

“Contributions to science, including to sciences such as biology and neurophysiology, are composed of two essential parts. One is the account of the findings. The other is the account of the scientists’ actions by which the findings were obtained. What discriminates sciences from other epistemic undertakings is the claim that its findings are reproducible, and that reproducibility is itself grounded in the claim that the results were arrived at by courses of action reproducible by anyone in principle. Other investigators can, by engaging in the same actions, arrive at the same findings.”

Sacks thus argued in favour of a social science that employs a rigid, reproducible methodology.

practices and tacit conduct of naturally occurring, ‘unmanaged’ (Bergmann, 2004c) interaction. What is more, recordings made research findings transparent (and, ultimately, challengeable), and thus contributed to efforts made towards attaining an observational natural science (Heritage & Atkinson, 1984).

Further technological developments, such as the advent of video technology, did not have comparable consequences. Although video analysis of, for example, gaze and gesture, as represented by eminent scholars such as Charles (1986, 2000a, 2000b) and Marjorie H. Goodwin (Goodwin & Goodwin, 1992), allows for unmatched accuracy and is today commonly used to augment audio data in order to arrive at a more complete analysis of talk-in-interaction, the development of methods for video analysis seems to be lagging behind by a good two decades (Luckmann, 2006), and methodological discussions remain largely implicit. Only recently have scholars started to tackle the problematic aspects of video analysis; this includes the ‘reactivity’ problem (the interactants’ reactions to the camera), the “fusion, in shared time, of two structurally disparate modalities” (Luckmann, 2006:30), i.e. the vocal and visuospatial modalities and their differing methods of processing²⁴, the complexity and abundance of data provided by videos and the resulting questions concerning data management, and finally the potential manipulation of videos for recipient design purposes (Knoblauch, Schnettler & Raab, 2006). In comparison to audio data, then, the use of audiovisual data for multimodal analysis is still largely underdeveloped.

With Goffman and Garfinkel providing the theoretical framework and in light of the technical advancements ushered in during the 1960s, CA emerged as a research programme that allowed for the structures of interaction to be studied. The key assumptions and concepts put forward in the research enterprise reflect this background, and will be treated in further detail below.

²⁴ Whereas the vocal layer is subject to a temporal structure, visual perception does not follow similarly clear linear structure. For standard video analysis, eye movement and visual processing is near impossible to trace, and the distinction between relevant and irrelevant details of an interaction poses problems (Luckmann, 2006). Experimental studies working with head-mounted eye trackers hope to provide more insights. For EMCA-driven research, the central tenet that only those features which are demonstrably relevant to participants remains true.

3.2.2. CA: Key assumptions and concepts

Conversation Analysis subscribes to the principle that in all forms of everyday interaction and communication, be it linguistically or non-linguistically-oriented, agents are engaged in coordinating their actions with others, in interpreting and making said actions intelligible. When talking, we are *doing* things, inviting, negotiating, arguing – and all linguistic levels can be associated with actions or activities (Drew & Heritage, 2006). In accordance with Ethnomethodology, Conversation Analysis consequently considers joint and coordinated action to lie at the core of human social behaviour. Utterances in this view are actions that are embedded, situated, and methodically produced in specific interactional and sequential contexts.²⁵ CA aims to discover the tacit mechanisms through which agents collaboratively produce meaningful actions and the ordering of events, or rather naturally occurring talk-in-interaction. These mechanisms and methods are recognised as being orderly and available to agents due to their membership in a given community (Bergmann, 2004a; Drew & Heritage, 2006). Here, CA assumes the following basic propositions, as Psathas (1995:2-3; my comments added, VS) concisely summarises:

- “Order is produced orderliness”: Order is integral and endogenous to the local setting in which the interaction occurs.
- “Order is produced by parties *in situ*”: Order is “situated” and “occasioned” rather than imposed on the interactants from the outside. Order is similarly not seen as the strict adherence to a set of rules. In line with the assumptions

²⁵ The claim that utterances can be analysed as social actions is, of course, the central component of Speech Act theory. Although Searle’s work parallels that of Sacks (Searle’s pioneering paper, “What is a Speech Act”, was published at around the same time as Sacks’ first lectures in 1965), the two research traditions follow separate tracks and exhibit marked differences regarding methodology and focus. Austin (1962) and later Searle work on classes of utterance (such as the speech act ‘promising’), and study the conditions necessary and sufficient to felicitously accomplish a given speech act. Speech Act Theory thus does not focus on contextualized, individual, and naturally occurring utterances as Conversation Analysis does, but rather on types of utterances and the actions they accomplish. Furthermore, rather than addressing “necessary conditions” (Austin, 1962:14) for the “felicitous” (ibid) performance of a legally or ritually binding action as Speech Act theory does, CA aspires to uncover the “practices” and “partial methods” used by interactants to methodically produce that action in a given interactional context. Sacks, too, was critical of Speech Act theory (Heritage & Atkinson, 1984; Schegloff, 1992), as it failed to grasp the importance of interactional placement for the accomplishment of an action.

proposed by Ethnomethodology, agents produce order themselves. Meaning is “locally accomplished, situated, and conventional” (Psathas, 1995:52).

- “The parties orient to that order themselves”: The order observed is not one that is perceived solely by the analyst or has been superimposed by a preformulated theory regarding that which is occurring; rather, the participants themselves are observed to orient to order.
- “Order is repeatable and recurrent”
- “The discovery, description, and analysis of that produced orderliness is the task of the analyst”
- The frequency of a given phenomenon cannot be considered to form the basis for research in CA, as the focus lies on “discovering, describing, and analysing the structures, the machinery, the organised practices, the formal procedures, the ways in which order is produced.”
- “Structures of social action [...] can be described and analysed in formal [...] structural, organisational, logical, atopically contentless, consistent, and abstract, terms”.

Research in Conversation Analysis therefore broadly follows two aims: one the one hand, it investigates the sequential order of talk-in-interaction while on the other it explores the normative and inferential characteristics of talk-in-interaction (Hutchby & Woffitt, 1998). To illustrate this approach, I would like to briefly introduce a key topic in CA, turn-taking. Up to this day, accounting for turn-taking systems forms the very backbone of the research enterprise. Sacks, Schegloff & Jefferson (1974) provided the first “systematics” (ibid) of turn-taking. In their seminal paper, they first of all understand turn-taking to be “locally managed, party-administered, interactionally controlled, and sensitive to recipient design” (ibid:696). There is, in other words, no rigid system imposed on turn-taking; it is negotiated *in situ*. Nonetheless, a systematic for turn-taking can be identified. In this respect, the authors differentiate between two central components of turn-taking: the Turn Constructional Component and the Turn Allocation Component.

Regarding the first, turns comprise Turn Constructional Units (TCU), “the smallest interactionally relevant complete linguistic unit” (Selting, 2000:477). Such TCUs vary in length (i.e. a TCU might be a single lexical item or an entire sentence) and are followed by a transition-relevance place (TRP), the point at which speaker

change is a legitimate next action. Participants project a TCU to be possibly finished on the basis of certain context-sensitive parameters, the complex and context-dependent interplay of pragmatic (is the proposed action possibly complete?), prosodic (turn-final intonation pattern), bodily (e.g. gaze direction and postural orientation) and grammatical features (syntactical completion).

The Turn Allocation Component of the model refers to the ways in which turns-at-talk may be assigned. At a TRP, the current speaker may either select a new speaker (through gaze, questions, addressing and the like), or the next speaker may self-select in environments where prior talk allows for speaker change. In short, then, turn-taking in this model is interactionally negotiated and enacted *in situ*, and centres around TRPs at which speaker change may legitimately occur.

When studying interactional phenomena such as turn-taking, CA endorses basic research tenets, the use of naturally occurring data, and the avoidance of a priori conceptualisations. The first tenet holds that only those interactional phenomena that would have occurred in any event, irrespective of the research project and the analyst, are chosen for analysis. Natural interaction is thus non-experimental action, or rather interaction that is situated in the unfolding of the interactants' lives. At the heart of this tenet lies the central assumption that 'ordinary conversation' is the "primordial scene of social life" (Schegloff, 1996a:4) provides a much richer domain for analysis than experimental settings or invented examples (Heritage, 1995). As CA focuses on social action in face-to-face interaction, the phenomena analysed include the interplay of all those visual, auditory, tactual and/or kinaesthetic features within a spatio-temporal framework (Psathas, 1995) that participants demonstrably rely on in order to jointly coordinate social action. Accordingly, the interactional phenomena analysed include vocal and visuospatial modalities (Stivers & Sidnell, 2005); this thereby incorporates lexico-syntactic and prosodic features as well as embodied actions, gestures, facial expressions, gaze, body posture and alignment. A major requirement is that the matters selected for study are those that participants in the setting are themselves demonstrably aware of and/or oriented to in the course of their actions. This departure from analysing purely 'vocal' aspects of interaction (which first generation CA research had still largely concentrated on) was made possible due to technical developments in video recording, and enjoyed widespread successful

promotion by Charles and Marjorie H. Goodwin.²⁶ The recognition first of all that crucial facets of action relate in some way to the physical environment in which the action is embedded, the use of artefacts and the interactants' bodies, and secondly that the sequential organisation of action is dependent on these factors (Ten Have, 1999) led to further theorising in EM, CA and Cognitive Linguistics, such as the Fused Bodies approach proposed by Hougaard & Hougaard (in press). This approach emphasises the important role of "whole bodies in interaction", and aims to describe "sense achieved interactionally in and through the use of all embodied resources (including language) in face-to-face-interaction" (ibid).

Furthermore, multimodal analyses have of late been used in order to shed light on cognitive processes. To this end, gesture studies are commonly applied, as gestures are seen as providing a window into the mind, or, as the title of one paper suggests, "Postcards from the mind" (Peter de Ruiter, 2007). Within MSCI research, multimodal analyses focusing on gestures have also become increasingly common. Sweetser (2007:210), for example, shows that "divisions of space in gesture similarly metaphorically represent *Mental Spaces* or areas of content" (original emphasis). Gesture analysis, she claims, can thus be used to determine the organisation of Mental Spaces. Despite these recent trends towards further inclusion of the body, however, CA remains predominantly interested in the interweaving of vocal and visuospatial aspects in sense-making processes (Stivers & Sidnell, 2005) with the vocal layer remaining prominent.

Secondly, CA calls for the avoidance of a priori conceptualisations, and instead champions what is referred to as "unmotivated looking" (Psathas, 1995:45). By this, CA hopes to describe and analyse phenomena without consulting the existing repertoire of theories regarding human conduct, and to be open to new discoveries 'overlooked' by previous theories. CA furthermore seems deeply critical of existing theories, often arguing that these theories would reduce the complexity of human interaction to simplifications and inflict "more scholarly and theoretical concerns [...] upon members of society" (Wetherell, 1998:391). CA alternatively endeavours to chart the endogenous understandings being displayed by the interactants. Despite this critical take on existing ideas on social order, CA cannot be accused of being 'a-

²⁶ Along with scholars such as Christian Heath (e.g. Heath & Hindmarsh, 2002; Heath & Luff, 2006) and Adam Kendon (e.g., 1990b, 1997, 2004, 2008), who mainly works in interactional kinesics and had already started research on gaze and gesture in the 1960s.

theoretical'. As outlined above, CA strives to highlight the "inherent theories-in-use" (Ten Have, 1999:32) that participants employ rather than imposing interaction-external models as a means of accounting for social conduct. Due to the rigorously empirical nature of such concepts, however, they cannot be regarded as imposing interaction-external theories on the interaction; instead, they arise from the systematic and accumulated study of interaction. Be that as it may, it should not be neglected that simply by selecting a particular instance of interaction for analysis, the analyst attributes this excerpt a particular relevance and neglects its diachronic dimension to a certain degree (cf. Wetherell, 1998). Despite coming from a different angle, Fauconnier (1997) also en passant points out that theory-independent observations and interpretations of 'facts' are impossible. Explication-only endogenous understandings can therefore never fully be achieved.

CA's focus on explicating the tacit structures of interaction without taking into account 'larger', 'interaction-external' models and concepts further led to fears that CA was a "bloodless" (Moerman, 1988:x) research enterprise that neglected aspects of 'context' and 'culture'.²⁷ This 'cultureless smack' becomes especially prominent in early CA research, where tape recording of conversations are analysed without much reference to 'context' (Bergmann, 2004c). Following the understanding of 'agency' in Ethnomethodology, 'context' in the broadest sense and sociological variables such as age, class and gender are taken into account only when it is evident that the participants in the interaction themselves orient to these variables and that this in turn becomes visible in the recipient design of their contributions. Arguing that the actual context of an interaction cannot fully be defined and that that which is relevant cannot be fully determined by the observer (de Kok, 2008), practitioners of CA therefore usually refer to the most proximate context, the previous turn (Drew & Heritage, 2006), when speaking of 'context' – a conceptualisation that becomes visible in, for example, the study of *adjacency pairs*. Conversation Analysis has found that interaction works in pairs. Such adjacency pairs consist of a First Pair Part (FPP), which projects an action, and a Second Pair Part (SPP), which completes it. A FPP might be a greeting, to which the relevant SPP would be an acknowledgment of the greeting (through a nod or another culturally accepted form of greeting). In adjacency

²⁷ This view is related to the notion of 'culture' found in CA work: for Sacks (1992), 'culture' is the gadgetry that enables participants to produce recognisable actions. Culture is thus commonsense knowledge, the competences that enable members to produce and infer meaningful actions (Ten Have, 1999).

pairs, each action is both context-defined (resulting from a previous action) and context-renewing (defining the action that follows). Context in CA is thus not static, but dynamic (Heritage & Goodwin, 1990). As a result, CA opposes the standard sociolinguistic tenet of an “intrinsic and causal relationship between language and the social context in which it is produced” (Hutchby & Woffitt, 1998:5), namely the tenet that all interaction is influenced by interaction-external ‘context’. Accordingly, mainstream CA – and especially the early work conducted in this research enterprise – is not interested in the “ethnographic particulars” (Psathas, 1995:45) in which the interaction under scrutiny took place. Rather, it is the structure of the interaction that CA hopes to explicate, as orderliness in CA is not considered to be dependent on said particulars. On the contrary, sociological and psychological characteristics are understood to be manifested in the medium of conversational practices. The focus now should therefore turn to structural features. Only after these features have been determined is it possible (if at all) to investigate the way in which sociological factors may be displayed during talk-in-interaction (Heritage, 1995). This ‘pure’ strand of CA has, however, been criticised over the past two decades. The criticism originated from two camps: those CA researchers working in ‘Applied Conversation Analysis’ (also referred to as the ‘Studies of Work Programme’) and the ‘culturalist’ and ‘humanist’ branch of CA.

Applied Conversation Analysis studies institutional talk to understand how a given interactional phenomenon might be tied to the larger context of the organisation in which an interaction is embedded (Psathas, 1995). It therefore aims to discern how social action is linked to social structures, and how epistemological and interactional asymmetries are produced in interaction. Drew & Heritage (1992) succinctly summarise three main aspects exhibited by institutional talk:

- Interactants normally aim to achieve specific goals that are linked to their “institution relevant identities” (Ten Have, 1999:168)
- Special constraints can be found that determine ‘allowable’ contributions; these are contributions that are relevant to the business or institutional goal at hand
- Specific ‘inferential frameworks’ become salient and certain procedures are found in particular institutional contexts only.

Applied CA therefore takes a slightly different angle to ‘pure’ CA by investigating the local competencies and special knowledge exhibited by members in certain institutionally-contexted interactions, such as hospitals (e.g. Heritage, 2010; Heritage & Maynard, 2006; Paterniti et al., 2010; Drew, 2006), courtrooms (e.g. Drew, 1992; Ehrlich, 2002, 2006) or auctions (e.g. Heath & Luff, 2007a, 2007b, 2010) amongst others. Consequently, the criticism levelled against more ‘pure’ CA involves the latter’s propensity to assume that any interactional sequence contains allegedly context-independent devices. Applied CA has, however, shown that interactants in fact exhibit an orientation to specific institutional contexts that manifests itself in, for example, the turn-taking system, where in certain contexts, such as doctor-patient interaction or courtrooms, pre-prescribed rights at turn-taking might prevail. Nevertheless, applied CA does not opt for a ‘container view’ of context, whereby interaction-external constraints are imposed on interaction from the outside; rather, it argues that institutional talk contributes to the discursive construction of the institutions’ actual ‘institutional’ character. Furthermore, it maintains that context-free resources are applied in context-sensitive interactions by means of knowledgeable agents that display to others an orientation to, and awareness of, a given context (Drew & Heritage, 2006; Heritage, 2004).

The ‘culturalist’ strand of CA voiced its criticism in Michael Moerman’s *Talking Culture* (1988), in which the author argues in favour of combining CA’s “high-powered lens”, which, he feels, lacks “range, breath, and *mise-en-scène*” (ibid:x), with ethnography and its “concern for context, meaning, history, and intention” (ibid:xi). As Moerman goes on to explain, this is necessary due to the fact that “we never merely exchange turns of talk. In all conversation, people are living their lives, performing their roles, enacting their culture” (ibid:22). What Moerman is thus proposing is what he refers to as ‘culturally-contexted’ Conversation Analysis. Here, ethnographic descriptions “provide meanings and material conditions of the scenes in which the actions occur” (Ten Have, 1999:56). Culturally-contexted CA therefore allows for a description that – albeit never complete – can illustrate the interconnection between cultural directives and individual action design; this is especially useful when working with communities which are unfamiliar to the researcher, as analysis of recordings of talk-in-interaction requires membership knowledge, foresight concerning a culture’s shared, tacit and taken-for-granted knowledge, or rather a community’s doxa (cf. Bourdieu, (1972)1977), cultural models

and practices (D'Andrade, 1984), if the occurrences in a given interaction are to be understood or differentiated between. As Geertz (1973) succinctly states, this essentially involves being able to tell twitches from. This cultural knowledge is especially important in humorous interaction (Kotthoff, 2006), as humour employs, alludes to, and comments on culturally salient 'humour scripts' (Raskin, 1985) as well as culturally salient phenomena, stereotypes and prototypes (Niemelä, 2010). Impersonations in particular often echo culturally shared voices that lie beyond the adjacent interactional context (Koven, 2002). Without awareness of the community's taken-for-granted cultural knowledge and models, the social actions which humour may encompass cannot be fully grasped by the researcher. This point is also of key importance to cognitive semantic research, which contends that linguistic structure activates frame structure, such as cultural scripts and other forms of shared encyclopaedic knowledge; these forms of structure impede on, and are at the basis for, operations such as conceptual integration.

In order to account for the roles played by shared knowledge and practices in the intersubjective coordination of meaning, I align with culturally-contexted CA, and utilise ethnographic information to describe and analyse the (context-free) operations with which (context-sensitive) meaning construction relying on conceptual blending is achieved. However, it is not just the treatment of 'context' within CA that needs to be reconsidered when applying it to a cognitive-semantic theory such as MSCI. Preliminary conceptual reconfigurations, too, are required for the interaction of the two approaches and the subsequent generation of useful results. These reconfigurations will be incorporated into the analysis to follow in later chapters, and will be reassessed on the basis of the results gathered in the process. A refined model of MSCI in interaction is offered in chapter six.

3.3. A social-interactional approach to MSCI: Preliminary hypothesis on conceptual reconfigurations

Employing a predominantly ethnomethodological approach to investigate a cognitive semantic theory that hopes to offer new insight into "backstage cognition"

(Fauconnier & Turner, 2002) initially seems to result in clashes with regard to conceptions of cognition and meaning. Ethnomethodological Conversation Analysis perceives interactants as socially embedded, as making sense of a social world through ordered practices. In this way, CA seems almost “agnostic” (Hopper, 2005) towards cognitive processes, or, as Gitte R. Hougaard (2008) understands it, as residing on the ‘communal end’ of a meaning-making continuum that stretches from individuality to communality. This focus on the social and the observable in contrast to the individual mind is reminiscent of the mentalism vs. behaviourism debate that dominated academic discourse for a good deal of the twentieth century, and resulted in the time-honoured division between ‘social’ and ‘psychological’ dimensions of human behaviour, respectively (cf. Croft, 2009). Yet while CA can in many ways be considered a behaviourist enterprise in the vein of George Herbert Mead (1934) and his later supporters, the mind is not negated through this approach.²⁸ Rather, the mind is conceptualised as a “shared activity” (Hougaard, 2008:198), with meaning emerging between interactants rather than in the individual mind alone. With its focus on the joint and shared, CA does not deny individual experience either, as Hougaard & Oakley (2008) point out. Instead, it puts its focus on the social and the enacted without denying psychological dimensions. In Schegloff’s (2003:39) words, cognition should consequently “be understood not necessarily by reference to the individual cut off from a world around, but by reference to an individual engendered and constituted by the world around in the first instance.”

This view is not altogether different from the one held in Cognitive Semantics, which subscribes to experiential realism (Lakoff, 1987). As outlined in chapter 2, ‘meanings’ in Cognitive Semantics are seen as being dynamically construed by an embodied mind that interacts with the world around it and with other cognising minds (Langacker, 1999, 2001; Williams, 2004). However, this latter dimension, namely interaction with other cognising bodies, has largely been neglected in actual research, or at least prior to the ‘social turn’ (Croft, 2009; Harder, 2010). As a number of recent projects (see above) have shown, though, EMCA can be successfully aligned with Cognitive Linguistics approaches such as MSCi due to its phenomenological

²⁸ It should also be noted that Conversation Analysis cannot be regarded as purely behaviourist. Interactants are understood to be sense-making individuals rather than being fully driven by stimuli. Still, interactants obviously cannot determine their conversational partners’ experiences by looking into their heads. Instead, Hougaard (2004) argues, meaning experiences have to be signalled and shared, leading to the emergence of interactional meaning.

roots²⁹ and resulting interest in shared knowledge. Marrying EMCA to MSCI helps to address key problems within the latter, but requires the rethinking of a few key issues.

This importance of preliminary hypotheses for conceptual reconfigurations becomes further evident when once again considering the *Achilles is a lion* example discussed in chapter 2.3., where one interactant draws an analogy between the shared friend Achilles' beard and that of a male lion. In this interaction, the interlocutors will do their best to align conceptualisation processes in order to attain intersubjective meaning construction. To do so, they will first rely on their shared knowledge of (a clean-shaven) Achilles, and the salience of his beard in their shared perceptual field. Secondly, the interactant will try to guide his/her partners' mapping processes by emphasising central words of the utterance, gesturing to his/her face and waiting for signals of recognition to be emitted by the communicative partner. Should he/her not instantly comprehend the analogy, further steps will be initiated to ensure that all interactants arrive at approximately the same conclusions. Meaning is, in short, an interactive achievement, a "social doing" (Sinha, 1999:232). Nevertheless, the question remains as to how the central importance of the intersubjective dimension of meaning-making can be accounted for in a cognitive semantic analysis.

3.3.1. Cognition as 'centre stage' rather than 'backstage'

'Mainstream' Cognitive Semantics attempts to discern the 'unconscious' processes of 'backstage cognition'. This approach has led to the development of what Anders Hougaard (2008:181) calls "hypothetical modelling" of seemingly inaccessible mental operations. An approach inspired by CA, on the other hand, conceptualises cognitive processes as happening 'onstage' rather than 'backstage', to continue with Hougaard's (2004) pithy exploitation of Fauconnier and Turner's original stage metaphor. As the interactional data accumulated within CA shows, cognitive meaning-making processes are shared, coordinated, "interactionally enacted" and "part and parcel of the interactional action" (A. Hougaard, 2008:22). This would be the case in the Achilles example described above, where the speaker makes certain to explicitly guide mapping processes by means of gestures and prosodic emphasis.

²⁹ For a further discussion of how Cognitive Linguistics (mainly in the vein of Lakoff, Johnson, Talmy and Langacker) relates to Phenomenology, see Zlatev (2010).

Cognition consequently takes centre stage due to the fact that it is publicly displayed. Conceptualising cognition as ‘onstage’ allows the joint *processes* of meaning-making to be investigated rather than having to start at the product and hope to trace the steps that were taken in order to arrive there. Investigating the step-by-step negotiation of meaning-making as it temporally unfolds and is publicly revealed will consequently address one of the major points of criticism (cf. Gibbs, 2000) voiced against MSCI.

3.3.2. Common Ground and Mental Spaces

Mental Spaces in such a social-interactional approach are intersubjectively coordinated and are not isolated packages in an isolated individual. Returning to the Achilles example, the speaker will need to be sure that his/her communicative partner is aware of the fact that male lions have manes and that he/she has noted Achilles’ new look. Mental Spaces constructed by an individual cogniser in on-line cognising are accordingly based on assumptions regarding shared ‘Common Ground’³⁰ as proposed by Herbert H. Clark (1996, 1982). In his definition, Common Ground refers to the “the sum of their [the interactants’] mutual, common, or joint knowledge, beliefs, and suppositions” (Clark, 1996:93). Without assuming a certain degree of Common Ground, of shared knowledge and awareness, this means that joint actions and, ultimately, successful communication are not possible. After all, Tomasello (2008:57), in building on Wittgenstein (1953), shows that “linguistic communication relies to a much greater degree than is readily apparent on uncoded communication

³⁰ Common Ground is sometimes also referred to as *common knowledge* (Lewis, 1969), *mutual knowledge* or *belief* (Schiffer, 1972) and *joint knowledge* (McCarthy, 1990; All in Clark, 1996). This notion of an assumed shared world is also prominent in Ethnomethodology, as in, for example, Garfinkel’s (1963:220, quoted in Hougaard (2004:cclxxviii)) “congruency of relevances”, in which

“The person [who] expects, expects that the other person does the same, and expects that as he expects it of the other the other expects the like of him that the differences in their perspectives that originate in their particular individual biographies are irrelevant for the purposes at hand of each and that both have selected and interpreted the actually and potentially common objects in an ‘empirically identical’ manner that is sufficient for the purposes at hand.”

Hougaard & Oakley (2008) go as far as to argue that Mental Spaces resembled Common Ground. For me, this seems somewhat ostentatious, as Common Ground is first of all understood to subsume all knowledge existing between participants, whereas Mental Spaces are built up for local meaning-making; for this reason, Mental Spaces only subsume those features of shared knowledge that are relevant to current meaning construction. Secondly, it is still the individual that experiences events as meaningful. Mental Spaces thus have to be coordinated between interactants in order to achieve approximation in their construal. While they are signaled in an attempt to achieve shared construal, they nonetheless remain the construct of the individual.

and other forms of mental attunement.” Consequently, interactants have to make a cooperative effort to take into consideration their conversational partners’ potential actions, background knowledge, beliefs and intentions (Tomasello, 2008)³¹; otherwise, communication would simply break down. This is of particular importance in figurative and creative language use. Research in the development of metaphorical competence (e.g., Glaznieks, 2011) has shown that people lacking detailed knowledge of a certain domain find it impossible to understand novel metaphors recruiting that domain. Humour that relies on the recruiting of knowledge schemas not available to the audience will be entirely unsuccessful. As such, the consideration of the communicative partners’ potential knowledge, actions and intentions is of key importance.

These assumptions with regard to Common Ground are reflexive and recursive, vary in quality and are therefore subject to judgment and evaluation, as evidence needs to be gathered to determine the validity of a given supposition. This evidence is established on the basis of such interactional resources as joint gaze or gestures. When wrong judgments are made as to the Common Ground, repair mechanisms are available to be utilised.

Mental Spaces in an interactionally-oriented Cognitive Semantics rely on Common Ground, a view that is also implicit in Turner (2000:46), who writes that

the effectiveness of the manipulations [i.e. rhetorical effects, VS] depends on the *shared* nature of these cognitive operations and conceptual structures – they are part of the backstage cognition of the members of the audience. [...] What can be recruited to mental work depends on social and cultural location. Parts of the repertoire are common and can be assumed for any audience while other parts are special to special communities or special situations (my emphasis, VS).

The effectiveness of cognitive operations, such as conceptual blending, therefore depends on assumptions made about shared knowledge and frames. As such, Mental Spaces are locally negotiated and interactionally displayed by participants in

³¹ A related notion can be found in ‘mainstream’ Conversation Analysis, where the term *recipient design* denotes the ways in which talk in conversation is constructed or designed to display an orientation and sensitivity to the particular other(s) constituting are the co-participants. Recipient design affects, for example, word selection, topic selection, admissibility, the ordering of sequences and the conventions for opening and closing conversations (Sacks, Schegloff & Jefferson, 1974). In contrast to Common Ground, however, it does not explicitly refer to epistemic, perceptual or cognitive dimensions.

interaction for joint sense-making purposes, and are based on the reflexive assumptions made concerning shared Common Ground.

3.3.3. Coordination and Mental Spaces

As Fauconnier (1994:2) points out, for Mental Spaces to be shared interactants need to “build up similar space configurations from the same linguistic and pragmatic data” in order to successfully communicate, because “to communicate is to trigger dynamic creative processes in other minds and in our own” (Fauconnier, 1997:182). What Turner (2000) and Fauconnier (1994, 1997) here seem to neglect, however, is that this building of “similar spaces” is a highly complex endeavour that is subject to careful coordination. Rather than trying to reconstruct the spaces which each individual builds, an interactional approach will necessarily focus on the *shared* Common Ground that interactants orient to in their conversations and that serve as a basis for joint meaning construction.³² For this reason, Mental Spaces are regarded as shared between interactants, whereby *shared* refers to the notion that a given piece of information or beliefs are *assumed* by the interactants to be available to all interactants. In this manner, Mental Spaces are also *reflexive* in that interactants are conscious that they and their partners are aware of a given proposition. This meta-cognition ties in with findings obtained in Theory of Mind research, which show that during childhood, humans develop an understanding of the mental states of others. Humans thus seem to know that others, too, have mental representations of the world and that these representations inform their respective actions. Humans consequently assume and orient to others’ motivational and mental states (Astington, 2001; Givón, 2005). People on the autistic spectrum and children under the age of four lack Theory of Mind, and therefore find it more difficult to understand figurative expressions, such as novel metaphors, irony or hyperbole, as their ability to ‘read’ the intentions, the mind of other, is not fully developed (Glaznieks, 2011). Moreover, it has been suggested in comparative psychology (e.g., Tomasello, 2008; Herrmann et al., 2007; Liebal et al., 2009) that humans possess a special set of cognitive skills, known as social cognition, to

³² In their discussion of conceptual blending and evaluation, Kok & Bublitz (2011), too, emphasise the importance of aligning our own mental worlds with that of our communicative partners, but do not go as far as to define Mental Spaces as relying on Common Ground; indeed, they do not even incorporate Common Ground into their blending model.

an extent that other primates simply do not. Social cognition is what allows humans to read intentions and cooperate in joint activities.

This awareness of other minds can, as Clark (1996) finds, be built on either communal and cultural or personal knowledge. While communal knowledge (such as that of a lion's appearance) is shared or assumed by way of belonging to a specific group or subculture, personally grounded knowledge refers to those assumptions that are based on previous, shared experiences accumulated by interactants in their personal interactions (such as Achilles' erstwhile appearance), as well as on a perceptual basis (such as Achilles' current look). That which is situated in the interactants' shared perceptual field and subject to joint attention can hence be considered to belong to the personal Common Ground.

Other aspects of the personal Common Ground relate to what Croft (2009) refers to as the "discourse basis", which is the previous conversations interactants have experienced together. Communal Common Ground encompasses the shared stories, concepts, things, genres or symbols that circulate in a given community and that can be assumed as being shared by virtue of joint community membership.

In this social-interactional approach, then, Mental Spaces are conceptualised as conceptual packets based on assumptions regarding Common Ground; they are publicly displayed and signalled, or rather interactionally enacted, and guided joint local processes of meaning-making and action. As reading minds is of course not possible, interactants can never be fully certain as to the quality and validity of a given feature (although the better that interactants know each other and the more personal Common Ground they share, the more the quality tends to enhance), and therefore constantly need to monitor, modify and coordinate their Mental Spaces.

3.3.4. Mental Spaces as interactionally relevant and phenomenologically evident

Conceptualising Mental Spaces as interactionally shared and substantiated by Common Ground serves to address another major criticism voiced towards MSCI, namely the question as to what Mental Spaces really are. After all, the founding fathers provide very little information as to a proper definition, and, depending on the analyst, many different elements are believed to reside in Mental Spaces: anything from perceptual and conceptual structure and structured and unstructured scenarios to emotions. Rather than perceptual structure, in this interactional approach Mental

Spaces now only comprise views, superstitions and knowledge, essentially *framed semantic ‘meanings’* that are interactionally relevant and shared (cf. Hougaard, 2004).

This also applies to the ‘contents’ of the blended space. Investigating blending in interaction means to explore blending on a *phenomenological* level and investigating *consciously* and *mindfully* set up blends, whose *dialectic* nature is explicitly oriented to in interaction. As a result, blending operations on the grammatical level (constructions blending with semantic content) or perceptual level (colour integration) are not considered, as the alleged input spaces are not explicitly *acted out* and *oriented to* in interaction.

3.3.5. Current Discourse Space as a basis for communication

It goes without saying that not all facets of the Common Ground understood to exist between interlocutors are treated as relevant at every moment in a given interaction. For this reason, I will draw on Langacker’s (2001) notion of Current Discourse Space (CDS). This discourse base space is used to refer to those aspects of the Common Ground that are “construed as being shared by the speaker and hearer as a basis for communication at a given moment in the flow of discourse” (Langacker, 2001:144).

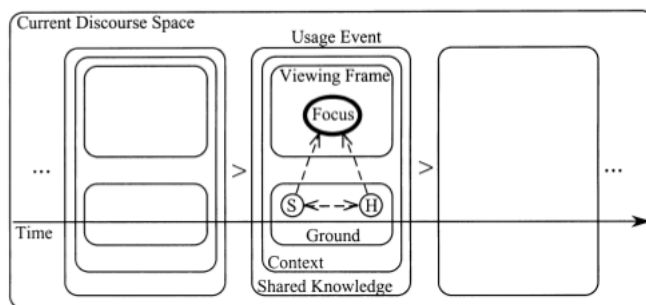


Figure 10: Langacker's (2001:145) Current Discourse Space

Langacker here proposes that interactants in usage events draw on shared knowledge, but also take into account what he refers to as ‘context’, i.e. the “physical, mental, social and cultural” (Langacker, 2001:145) aspects that are treated as shared and relevant in the usage event, as well as discourse history – the Common Ground discussed in the previous chapters. A usage event in his model refers to the “action carried out by the speaker and hearer” (ibid:144).

It is not only this focus on social action that makes Langacker's model compatible with discursive approaches to MSCI; he also accounts for the joint negotiation of meaning construction. His viewing frame in this regard describes the joint managing of the current attention range, i.e. the coordination of beliefs about currently relevant conceptual entities and the focus of attention (visually represented via --->). The viewing frame comprises what Langacker (2001) refers to as conceptualisation and vocalisation channels. These channels consist of factors such as speech management (turn-taking behaviour), as well as intonation and gesture, which are employed in the management of attention frames. The Ground in Langacker's (2001) model refers to the process of meaning negotiation and coordination that the speaker (S) and his/her addressee, the hearer (H), are engaged in. Multiparty interaction is, however, more complex and the Ground of these interactions involves more than a speaker and a hearer. It has to be able to account for the "participants" and "non-participants" (Clark, 1996:14) witnessing the interaction, as awareness thereof has a significant impact on contributions. By way of example, in addition to coordinating their contributions with their addressees and fellow panellists, speakers in televised interaction also shape their contribution whilst keeping in mind the non-participating audiences. These audiences can be divided into two types: the audience present in the studio, which can signal understanding through laughing or clapping, and the non-present audience watching via their TV screens. The coordination of blending operations in a way that allows the latter group to follow the exchanges and 'participate' in the joint project, so to speak, constitutes a major challenge for the speakers. In order to account for the multiplex structure of multiparty, televised interaction, potential participants (i.e. all panellists), non-visual bystanders (i.e. the studio audience) and all other listeners (i.e. the TV audience) are included in my reconfiguration of Langacker's (2001) Ground. Figure 11 provides a schematic visualisation of the participant structure in panel shows.

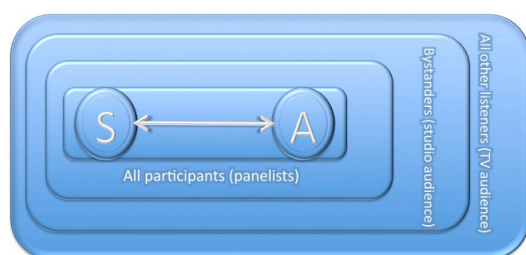


Figure 11: Participants in the Ground of televised multiparty interaction (cf. Clark, 1996:14)

For the analyses implemented in chapters 4 and 5, Langacker's (2001) model is chosen over Fauconnier's conception of a base space and Brandt & Brandt's (2005) reworking thereof for the following reasons. First of all, the CDS allows for the description of a steadily evolving mental space that serves as the basis for the interaction at hand. It subsumes those features of the vast Common Ground that are deemed to exist between interactants treated as shared and relevant at the moment of interaction, such as previous interaction, cultural paragon, perceptually salient items or culturally shared values, and constitutes the background for the various construal operations – these are features which are represented in the base space in the manner conceived in canonical MSCl.

The CDS not only licenses and makes relevant blending scenarios by subsuming such features as discourse history and roles, but can, as a *discourse base space*, also serve as an input space that is blended with alternative scenarios, as is the case in the phenomenon discussed in the analysis of impersonation humour.

Second of all, other than the model offered by the Aarhus research group in cognitive semiotics (cf. Brandt & Brandt, 2005), which differentiates in great detail between expressive acts, circumstances and pheno-world spheres in a semiotic base space but refrains from explicitly considering shared knowledge, Langacker's (2001) Current Discourse Space allows the shared, yet unstable and fragmented nature of Common Ground to be explicitly acknowledged whilst simultaneously being able to depict its historically grounded and evolving nature with an focus on joint action, conversational relevance and cognitive construal.

3.3.6. Blending analyses in a social-interactional Cognitive Semantics

Regarding meaning construction in conceptual blending as jointly coordinated and recipient-designed further leads to the question concerning how blending analyses are conducted in this social-interactional approach. More specifically, the interactional, phenomenologically grounded approach taken here raises the question as to how to deal with the contested “generic space”. The generic space as conceptualised in canonical blending theory is, after all, not interactionally signalled and the ‘content’ of this type of *tertium comparationis* cannot be inferred from the interactants’ actions. Secondly, generic spaces are ‘post hoc’ constructed by analysts, and consequently do not provide any information as to the motivation behind and the different readings

arising from a blend in context. While representing a generic space in the analysis of, say, metaphors might be interesting and relevant to cognitive and philosophical studies, I here follow the standards adhered to in the most recent interactional blending research (see for example Brandt, 2008; Brandt & Brandt, 2005; Hougaard, 2004; Oakley & Coulson, 2008; Oakley & Kaufer, 2008; Williams 2004, 2008). In an interactional and contextualised approach to Cognitive Semantics, a ‘post hoc’ constructed structure which lists the similarities between the inputs cannot faithfully be represented unless this generic structure is explicitly referred to by the interactants. Instead, the analyses focus on jointly coordinated and contextualised processes of meaning-making.

The social-interactional approach to Cognitive Semantics championed in this project does not study meaning as processed in the individual mind, but as socially and locally constructed and interactionally relevant. In this approach, Mental Spaces are therefore understood to be shared and interactionally enacted for purposes of local understanding. They are thus based on Common Ground (Clark, 1996) in that they are rooted in communally or personally shared beliefs, suppositions or knowledge. They are publicly displayed and held for purposes of local meaning-making, and are embedded in the *haecceitas*, the here and now of the interactants.

This approach consequently addresses the criticism brought against MSCI that targets the theory’s apparent focus on the logic of hidden meaning construction (cf. Brandt, 2005). The meanings emerging in blending operations are conceptualised as processes of creative meaning construction which are shared and designed by the recipient. Accordingly, the MSCI analyses only ever model projections which follow from the interaction under scrutiny, and avoid the ‘hypothetical modelling’ which was a common feature of early MSCI research. This is not to deny that the analyses and the graphic representations that subsequently appear in conjunction with them are completely void of the analyst’s membership knowledge or introspection. An interactional approach does, however, require empirically solid interactional evidence for all assumptions made as to the meaning construed by the participants. This focus on interactional evidence naturally leads to the analyses being less fine-grained and detailed than those offered in more traditional MSCI research, as Anders Hougaard (2004) has already pointed out when outlining his interactional approach to Cognitive Semantics. Yet, analyses conducted within a situationally and interactionally

grounded framework address those contextual and social aspects of cognitive theorising which are commonly neglected in canonical MSCI writing. It therefore leads to analyses that are answerable to the social worlds in which we debate, joke, play, and interact.

Of course, it is important to note that this approach neither refutes individual cognition nor denies that it is the individual who ultimately experiences blends as meaningful. Rather, it explicitly acknowledges the social and material dimension of meaning-making processes and addresses the problem of ‘hypothetical modelling’ common in canonical MSCI research by concentrating purely on those aspects of cognition which are publicly displayed and oriented to by the interactants themselves in processes of joint meaning-making activities – in short, this approach investigates cognition as enacted and revealed in interaction.

3.4. Data collection and data analysis

3.4.1. Domain for study: Impersonation humour

When browsing the literature on MSCI, it seems that blends occur everywhere: letters, numbers, grammar, religion and literature have all been described in blending terms. This great variety of tokens has lead to blending being labelled a “ubiquitous” (Bache, 2005) theory that is able to smartly account for everything and nothing at the same time. To address this problem, Pascual (2008:105) proposes the study of “blending types” in detail rather than “anecdotal tokens”, as the in-depth analysis of one type may contribute to a more comprehensive understanding of the structure underlying the type. Gibbs (2000) and Bache (2005), too, suggest that in order to address MSCI’s ‘ubiquity problem’, which targets the “too general nature of blending” (Bache, 2005:1617), “localized hypotheses” (Gibbs, 2000:349) should be devised so as not to lose sight of the phenomenon.

Following this line of argumentation, my analysis will concentrate on one specific phenomenon, i.e. impersonations in interactional humour. I have chosen this phenomenon firstly because the “other-character-mode” (Chilton, 2008:241) is a prime example of conceptual blending given that it integrates two personae (the

‘reference’ and the ‘presentation’, to employ the semiotical terminology coined by the Aarhus research group in cognitive semiotics) and thus results in discursive dissonance. Brandt (2010), who subsumes impersonations under the category of ‘virtual identification blends’ similarly argues that they constitute the “most prototypical” and “most intriguing” (331) blending type.

The clear presentation of semiotically distinct, i.e. dissonant, input scenarios evident in impersonation blends subsequently helps to address one of the fundamental problems in MSCI research: that of delineating Mental Spaces. In a scenario where one ‘real’ space (the impersonator, or *animator* in Goffman’s (1981) framework) signifies another ‘fictive’ space (the impersonated, or *figure* in Goffman), it is clear that two spaces are at work (cf. Brandt, 2008; Brandt & Brandt, 2005). These spaces are ‘theatrical’ in nature (cf. Brandt, 2005), and contain imaginable scenarios rooted in the interactants’ Common Ground.

Furthermore, such voice changes seem to be pervasive in interactional humour, where they constitute a “recognisable structure of activity” (Sidnell, 2006:380) that largely relies on the creation of unconventional meanings (cf. the incongruity theory of humour). Impersonations in particular and polyphonic humour in general hence seem to constitute a popular and successful blending type that is characterised by *polyphony*, *performance* and *playfulness*. The following pages set out to describe the phenomenon in further detail, drawing on research in linguistics, humour studies, performativity and literary studies. Following this, the phenomenon is discussed in an MSCI framework, thus demonstrating the extent to which impersonations recruit various levels of frame-integration complexity.

3.4.1.1. Polyphonic narrative, double-voiced discourse and impersonations

Although the importance of discursive dissonance for a clear delineation of Mental Spaces has been discussed above, this dissonance in impersonations takes on a further characteristic, that of polyphony. The term derives from Bakhtin’s treatment of Dostoyevsky’s works. In his *Problems of Dostoevsky’s Poetics* ((1963)1984), Bakhtin proposes that in order to account for Dostoyevsky’s highly complex work, his characters need to be thought of as speaking in unmerged, polyphonic voices. In what Bakhtin calls the ‘polyphonic novel’, the author is placed alongside his/her characters as an equal rather than merely as their creator. We are thus faced with a plurality of

consciousnesses, and the author, the characters and the audience together participate as equals in the creation of the novel's 'whole'. For Bakhtin, the author speaks with (rather than about) a character as someone who is actually present. This notion of polyphony is highly analogous to phenomena found in spoken interaction, where participants quote, mimic and otherwise make use of 'voices' other than their own.³³ Polyphonic narrative can thus also be witnessed in everyday interactions such as impersonation humour.

Impersonations as polyphonic narratives

Impersonations are fragmented polyphonic and "scenic" (Kotthoff, 2007) narratives that mainly achieve polyphonic voice through prosodic (Günthner, 1999) and bodily means.³⁴ In this regard, research on conversational narrative has shown that animated, multimodal enactments of reported speech often appear at a story's high point (Niemelä, 2010), thereby constructing a "sense of immediacy", making the story more vivid, and ultimately attaining conversational involvement (Pascual, 2006; Tannen, 1989). Humorous impersonation seems to present such high points in a 'story' whose context has, in accordance with Bakhtin's idea about the polyphonic narrative, to be co-constructed by the recipient: in impersonation humour, we are confronted with at least two voices, or consciousnesses: the creator's and the character's voices are co-present and contribute equally to the 'whole' of the emerging impersonation. It is the creator, through whose body and voice we perceive the character(s), who tailors his contributions to the specific requirements of the speech event at hand and the audience in question, and who, by means of his/her quoting, evaluates the quote at hand and takes a stance on it. Apart from the creator's voice, we also witness the created or quoted character, who can agree or disagree with the creator or 'quoter', who is the subject of his/her "own directly signifying discourse", and whose consciousness "is given as someone else's consciousness" (Bakhtin, (1963)1984:7). The voice of the created or quoted character is thus independent of

³³ See for example Tannen (1989), Günthner (1999) and Couper-Kuhlen (1999) on polyphony in reported speech, Kotthoff (2002, 2007) on double-voicing in playful, reported dialogue and (2009) on irony development, Koven (2002) on voicing in conversational narrative, Glick (2007) on voice and chronotope in stand-up comedy, and Priego-Valverde (2009) on double-voicing in failed humour.

³⁴ Following Günthner (1999), I subsume features of pitch, loudness, pausing and duration under *prosody*. The term *voice quality* will be used to describe paralinguistic features such as whispering, aspirated voice or falsetto voice. Bodily resources refer to features such as gesture, facial expression and body positioning and alignment.

that of the creator's.³⁵ Finally, there is the audience. In Bakhtin's view, the audience is an active participant, and adds its consciousness to the emergent whole of the impersonation. The importance of adding the recipient's "own consciousness" becomes especially transparent in impersonations that are 'best-of' performances and that with their elliptical speech style (Kotthoff, 2007) only provide fragments of the ensuing impersonation's context or "storyworld" (Günthner, 1999). They consequently metonymically represent, or provide dramatic illustrations of, whole stories. Impersonations thus exhibit that which Ohmann (in Wirth, 2002:26) calls a "world-generating function" (*welterzeugende Funktion*) in reference to poetic language, as the communicative context of the fragments provided in the impersonation needs to be imagined and co-constructed by the recipient – this is a feature that Tannen (1989) claims heightens conversational involvement. As a result, impersonations heavily rely on Common Ground, such as personally and culturally shared stereotypes and prototypes whose most salient features provide prompts for construal and enable coherence (Kotthoff, 2004, 2007).

While interactional humour has often been described as a case of sheer 'layering' (Clark, 1996), i.e. putting an additional semantic layer onto non-marked interaction, or as the creation of a 'play world' (Bateson, (1955)1972) that is somehow set apart from the 'real' world, the Bakhtinian notion of polyphony now allows the interactionally demonstrable co-presence of multiple 'consciousnesses' to be accounted for in impersonation humour. Together, they form an interwoven network, and do not, as other notions would have it, add additional layers or construct worlds that are set apart from 'reality'. On the contrary, all of these different voices are combined, and it is only through their interaction that the emerging 'whole' of the impersonation attains its communicative status. Impersonations are thus polyphonic.³⁶

In his treatment of monologue and dialogue, Bakhtin introduces yet another useful distinction, that of monologic and dialogic discourse. For Bakhtin, monologic voice is single-voiced, and hence does not explicitly recognise other people's words. Dialogic discourse, on the other hand, is double-voiced, and thus contains explicit

³⁵ As Glick (2007) points out, it should be noted that due to the performative context in which the humorous exchanges under investigation in this project are embedded, there are already at least two 'voices' present: that of the professional entertainer and that of the individual. Impersonations thus add a third, and sometimes even fourth voice.

³⁶ Polyphonic humour is not restricted to impersonations alone, as other humour genres such as irony are also clearly double-voiced. This means that impersonations only constitute one type of polyphonic humour.

references to words that have been uttered before. Double-voiced discourse adds “a new semantic intention into a discourse which already has, and which retains, an intention of its own” (Bakhtin, (1963)1984:189). The ‘original’ utterance is thus “decontextualised” and becomes “recontextualised” in a new speech event (Günthner, 1999; Koven, 2002). As a prime example of double-voiced discourse, Bakhtin mentions parody, which he defines as adapting someone else’s discourse, but introducing “into that discourse a semantic intention that is directly opposed to the original one” (Bakhtin, (1963)1984:193).

As with polyphony, the concept of double-voiced discourse can again be adapted to account for impersonation humour. The impersonator either directly quotes or introduces a fictive quote that s/he then equips with a new communicative intention, that of displaying the quoted sequences as a laughable. Impersonations are thus a means of taking a stance, of publicly commenting on the character, events and utterances presented (Günthner, 1999; Labov, 1972a).

Impersonations in interactional humour are, in sum, fragments of polyphonic and therefore dialogic narratives that only provide brief, ‘best-of’ insights into a storyline that has to be co-constructed by the audience. They can hence be subsumed under what I propose to call *polyphonic humour*. Polyphonic humour gains its special qualities from the interactional co-presence of multiple voices that taken together account for the ‘whole’, the communicative effects of the impersonation. Additionally, polyphonic humour is dialogic in that the original communicative intention, the illocution of the (fictive or non-fictive) quoted segment is replaced by a new semantic intention.

3.4.1.2. Performance and theatr(ic)ality in impersonations

Mimicry, parody and other forms of impersonation humour all show one special feature that sets them apart from other forms of polyphonic discourse, such as mere quoting or irony: they are emerging, embodied performances that are playfully and comically keyed (Kotthoff, 2006, 2007). Impersonations can therefore be regarded as ‘dramas’ that are staged before an audience. This staging capacity is what differentiates them from joke telling: where a joke relies solely on verbalisation, impersonation performances can utilise bodily illustrations. As Kotthoff (2004) argues when considering the sketch, the strength of humorous performances lies in its

theatrical element. Impersonations therefore rely heavily on the comical effects that embodied performances and stylisations can cause, as it is just such animated, theatrically and dramatically staged performances that allow us to experience events more intensively through the use of various semiotic resources and media (Sidnell, 2006; Wirth, 2002).

Before delving deeper into the finer details of humorous performances, it will, however, be necessary to delineate the concept of ‘performance’ itself, as this notion (German: *Performanz* and *Performativität*) has become an almost ubiquitous concept within the humanities.³⁷ For the sake of clarity, at this point I would like to adopt Fischer-Lichte’s (2002) treatment of performance, which is linked to theatricality³⁸ and staging³⁹. Coming from a theatre studies background, she conceives of performance as the theatrical staging of a role-swap, and adds that performances are acts of presentation through bodily and vocal means before an audience, and which rely on the interplay of staging, corporality and perception. An utterance is transported from the “living environment” to a “theatrical stage”, and during this process undergoes modulating transformations. According to this definition, then, theatricality and performance constitute a perceptual and semiotic mode (Fischer-Lichte, 2002) that shares the notion of iterability⁴⁰ and modulation with Bakhtin’s (1984) concept of double-voiced discourse. However, performance is not all about aesthetic pleasure; rather, it serves two other functions as well: framing and typification⁴¹, which can

³⁷ Wirth (2002:9) identifies four key concepts that performance can denote: (1) the “serious” performance of speech acts, (2) the staged performance of theatrical or ritual acts, (3) the constitution of imagination, and (4) the material embodiment of messages. He does, however, note three tendencies which the concept seems to develop within culture studies: *theatricalisation* (*Theatralisierung*), *literalisation* (*Literalisierung*) and *medialisation* (*Medialisierung*).

³⁸ The term ‘theatricity’ is here, in accordance with Alter (1981), utilised in order to avoid the negative connotations attached to the more common notion of ‘theatricality’. Under theatricity, Fischer-Lichte (2002:299) subsumes four aspects: *staging* as a specific semiotic mode, *corporality*, *perception*, and *performance*. Theatricity, she claims, only applies when all four aspects are present.

³⁹ Fischer-Lichte (2002:295) defines staging (German: *Inszenierung*) “as the application of specific aesthetic practices whose performative and corporal ‘excess’ achieves a certain effect” (my translation).

⁴⁰ Here used to denote the creative recontextualisations of ‘replica-tokens’ (Wirth, 2002:52).

⁴¹ Typification according to Schütz and Luckmann (1979) refers to the human tendency to perceive and structure the world in ideal types for which typical kinds of behaviour are assumed. In typification processes, pragmatic reduction takes place and idiosyncratic features of a certain type are ignored, and only the attributes relevant to the specific situation at hand are considered. *Typification* will here be used synonymously with *stereotypisation* in cases where groups rather than individuals are portrayed (Plum, 1998).

most concisely be described as the coordination of Mental Spaces which building in/incorporate interactional meaning-making.

First of all, the resources ostentatiously and markedly exploited in impersonation performances are used in the coordination of inferences. Ostentatious, marked performances thereby serve a framing function. Within the “theatre frame” (Kotthoff, 2007; Wirth, 2002) construed in impersonations, performative gestures take on specific theatrical purposes. Their transformation from one context to the other requires reflection on both their original context as well as their ‘new’ one. It requires the simultaneous maintenance of the inputs contributing to the integrated performance.

Secondly, mimetic performative resources such as body positioning, facial expression, code-switching or prosodic cues serve as prompts to either construct a particular type (Günthner, 1999) or to mimic previous discourse. By playing with and relying on (assumptions about) shared knowledge regarding stereotypical ways of speaking and acting as well as of cultural prototypes, performed impersonations thus act as cues for the co-construction of both the “storyworld”, or “reporting space” (Niemelä, 2010) in which the impersonation is set and the types inhabiting said world or space. In impersonations lacking a diegetic introduction, the storyworld is not described, but evoked solely through stereotypically performed details that make use of bodily semiotics in addition to lexico-syntactic and prosodic features. Here, a high degree of typification is required, and subtleness must be avoided, as the scene needs to be quickly understood by the recipient (Kotthoff, 2004); after all, impersonations are meant to depict and demonstrate referents rather than describe them (Niemelä, 2010). In this respect, impersonations with their exaggerated imitations of people and hyperstylisations of cultural prototypes and stereotypes resemble caricatures and cartoons.⁴² Although these art forms arguably constitute predominantly visual and static media, caricatures, cartoons and ridiculing impersonation nevertheless share striking similarities. First of all, the humour of all three types heavily relies on the interplay between two semiotic channels, the verbal and the visual. Psychological research has shown that images exert a stronger and longer-lasting influence than words (Plum, 1998), which might serve to account for the attractiveness of impersonations. As Samson and Huber (2007:14 in Tsakona, 2009:1172) point out

⁴² While cartoons are “drawn jokes”, caricatures always imply criticism in the broadest sense (Plum, 1998:28).

when discussing cartoons, bodily semiotics can augment the verbal channel in three ways: by simply illustrating a verbally told joke without adding to the humour, by providing additional information to the lexico-syntactic layer and thus contributing to the humorous effect, and, thirdly, by providing essential contextual information without which humour would not be possible. In a similar fashion, bodily performances elaborate on the verbal channel and vice versa. It is through the interaction of these two channels that hyperdetermination (Tsakona, 2009), in effect the facilitation of high density of metonymic and metaphoric inferences⁴³, can occur. Hyperdetermination and the interaction between visual and verbal semiotic modes constitute impersonation humour's asset.

Secondly, polyphonic impersonations show resemblance to caricatures, which can be defined as visual "representations [...] in which distinctive aspects are exaggerated" (Rhodes & Brennan, 1987:474). In caricatures, the process of simplification takes place, trivial features are largely neglected or even omitted while the most distinctive features of a character or type, typically the ones that deviate most from an idealised norm, are magnified, hyperbolically intensified, and thus emphasised as a means of facilitating recognition. In impersonations, the depicted type or character's most salient features, the features that most deviate from the idealised norm, be it body posture, facial expression, dialect or voice quality, are exaggerated in the performance to make recognition possible. Furthermore, caricatures, in much the same way as impersonations, present the depicted as laughable (Plum, 1998), as deviating from the expected norms adhered to by the in-group.

Humorous impersonations are therefore caricaturing performances, staged role-swaps which are achieved through bodily and vocal means. Changes in body posture and alignment, facial expression, pitch, code and register are all semiotic performative resources commonly employed to signal a type or person and the world inhabited by this person during impersonation performances. However,

⁴³ The linguistic and visual cues provided in caricatures, cartoons and impersonation performances function as triggers for metonymic and metaphoric inferences, as recipients are able to access aspects of shared knowledge or frames solely by alluding to a salient aspect of that frame. For this to happen, at least one visual or verbalised constituent of the cartoon, caricature or impersonation needs to be sufficiently salient (Brône & Feyaerts, 2003) in order to ensure activation of the intended frame. Due to the great variety of inferences possible in humour that relies on multiple semiotic channels, variations in humour appreciation can be accounted for: recipients may choose to focus on one or two aspects and draw inferences from them, and choose to neglect others (Tsakona, 2009).

impersonations are only ever successful if they are perceived as being funny or witty. Here, it is of central importance that the recipient receives enough cues to co-construct the narrative and types underlying the performance.⁴⁴ The stylisation of the character, be it a fictive or a ‘real’ one, thus needs to be easy to identify for the recipient if they are to achieve intersubjectivity and attain and ensure Common Ground (Clark, 1996; Kotthoff, 1998, 2007).

Humorous impersonations are performances in that they transform and modulate utterances through the ‘excessive’ use of bodily and vocal means before an audience. Performative resources thus provide cues as to the character or type portrayed, and at the same time constitute a framing device that helps the recipient to construe the speech situation at hand as polyphonic and humorous and playful in nature. It is this latter point that I shall be turning to now, as impersonations attain their framing as humorous through their specific keying as play rather than through performance features such as code-switching or ‘crossing’ (Rampton, 1999) into varieties other than one’s own.

3.4.1.3. Impersonations, play and humour

Since the groundbreaking work of Huizinga (1949) and Bateson ((1955)1972), a plethora of research on play has become available. Summarising the discussions about the nature of play would exceed the scope of this chapter, which is why, for the sake of brevity, I would like to use Goffman’s (1974) comprehensive definition of the “rules” that “transform serious, real action into something playful”:

- a. The playful act is so performed that its ordinary function is not realized [...].
 - b. There is an exaggeration of the expansiveness of some acts.
 - c. The sequence of activity that serves as a pattern is neither followed faithfully nor completed fully [...]
 - d. A great deal of repetitiveness occurs.
 - e. When more than one participant is to be involved, all must be freely willing to play, and anyone has the power to refuse an invitation to play or [...] to terminate the play once it has begun.
 - f. Frequent role switching occurs during play [...].
 - g. The play seems to be independent of any external needs of the participants [...].
 - i. Signs presumably are available to mark the beginning and termination of playfulness.
- (1972:41-43)

⁴⁴ In Fischer-Lichte’s (2002:283) words, “Jeder Zuschauer erschafft sich seine eigene Aufführung.” Theatrical performances only provide prompts for the creative construction of scenic narratives and/or types in the recipient.

Regarding impersonations in humour, it is obvious how they constitute playful performances: apart from being double-voiced due to the fact that the (projected) original communicative intentions are transformed and recontextualised, we find exaggerations, dramatisations and ‘best-of’ performances that ignore the (projected) original sequential pattern. More importantly, the impersonations are playful, as they are not seen as contributing to ‘serious’ communication, or rather communication that is employed to attend to “external needs”. Furthermore, impersonations are based on role-switches, and need to be systematically signalled as set apart from non-play and other voices through the use of disjunct markers and comical keying.

However, not all play and not every performance is humorous or makes us laugh. As play, a definition of humour as a remains elusive, and has for millennia been a major subject of dispute for philosophers, psychologists, linguists, anthropologists, and neuroscientists alike. Humour has many facets, and relies on the interplay of cognitive, affective and physiological systems and socio-cultural factors (Marin-Arrese, 2008). Within the large body of research available on humour, three major theoretical strands can be identified (Attardo, 1994; Brône, 2010; Zillmann, 1983):

The first is *cognitive theories*, which focus on the notion of incongruity (or: inconsistency) and, consequently, its resolution. Incongruity theories seem to be among the least disputed within the context of humour research, and are based on the assumption that there are two different frames or scripts that are mutually incompatible, but have a common element that makes it possible to shift from one frame to another. A recipient normally reduces information to the most salient script, and proceeds in this script until s/he encounters a semantic problem which forces the recipient to access a different script in order to solve the communicative obstacle. This previously hidden script allows for understanding, and hence leads to a feeling of surprise, satisfaction and, ultimately, laughter. Historically, research in this vein has been conducted by philosophers including Arthur Schopenhauer, Søren Kierkegaard and Immanuel Kant, authors such as Arthur Koestler and psychologists (Jerry M. Suls being the prime example). Influential cognitive humour theories have been proposed by Victor Raskin and Salvatore Attardo (e.g. Attardo, 1994, 2001; Raskin, 1985), whose *Semantic Script Theory of Humor* and *General Theory of Verbal Humor* have proven to be highly influential in linguistic research. Incongruity humour has previously been described as relying on processes of conceptual blending (cf. e.g., Brône, 2010;

Coulson, 2005), as the theory is especially suitable for depicting the simultaneous ‘juggling’ of distinct, incongruous scenarios.

The second theoretical strand is *social theories*, which discuss issues of superiority, aggression and disparagement in humour. These theories concentrate on the social and interpersonal dimension of humour and the negative attitude of the producer of humour towards the target. Research in this field has been conducted by Charles R. Gruner, Henri Bergson and, historically, by philosophers such as Aristotle in *Poetics*, Plato in *Philebus* and Thomas Hobbes in his work *Human Nature*. According to social theories, humour can be perceived as an aggressive game during which “wittiness becomes a symbol of intellectual and social superiority” (Brône, 2008:2029). Gruner (1997) has developed a game-theoretic account of humour which yields important insights into the understanding of the social functions of conversational humour.⁴⁵

Psychoanalytical *tension-release models* constitute the third strand. Inspired by Sigmund Freud (1905) and Herbert Spencer, these models focus on structures and psychological processes which lead to the production of laughter. In Freud’s theory, laughter and humour provide a release mechanism for suppressed and socially unacceptable behaviour, such as taking pleasure in another’s misfortune and shortcomings or talking about sexuality. The theory is thus especially appropriate for accounting for obscene humour, which in Freud’s view lead to amusement “because it constitutes a transitory overcoming of sexual repressions” (Zillmann, 1983:99).

Although all of these theories seem to be able to sufficiently address specific forms and levels of humour appreciation, it hardly comes as a surprise that no single theory has yet managed to satisfactorily account for all humour found in human interaction. Integrative models have, however, been proposed, such as the one put forward by Suls (1977). Yet these models still leave many questions unanswered.⁴⁶ As such, impersonations could certainly be discussed from all perspectives, most notably incongruity and superiority theories. According to incongruity theories, the humour emerging from impersonations could result from the clash of the two voices, those of the the parodist and the ‘parodee’. The humour arising from joint fantasising (Kotthoff, 2007), or more specifically the performance of fictive stories and cultural

⁴⁵ For an in-depth historical overview of research adopting superiority theories, see Zillmann (1983).

⁴⁶ Regarding Suls’ (1977) attempt at unifying superiority and incongruity theories, for example, Zillmann (1983) criticises the theory’s apparent failure in explaining the very reasons for the arousal of mirth.

prototypes, may rely on culturally incongruous characters and roles and the underlying celebration of shared cultural knowledge.

Superiority theories, on the other hand, could argue that impersonations facilitate the implicit evaluation of the original voice, which could allow for a ‘put-down’ humour element to be added to the impersonation. After all, it is such ‘direct’ quotes as those found in impersonation humour that enable “implicit person characterisation” (Kotthoff, 2007:268), and it is through such implicit characterisations and typification that evaluative stance can be communicated. Here, issues of affiliation and disaffiliation, of cultural identity and alterity, seem to play a role. Funniness arises not from the belittlement of any particular group or person, but from those with whom we feel little or no affiliation. We thus seem to prefer to ridicule, to laugh at a perceived outgroup or person, but laugh along with people we feel we affiliate with. As a result, the appreciation of humour seems to be a matter of disposition and affiliation (Zillmann, 1983). In this regard, the humour arising from impersonations might depend on the sub-genre: mimicry and the restaging of past dialogue (Günthner, 1999) uttered by persons present on the one hand show similarities to forms of teasing by focussing on and publicly exposing the perceived weaknesses of the ‘butt’ that is present in the interaction. One could thus argue that mimicry is a means of exerting power and of reinforcing social norms and normative morality by displaying deviant behaviour (as reflected in the original utterance) as laughable within the group. Teasing and other forms of put-down humour, such as ‘banter’, have sometimes been labelled as being ‘pro-social’ (Boxer & Cortés-Conde, 1997; Habib, 2008), as they are playfully keyed, and can be used to signal intimacy within established groups by indicating that the interlocutors are so intimate that they can ridicule each other without having to fear any serious consequences.

However, other psychological research has shown that “teasing is more a means of achieving superiority than a means for establishing solidarity” (Tragesser & Lippman, 2005). The teaser in this view is perceived as dominating the ‘butt’, as elevating his/her own status by displaying social and ‘moral’ superiority. In the end, the perceived level of aggressiveness of teasing impersonation humour will depend on a large number of variables, such as the degree of group integration, one’s status within the group, and the extent to which the teasing is framed as play (Plester & Sayers, 2007). Mimicry and the restaging of past dialogue are nonetheless a means of reinforcing the social norms and values that are held as valid within the group, as such

jocular abuse of people present is one way of presenting deviant behaviour as inappropriate and laughable. In parody, on the other hand, the ‘butt’ is not personally present, but represents the deviant behaviour of an out-group that is perceived as ‘the other’. Parodies thus constitute a means for establishing and celebrating in-group status and values through shared humour⁴⁷ by indexing the ‘other’ and projecting disapproved values and characteristics onto an out-group (Davies, 1982, 1998). Issues of social identity and alterity, of elevating ones own value system over that of a perceived out-group, come into play here.

I would also like to add another aspect that may contribute to the humorous effect of impersonations and which is not sufficiently addressed by the humour theories illustrated above. It occasionally seems to be the case that the hyperstylisation and typification of cultural prototypes and culturally salient, famous personalities alone seem to suffice for the evocation of mirth among recipients. Here, Kotthoff (2004), arguing with Bergson (1900), claims that it is the mimetic approximation of the impersonation that is responsible for the humour and resulting pleasure. She adds that such “comedy humour” in this respect shows resemblance to antique comedy, which also exhibits a high degree of typification responsible for evoking mirth. A similar explanation was proposed as a means of accounting for caricature, where Gombrich (1978 in Plum, 1998:41) emphasises the importance of resemblance and equivalence. In essence, while the constituent parts of a caricature pursue divergence, as a whole it seeks equivalence to the ‘original’. As Aristotle has already pointed out in his *Poetics* (1996), humans thus seem able to find pleasure in mimesis alone, in the recognition of the object behind skilful imitations. This point, of course, also leads to humour becoming a risky, potentially face-threatening⁴⁸ test for understanding (Sacks, 1974), as not being aware of the referent in humour will lead to a failure in showing and celebrating in-group status.

However, it should be kept in mind that humour is still a highly complex and situational phenomenon whose amusing effect can be attributed to numerous idiosyncratic and highly context dependent factors which can hardly be fully accounted for; rather, they need to be investigated on a case-by-case basis.

Conversational humour is full of playfully transformed and humorously keyed references to previous discourse(s), of citations, allusions and references to other styles

⁴⁷ Cf. the notion of Joking Cultures (Fine & DeSoucey, 2005).

⁴⁸ See e.g. Haugh (2010) and Zajdman (1995) for research on face and humour.

of speaking, and as such provides a means for commenting on social norms and indexing identity. It can therefore be said that humorous impersonations subsume polyphonic and multimodal performances of fragmented, non-contextualised narratives that are comically and playfully keyed, and rely heavily on shared knowledge, stylisation and typification. Humorous impersonations are locally occasioned, and systematically signalled as distinct from previous discourse, but are also inextricably intertwined with and made relevant by adjacent talk, exhibit a specific interactional structure, and are subject to practices that signal the end of the performance.

3.4.2. Impersonations as conceptual blends

Impersonations are prime examples of conceptual blending, as two distinct voices are present and one persona is presented through the body of another. In these “Virtual Identification Blends” (Brandt, 2010:91) the impersonated is understood as ‘being’ the impersonator. Yet while the impersonator and the impersonated might be perceived as identical during the performance, they are not identical once the inputs become available and qualitative inferences emerge. Decompression is consequently of vital importance in impersonation performances.

Imagine a person impersonating a character from a horror movie by quoting a salient, central line uttered by that character; this person not only copys the verbal performance, but also assumes the body posture and facial expression displayed by the original character. This impersonation constitutes the blend, yet for it to be meaningful the performance has to be decompressed in order for the story world of the performance to become present. In impersonations, the character in the blended performance metonymically stands for the narrative space⁴⁹ (Dancygier, 2008) which

⁴⁹ Dancygier (2008:54) conceives of a narrative space as a mental space that is required for story comprehension. The text itself only provides partial elements of the narrative space, which is why it necessitates further structure and elaboration to provide full(er) understanding. As she analyses complex narratives, such as Margret Atwood’s *The Blind Assassin*, Dancygier perceives of narrative comprehension as a hierarchical blending process. The narrative space is, in her model, at the topmost level, and integrates other Mental Spaces that were introduced to the text and needed to be integrated with other spaces to attain an emergent, coherent story.

While the notion of a narrative space as the blend of a variety of text elements is conducive to the analysis of complex literary narrative, my use of the term Narrative Space does not presuppose blending operations. Rather, I perceive of the narrative space as a mental space that is set up for meaning construction in narrative, and hence is subject to certain presumptions, such as event structure (e.g. presentation of a conflict) and the furnishing of further elements as the text evolves.

s/he inhabits. In such an impersonation, a whole story is compressed into a few lines uttered by one of its characters. In order to access this narrative space, the interlocutors are required to decompress the performance they are confronted with.

Decompression is required for the polyphonic nature of the blend to become manifest. It is only through the ‘unpacking’ of the blend that the co-presence of two voices becomes evident. The interactants firstly need to be (made) aware that the performer is not speaking as ‘herself/himself’, and secondly need to be able to infer the role which s/he is playing. The blend is thus an integrated performance that compresses the represented and the representer into uniqueness via the vital relation of representation (cf. Fauconnier & Turner, 2002:97f). However, the polyphonic dimension of this blend only becomes discernible when co-participants are able to unpack this performance and identify its input domains, as figure 12 illustrates. Polyphony consequently arises from decompression and the resulting interplay of the input spaces that are brought together in the integrated performance.

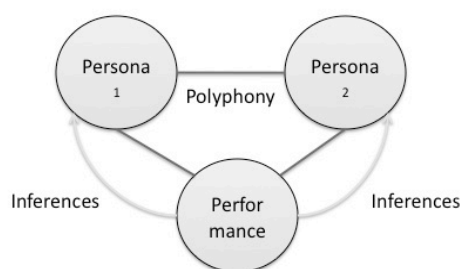


Figure 12: Polyphonic blends

Mirror, Single- and double-scope impersonation blends

Impersonations make use of various types of blending, from mirror to double-scope⁵⁰, in the classification offered by Fauconnier & Turner (2002). In impersonation blends, the dividing lines depend on two frame structuring aspects: *discourse scenario* and *semiotic codes*⁵¹ employed for meaning-making purposes. A human impersonating another human of the same discourse scenario, as in the case of mimicry or the restaging of

Unlike a blend, there are no cross-space mappings; rather, elements are added to the narrative space *itself* as the story unfolds during telling.

⁵⁰ As simplex blends are not considered ‘true’ blends for the reasons outlined in chapter 2.3., they will not be dealt with in the analyses.

⁵¹ By semiotic codes I refer to conventionalised processes by which signs are attributed meaning in a given speech community. Due to each semiotic code having various sub-codes, frame clashes are placed on a continuum from highly incompatible (e.g. impersonations of horses or chickens) to fully compatible (no clash of semiotic codes, as the same accent and pitch are used in the unmarked language of both characters, i.e. the impersonator and the impersonated).

past discourse, is thus classified as a mirror blend case. In mirror impersonations, no frame clashes in discourse scenario and the signs used in meaning-making occur, as the impersonator remains oriented to the original discourse world, and uses the semiotic code already employed in the original utterance.

In single-scope impersonations, however, humans impersonate another human from a different (fictive or non-fictive) discourse world. Here, frame clashes occur on the discourse scenario level, and is rather limited on the semiotic codes level. The emerging blend is structured by the fictive discourse world alone.

In this typology, then, humans impersonating machines or plants constitute cases of double-scope blending, as frame clashes occur both with regard to discourse scenario and the semiotic codes employed by the impersonated. Machines or plants do not use human means of communication, yet impersonation blends rely on the visual and verbal integration of input domains, and are acted out, or rather presented, through a human body. This means that semiotic codes employed for the representation of a non-human sign-system are human. For this reason, impersonations that equip plants or other non-human objects and beings with a ‘voice’ face frame clashes that ‘human’ impersonations do not. The emergent structure of the impersonation blend contains elements from both semiotic and discursive worlds. It is double-scope.

Table 1: Typology of Impersonation Blends

Blending Type	Discourse scenario	Semiotic system
Mirror Blend (e.g. restaging of past discourse)	Same discourse scenario in both input spaces, as designated person from CDS is portrayed → no major frame clashes	Human semiotic system in both input spaces → no major frame clashes
Single-scope Blend (e.g. impersonation of non-present fictive human)	Discourse scenarios are clearly distinct from each other; impersonated character inhabits another (fictive or non-fictive) space that contributes the emergent frame structure	Human semiotic system in both input spaces → no major frame clashes
Double-scope Blend (e.g. impersonation of a non-human or inanimate entity)	Discourse scenarios are clearly distinct from each other; the space inhabited by the impersonated contributes only <i>some</i> frame structure to the blend.	Non-human sign process portrayed through human semiotic resources → major frame clashes

As Table 1 illustrates, in this project the point of distinction between single and double-scope blends is drawn along the dividing line of discourse worlds and semiotic codes used in joint meaning-making processes. Whereas blends that involve inanimate objects failing to master human communicative codes will be classified as double-scope networks, blends that recruit (fictive or real) humans inhabiting clearly distinct

scenarios will be classified as single-scope. The restaging and mimicking of previous discourse is, as it is set within the same discourse scenario, classified as mirror-blend.⁵²

Why study impersonation blends?

As humour is a deeply social phenomenon, it relies heavily on the celebration of shared knowledge and values. For this reason, impersonations require careful interactional coordination and the recruiting of shared knowledge frames for their polyphonic nature and, consequently, the story world presented to become evident and shared. Investigating polyphonic humour in interaction should therefore help to address key problems in the analysis of conceptual blending:

- *Ubiquitous blends*: Blending theory has been employed to account for a wide variety of phenomena, and has subsequently been criticised for being unable to provide much insight into the question investigated. By addressing a special blending type in its interactional context rather than assessing eclectic, isolated examples from diverse fields, potent and meaningful results can be gathered that illustrate the theory's productivity without neglecting the specifics of the phenomenon it describes.
- *Delineating Mental Spaces and inputs*: As impersonations are characterised by two distinct and interactionally explicitly signalled characters inhabiting discrete epistemic and semiotic domains, the identification and delineation of input spaces is generally indisputable. In impersonations, the interactants explicitly orient to the distinct personae presented whilst simultaneously maintaining and aligning to different discourse scenarios. This explicit orientation to discursively dissonant scenarios helps to avoid the rather arbitrary space partitioning found in canonical MSCI.
- *Processes vs. products*: Rather than starting at the 'product' of blending and trying to reconstruct the cognitive steps taken to arrive at this product, investigating impersonations in interactional humour allows for the contextualised tracking of the processes behind interactional blending.
- *'Solipsistic' analysis*: As the focus lies on 'interactional blending', conceptual integration is first of all taken into the social (and hence visible) realm. Special attention can be paid to the step-by-step *co*-construction of meaning between

⁵² See chapter 4.1. for detailed case studies of each blending type.

performers and their audience. The accusation of focussing too much on the individual mind, of neglecting the social dimension of meaning-making, is consequently countered. Secondly, a social-interactional analysis that pays attention to the co-construction of meaning avoids hypothesising about psychological processes which occur unconsciously within milliseconds without the aid of appropriate neuroscientific tools. Instead, only blending that is explicitly signalled in interaction and temporally unfolds in interaction is considered in order to allow for empirically demonstrable results.

- *Empirical grounding*: Canonical blending analysis has been accused of relying on introspection alone and lacking empirical grounding. By studying blending as an interactional achievement in impersonation humour through the use of the methodology devised by an established research endeavour (EMCA), empirically valid results can be gathered.

A detailed interactional study of different impersonation types will thus provide the tools for testing the blending typology (RQ2) advocated by Fauconnier & Turner (2002) and describing the ways in which processes of conceptual blending are coordinated in interaction (RQ1).

3.4.3. Data Source

3.4.3.1. Studying interaction-as-broadcasted in *Never Mind The Buzzcocks*

The collections which this study recruits were taken from the British comedy panel game show *Never Mind The Buzzcocks* (NMTB), and were recorded between 2006 and 2010. NMTB is a long-running British quiz show (1996-today) which is produced by Talkback Thames and is aired on the BBC and other British channels such as DAVE⁵³. The show, which the BBC advertises as an “irreverent pop quiz”, is targeted at a young and mostly male audience. Pop music knowledge is used to test two competing teams⁵⁴, and the show has earned a reputation for providing especially

⁵³ DAVE markets itself as the “home of witty comedy banter” and shows mainly reruns of programmes that target a young, predominantly male demographic (16-34 years). Examples of shows include the car series *Top Gear* and the comedy sketch show *Little Britain* (BBC News 20/09/2007).

⁵⁴ The way in which the programme focuses on pop music shows the extent to which ‘knowledge’ in quiz shows is configured by cultural and social contexts (Holmes, 2006; Fiske, 1987). The

provocative and harsh humour that according to some commentators (e.g. Wignall, 2008) seems to constitute one of the show's major attractions. This conclusion can also be drawn when looking at the BBC's promotion of the show: "With the show's reputation for the unpredictable, guests are always primed for the occasional embarrassment and harmless ribbing as their knowledge of music is put to the test" (BBC Press Office).

The teams consist of two captains that like the host feature in most episodes. Each captain has two guests on his team that change from week to week. Guests are usually British (or Anglophone) celebrities of various denominations: pop stars, comedians, actors or even news presenters. The show's general appeal thus seems to result not so much from the competition element, but rather from the opportunity it provides to 'consume' a celebrity. As Kurzman et al. (2007:353) argue, postmodern celebrity culture 'honours' celebrities for their distinctiveness whilst at the same time trying to undermine said distinctiveness, thereby exposing their ordinariness. NMTB makes such consumption possible. Although the celebrities' achievements are recounted (some use the show to promote a new album, autobiography or film), they are mostly observed to be guessing answers, failing to recognise popular songs, becoming the 'butt' of somewhat opprobrious jokes, exchanging repartees, (in one case) getting drunk, and telling anecdotes. Their ordinariness and humanity can be traced, and they become what Schickel (1985) calls "intimate strangers" with whom the fan can enter into a unilateral relationship. Over the past few decades, such panel quiz shows have become extremely popular in British television, as other formats such as *A Question of Sport* (BBC, 1970-today), *QI* (BBC, 2003-today), and *Have I Got News For You* (BBC, 1990-today) all rely on similar concepts.⁵⁵

The participants are seated behind desks in what Kendon (1990a) terms 'Formation': all panellists and the audience present in the studio are arranged around a

epistemological basis assumed in NMTB illustrates the form of 'knowledge' believed to be valued by the show's target audience; in effect, knowledge drawn from "common social experience" (Fiske, 1987:267) is presented as being relevant and important.

⁵⁵ Quiz and game shows have a long tradition in British radio and television broadcasting. While many focus on 'schooled' knowledge (such as *University Challenge*), quiz shows that centre on issues of popular culture are less common, and constitute a more recent development. *Have a Go!* (BBC, 1946-1967) was among the first to venture into the realm of the popular and the carnival; like later *Never Mind The Buzzcocks*, it was also not primarily a quiz show, but a show that relied on the participants ('ordinary folks') recounting anecdotes and exchanging cheerful banter (Holmes, 2006). With NMTB and other current panel quiz shows, however, it is not the 'ordinary' people that do the guesswork and the story telling; it is celebrities.

central, triangular space and are thus able to co-orient to each other, as figure 13 illustrates.

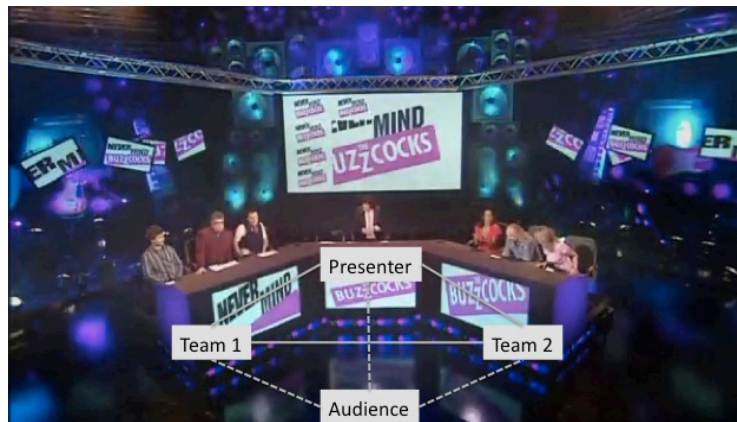


Figure 13: NMTB panel set-up

The triangular set-up is headed by the presenter, whose desk sits at the ‘top’ of the triangle and is framed by the desks of the competing teams. The teams’ captains are seated in the middle, with their two guests to their right and left, respectively. Opposite the presenter, and invisible to the TV audience, is the studio audience. The major screen, which is used to show music videos that are related to quiz questions, is positioned behind the presenter. Further screens are built into the teams’ desks, and it is these inbuilt screens that the teams mostly orient to when working on the quiz questions. The show opens with titles and theme music, after which applause from the studio audience is followed by the presenter (British comedian Simon Amstell and selected guest hosts in the seasons under investigation,) opening with a scripted monologue containing a few jokes that he delivers looking into the camera, ‘directly’ at the TV audience. The guests are subsequently presented. Video clips show the guests in their familiar occupational fields: performing in music videos, delivering stand-up comedy or presenting TV shows. At the same time, the host (in voiceover) introduces the guests with the help of a few, usually face-threatening comments that are acknowledged through (sometimes canned) audience laughter.

These comments already establish the tone of the show, and legitimise face-threatening jocular abuse. After these scripted introductory remarks, the games begin. The first game is usually ‘Sorry, no refunds’, in which teams have to guess why a given concert was cancelled. Alternatively, the games ‘What have we pixelated?’ or ‘whoop whoop it’s the sound of da police’ are played. As each team is asked a different

question, they are not in direct competition and no time pressure is exerted. The teams are free to choose from three possible answers provided by the host, and the guesswork provides ample opportunities for further exchanges of ‘banter’ that are once more explicitly acknowledged with audience laughter. The guessing is largely non-scripted (the guest may be given a set of jokes written for them), but is subject to editing. Should the team not fail to provide the correct answer, the question is posed to the other team; this team may then indeed offer the solution – this is an arrangement that is also valid for the games played later in the show. After this first game, the host puts forward several questions to his guests. These questions usually serve as a trajectory for further spontaneously-occurring jocular abuse. The games resume after these exchanges with the ‘Intros round’, in which two members of the team perform the opening beats of two pop songs without instruments or vocals for the third member to guess. This game also requires approximately ten minutes, and is again followed by an intervention of jocular-abusive chitchat chaired by the host. The next game, the so-called ‘Identity parade’, is then introduced. In this game, each team is presented with a line-up of five people in similar attire, and is tasked with determining which person was a former pop star. The line-up is not allowed to speak or move, which is why their appearance provides the only cues for guessing the correct answer. As a result, this game generally leads to an abundance of impersonations, as impersonations constitute one way of assigning identities to the members of the line-up. After this, the final game, ‘Next lines’, is played. The teams are provided with the first line of a song, and must provide the following one. Afterwards, the host announces the winner of the week’s show.

Despite this rather straightforward structure, point allocation is slightly arbitrary and opaque (as it is in all panel quiz shows). As panel game shows go, the overarching object of the show is not to win the quiz (there are no prizes either, as is the case in other television quiz shows, such as *Who Wants To Be A Millionaire?*²), but to entertain the audience in the studio and behind television sets through witty repartees, the telling and joint celebrating of personal anecdotes, and general ‘banter’. The high density of humorous exchanges, and in particular of impersonations, makes the show a highly suitable resource for studying the various aspects of humour in interaction.

However, within Conversation Analysis, TV broadcasts are sometimes considered to reflect ‘artificial’ language use, as the interactions might have been scripted or subject to rehearsal. Alternatively, the recordings may have been edited

before airing. Despite these legitimate doubts, quite a large number of research enterprises consult TV broadcasts as a source. These projects either focus on aspects of ‘institution talk’ (e.g. Heritage, 1985; Heritage & Greatbatch, 1986; Heritage & Roth, 1995; Greatbatch, 1988) or of general talk-in-interaction – arguing with Schegloff’s (1978) assertion that interactants do not shift to an entirely different apparatus of interactional procedures when being broadcasted, (e.g. Norrick, 2010b). Whilst the first group investigates the formalised speech exchange systems found in broadcasted interaction and focuses on the distribution of speaker roles and turn-taking rights, the latter argues that under certain circumstances, conclusions that apply to ‘more natural’ data can be drawn as well. To this end, careful considerations are required, and reflection regarding the ‘product-ness’ of the data should not be neglected. As videos are “locally and contingently produced” (Mondada, 2006:51), it is not ‘pure’ talk-in-interaction that is at the heart of the investigation, but rather *interaction-as-broadcasted*. As Mondada (2009:67) succinctly points out, video data in whatever shape or form do not constitute “transparent window[s] on social interaction”, but are “situated product[s] of video practices” that are subject to shooting and editing choices, such as cutting or reassembling shots, and post-production arrangements, such as the adding of canned laughter.⁵⁶ Thus, broadcasted interaction is not ‘just’ talk-in-interaction, but interaction that has already been interpreted in a certain way and is actively shaped by professionals:

Interactional order [in video data] is the local product of the participants’ practices engaged within interaction, the product of participants’ orientations to the camera, as well as the product of the video practices themselves. Participants produce the interactional order, either by engaging in the interaction or by filming and editing it. By paying attention to the actual timed details of the interaction, *cameramen and editors orient to the same linguistic and multimodal resources as the participants* and exploit them within the professional practice of configuring videos. Cameramen and editors mobilize local resources in order to achieve a specific accountability of what is videotaped, achieving an interactional organization which is not only recognizable but also rearranged, highlighted, and emphasized in a specific way, thus *reshaping* the ongoing interaction (Mondada, 2009:72; my emphases).

⁵⁶ In NMTB, an audience is present. However, canned laughter is added in post-production to further key/identify the interaction presented as humorous and encourage the television audience to share the sentiment. Research on laughter (e.g. Freud 1905; Glenn, 2003) has shown that laughter is (at least partly) socially mediated and that people are more likely to laugh when they believe their in-group is amused as well. Additionally, shared laughter does not occur simultaneously, but is prompted by one conversationalist who is then joined by the others. It is these interactional effects that canned laughter hopes to exploit (Platow et al., 2005) as a means of triggering a jovial effect among the TV audience.

Mondada's research has shown that TV directors of talk show formats locally orient to the same lexico-syntactic and visuospatial resources as the interactants in their interpretations and, consequently, editing decisions. The editing choices are recipient-designed according to the directors' interpretations, as the images shown and the information added (such as canned laughter) highlight embodied features of talk that are deemed relevant for the interpretation of the action at hand. This can, for example, be seen in the camera movements' relation to turn-taking. Here, the shots shown predict and follow turn constructions, as is confirmed in my analyses (see chapter 5.2.2.). When the turn is complete, the camera moves to the participant projected to provide the next turn, thus providing and highlighting the multimodal details the TV audience is thought to require in order to interpret the joint action as co-constructed by the director and presented to the TV audience.

Consequently, interaction-as-broadcasted is recipient-designed for a wide audience, and is rendered intelligible and accountable by television professionals. The action presented exploits the same grammatical and multimodal resources that the 'original' interactants relied on, but certain resources and sequences are highlighted and foregrounded by the professionals reworking the original recordings. The analysis of broadcasted data hence provides insight into the resources employed by members to accomplish a given action, but at the same time these resources can only ever be analysed as forms of social action whose further participants are the editors and directors who have actively reshaped the 'original' interaction.

3.4.3.2. Transcription

The transcription of talk-in-interaction, which graphically represents selected features of interactants' behaviour in conversation (Kowal & O'Connell, 2004), is integral to data analysis in CA, as in the process of detailed transcribing, previously unnoticed interactional phenomena can be identified. Consequently, transcriptions should faithfully capture the 'nature' of interaction in minute detail whilst at the same time making interactional details available for analysis. However, transcripts are not regarded as the primary data, but are a tool for analysis, a way of presenting the 'fleeting' nature of the phenomena under investigation in written form. Due to the complexity of spoken interaction, transcripts are nevertheless necessarily incomplete and impressionistic reductions, and constitute selective 'translations' (Ten Have, 1999)

of the actual interaction into written form; they mediate between the recordings – themselves already reductions, reconstructions and subject to recording practices and choices – and the images constructed thereof post-hearing or seeing. These ‘translations’ are therefore contingent on a number of choices relating firstly to the orthographic rendering of the words spoken,⁵⁷ and secondly the interactional details included in the transcript.

Due to the great variety of research questions and interests, a number of transcription systems are in circulation that each aim to represent certain interactional details. Most of these systems are based on the one developed by Gail Jefferson (as in for example Jefferson, 2004) dating back to the early days of Conversation Analysis. This system is a ‘vertical’ one, and aims to explicate the sequential nature of talk. In Jeffersonian transcriptions, one utterance is printed below another to show the onset of a new turn. This vertical alignment leads to the question concerning when to divide a line and start a new one. A variety of solutions have been proposed for this problem: phrasal or clausal units could constitute a line, as could breath units, intonation units or Turn Constructional Units (Ten Have, 1999).

Apart from this focus on the sequential nature of talk and said problems arising from the ‘vertical’ set-up, the Jeffersonian system has a further disadvantage. As it was developed in the heyday of the tape recorder, it is unimodal, focusing on the verbal channel only, and is hence not suitable for representing the multimodal⁵⁸, visuospatial dimensions of talk-in-interaction. Nevertheless, video data provides helpful information for interaction analysis, as it “preserves the temporal and sequential structure which is so characteristic of interaction” (Knoblauch et al., 2006:19). However, representing the synchrony and integration of visual and vocal semiotic channels and their structural differences provides a problem for transcribing audiovisual data. To address this issue, multilayered transcription systems like the

⁵⁷ The question here is whether the word ‘spoken’ should be orthographically represented in the standard variety, as it runs the risk of losing important interactional details such as code or style-switching, or whether phonological approximation to the actual delivery should be attempted, which would thereby risk the transcript’s readability. A common solution to the problem is to use ‘eye dialect’ (Kowal & O’Connell, 2004). Here, standard orthography and variety are employed most of the time, but the spelling is modified to represent utterances that strongly deviate from the norm (Ten Have, 1999), as in *gotcha* for “got you”. As the decision regarding which deviations to include seems to be a highly subjective one, only those variations and switches that interactants themselves notably orient to in the interaction under scrutiny will be included in the transcript.

⁵⁸ Multimodal is, following Allwood (2008), understood as referring to transcripts that hope to account for at least two “sensory” or “production” modalities and hence include aspects such as gesture and postural orientation in addition to speech.

robust, yet readable German Gesprächsanalytisches Transkriptionssystem (GAT and GAT 2) (Selting et al., 1998; 2009) were developed. With GAT, a ‘basic’ transcript can be augmented with additional information, following what Selting et al. (ibid) refer to as “Zwiebelprinzip”. These additional layers allow the representation of prosody, stress, pitch, tempo, rhythm, nonverbal actions such as sneezing, proximity, kinetics, facial expressions and gaze in various levels of detail. Nonverbal features can, as in *score writing* (Luckmann, 2006), be inserted below the line that represents the verbal layer, and annotated screen shots are inserted to further illustrate visuospatial aspects of talk-in-interaction. The temporal ‘location’ of these non-verbal activities is represented by dashes and lines connecting them to the corresponding position in the verbal layer. The adding of frame grabs or screen shots have of late prompted discussion on the changing nature of transcripts and transcription procedures. Knoblauch et al., for example, argue that a “visual mentality” (2006:16) seems to underlie transcription practices that rely on frame grabs; it is, they claim, a mode of analysis which depends more on “visualisation and imagination” (ibid) than on thick written descriptions and that might consequently impede on the analytical, observation-generating function which the transcription process facilitates.

Despite these doubts, partially ‘visual’ systems such as GAT are used by a number of other European and Anglophone researchers: Goodwin (e.g. 2003a, 2003b) works with annotated screenshots to illustrate the interactional role played by the body, gestures and gaze in interaction. In a similar vein, Mondada’s (2009) multilayered transcript, which she uses to analyse TV broadcasts, represents the talk as spoken in a first (numbered) line. In the (unnumbered) line below, the interactants’ multimodal actions are described. The third (unnumbered) line chronicles or shows in one or more screenshots or frame grabs the image(s) broadcasted at this point in the interaction. Mondada’s system thus allows the illustration of the temporal synchrony of unfolding actions.

Furthermore, multimodal corpora, such as the *Idiap Wolf Corpus*, have of late been developed. These corpora contain annotated audiovisual recordings of natural face-to-face interaction. While these corpora facilitate simultaneous access to recordings and transcription, they face similar problems to paper-based transcripts of audiovisual interaction. Thus, synchronising information pertaining to the modalities represented is still problematic (Allwood, 2008), but projects such as *Social Signal*

Processing Network (SSPNET) have enjoyed considerable success (see e.g. Bousmalis et al., 2011) in automatically identifying non-verbal cues.

Due to its varied research questions that require different levels of granularity in analysis, this project does not make use of automated systems, but largely follows the standards set out by the GAT system, and recruits additional ideas and systems to represent aspects that are especially relevant to the action studied, and interaction-as-broadcasted. Audience clapping will be represented by **xxx**, as suggested by Atkinson (1984), and preceded by **AU** in the agent column. In order to account for the temporal dimension, clapping that occurs simultaneously with spoken interaction will be indicated below the vocal layer with which it co-occurs using square brackets to indicate co-temporality. Clapping that occurs outside of other interaction will be measured and its duration given in brackets after the symbols indicating the clapping. Clapping that is notably louder than before will be indicated using capital letters (e.g. **AU: xxxxx**). Audience laughter, whether canned or more naturally occurring, will be indicated using the **@** symbol. Jefferson and other CA researchers (Jefferson, 1979; Jefferson et al., 1977; Jefferson et al., 1987) have emphasised the importance of rendering laugh particles onomatopoeically and syllabically (such as *huh-huh*, *hahaha*, or *heh-heh*) to better account for their various interactional functions. As the chorus that is audience laughter cannot be broken down into these subtypes, the **@** symbol, following the same standards as set out for demonstrating the temporal unfolding of audience clapping, will be used. Participant laughter will, however, be symbolised onomatopoeically and syllabically where possible. Furthermore, screenshots detailing the visuospatial dimension of the interaction-as-broadcasted will be supplied next to the line providing the verbal channel, and will again be synchronised to account for the temporal and synchronous unfolding of the interaction-as-broadcasted. Additionally, voice changes in impersonation humour will be signalled with **<@** at the onset and **@>** at the offset, as suggested by Couper-Kuhlen (1999) and employed by Niemelä (2010). An overview of the transcription symbols used can be found on p.xi.

3.4.4. Data analysis: Process

In a first step, all instances of impersonations occurring in the corpus were identified, transcribed and grouped according to impersonation type. Of the 109 impersonations,

23 were classified as mirror-impersonations, 62 as single-scope impersonations and 24 as double-scope impersonations.

The analyses in this project were twofold, consisting first of all of cognitive semantic and secondly of interactional analyses. The cognitive-semantic analysis depicts the sequences under investigation in traditional MSCI terms, albeit in lesser detail than found in the standard literature, as only those aspects that interactants orient to are included in the analysis (replete with an added Current Discourse Space to account for in situ cognition). This discursively-informed cognitive-semantic analysis sets out to describe how interactants coordinate processes of conceptual integration in situ, thus addressing RQ1. The interactional analysis subsequently investigates the extent to which interactants orient to the blending typology advocated by Fauconnier and Turner (2002), thereby focussing on RQ2. The following will explain the analytical steps taken in each ‘input’ discipline, MSCI and CA, before describing in further detail the synergetic analysis applied in this project.

3.4.4.1. Cognitive-semantic analysis: General principles

MSCI analyses consist of modelling the conceptual integration network that is assumed to exist behind a blend – a product of human creativity that is assumed to be based upon various inputs. The input spaces, mappings and compressions that are seen as contributing to this understanding are diagrammed and described. In a first step, the input spaces that contribute to the network are identified. In a next step, the analyst would normally look for shared structure in the input spaces; this structure would then be diagrammed in the generic space. For reasons outlined above, I will refrain from abstracting away from the input spaces and the depiction of a generic space. Following the description of input spaces, cross-space mappings, the “vital relations” that connect the various entities within the input spaces are identified. These relations are then explored with regard to their compression type. In a further step, the frame structure of the blended space is investigated to determine which conceptual integration type (e.g. single-scope or double-scope) is at work. The blended space is subsequently examined with respect to its emergent properties, the inferences it allows for, the governing and optimality principles at work, and the extent to which the blend gets elaborated upon later on in the interaction.

The result of the analysis is then depicted in a blending diagram. Williams (2004:27) points out that such diagrams are “necessarily idealised”, as “the actual integration network constructed by a specific person at any given moment will depend on many factors, including (among others) past history, the present setting and situation, and the current focus of discourse” – in short, aspects of Common Ground. Intersubjectivity, Williams continues, is only achieved as a result of alignment between the networks individually constructed by the interactants. As the approach taken here refrains from modelling the processes underway in the minds of the individual, only the emergent shared meaning and those aspects that the interactants publicly display in their interaction are depicted in the diagrams.

3.4.4.2. Interactional analysis: General principles

Like most qualitative research endeavours, CA does not subscribe to any one ‘silver bullet’ technique of analysing data. Rather than superimposing formal rules for analysis, CA encourages the utilisation of a ‘conversation analytical mentality’. This reluctance to incorporate formal rules, Bergmann (2004a) claims, derives from the assumption that in qualitative research, the appropriateness of a given method can only be assessed in hindsight, as too little is known about the issue investigated prior to the actual analysis (this state of affairs being the very reason for investigation in the first place). In qualitative analysis, canonical and mechanical methods, he goes on to argue, are prone to only ‘discover’ phenomena which are already well known. In accordance with this premise, EM and CA promote the development of a methodological approach that suits the phenomenon under investigation. To this end, the literature recommends that the process outlined below should be generally followed (Hutchby & Woffitt, 1998:94-95; Ten Have, 1999:103).

1. Through ‘unmotivated looking’ (Sacks, 1984) an interactional phenomenon that could be of interest is located. By ‘phenomenon’, CA does not understand linguistic objects as such, but rather the interactional effect that is accomplished. This interactional phenomenon then constitutes the first instance. Its occurrence, as Psathas (1995:50) points out, shows that “the machinery of its production is culturally available, involves members’ competencies, and is therefore possibly (and probably) reproducible.”

Therefore, any specimen is essentially worthy of investigation. This first instance is described in detail, and special focus is paid to sequential patterns. It needs to be shown that participants are oriented to the action described and display this understanding in subsequent talk. The instance that is studied is not seen as a “reflection of reality”, but rather “as part of the reality studied” (Ten Have, 1999:38).

2. Similar instances are identified in the data. The phenomenon or action that has been identified needs to be supported and exemplified through extensive data displays and descriptions. To this end, ‘collections’ are built that comprise larger sets of data. However, these collections are still studied as multiples of single instances. In CA, quantification does not constitute an alternative to the in-depth analysis of single instances, and is merely used to reinforce the robustness of an argument.
3. The description is tested on other instances in the data, and the account is refined to arrive at a formal description that can account for all examples in the collection. It should be shown how particular practices, procedures or devices are methodically used and normatively oriented to as a means of generating particular and recognisable actions. In this process, the analysis of deviant cases can be helpful, as deviant cases may make it possible to extract the *conditional relevance* of a turn or feature.

Although CA does not posit a ‘royal way’ or any ‘golden rules’, the research enterprise nonetheless insists on rigorous, general and formal descriptions of interactional phenomena.

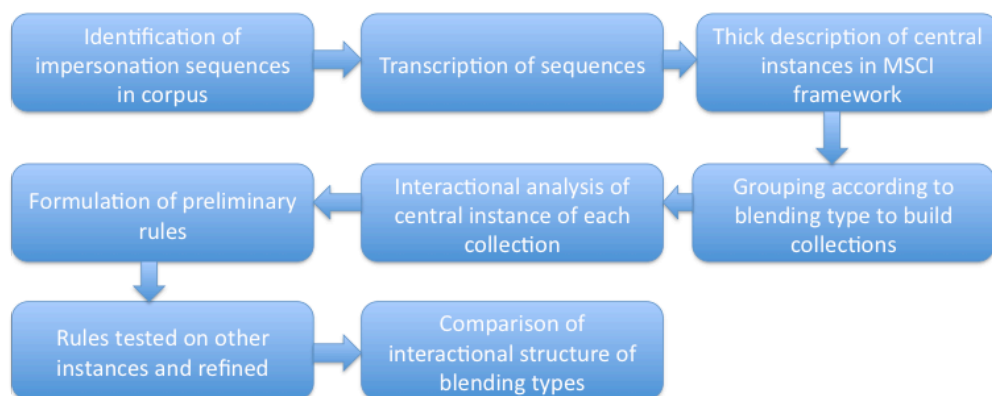


Figure 14: Data analysis: Process

This project integrates the two methodologies outlined above and follows the procedure outlined in figure 14. Instances of impersonation humour which show ‘external’ voice changes, i.e. assuming the voice of a person clearly different from one’s own persona (see chapter 3.4.1. for full characteristics), were identified in the data set and transcribed according to the standards set out in chapter 3.4.2.2. The instances were then grouped according to blending operation subgroups in order to build collections that contain instances of mirror, single- and double-scope integration networks, and discussed in accordance with the standards set out in CA. One central instance of each blending type was subsequently described and analysed within the framework of discursively oriented Mental Spaces and Conceptual Integration theory outlined in chapter 3.3. In a next step, the interactional organisation of one central instance of each collection was described. Following this, a second step describe the vocal (lexico-syntactic features, prosody, pitch, pausing) and the visuospatial channel (body alignment, gaze, gesture, facial expression) was taken in order to analyse the methods relied on by interactants to systematically and step-by-step signal the blend constructed. Following this, preliminary rules for each blending type were formulated that were then tested on other instances from the respective collection. During this process, the rules, methods and procedures identified were redefined to maximise generality. In a final step, the interactional structures of all blending types were compared in order to determine potential differences that would validate the distinction proposed in previous research.

The results of the analysis are presented in the next two chapters in a format that does not reflect the steps undertaken in a chronological manner; instead, the method of expression has been chosen to best reflect the findings gathered in the process. While chapter 4 will provide initial suggestions as to the issues raised in RQ1, chapter 5 is dedicated to RQ2. Chapter 6 will consolidate the results gathered and propose a refined model of in situ meaning-making which employs MSCl.

CHAPTER 4: IMPERSONATION BLENDS: CASE STUDIES

This section provides detailed case studies of mirror, single-scope and double-scope impersonation blends in an interactional Cognitive Semantics framework. The case studies first of all offer sequential analyses of the impersonations, addressing not only the means by which impersonators link impersonations to prior turn-by-turn topical talk, but also the overall interactional organisation of the performance. In a second step, the findings are subsequently utilised as part of a cognitive semantic analysis that provides detailed illustrations of the joint negotiation of conceptualisation processes.

As the data provides insights into interaction-as-broadcasted, the visual grammar of the show is also taken into account to provide a more holistic picture of how blending processes are negotiated in televised interaction (cf. chapter 3.4.3.1). Furthermore, each case study focuses on marginally different aspects in order to illustrate the intricate complexities behind each blending type and to avoid redundancies in aspects that are shared across blending types. A conclusion will summarise and critically discuss these initial findings; this will lead into the interactional analyses presented in chapter 5 providing more in-depth studies of the different means employed for jointly building impersonations and comparing these methods across blending types. In this manner, the case studies trace the steps taken to jointly set up impersonation blends in interaction-as-broadcasted, and offer further details regarding the terminology and criteria used for classifying blends in this project.

4.1. The Sports Commentator: A Mirror Blend Impersonation

Mirror impersonation blends all demonstrate similar frame structure and semiotic codes. All spaces contributing to the emergent impersonation blend are largely structured by similar events, and feature the same characters. This can be observed in The Sports Commentator impersonation.

The impersonation occurs during the intros round, in which guest panellist GL tries to guess a popular song from its introduction, which is performed a cappella by her fellow team members. Important to note is that GL is known (and has been introduced) as a sports presenter for the BBC.

NMTB SE23E02_2 3:30 (The Sports Commentator I)

01 GL: something like (-)



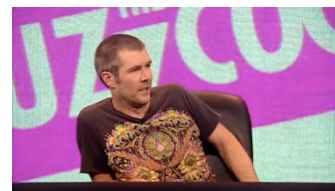
02 like the (.) FOUR tops



03 is it, (.) jackson five?



04 RG: you just gonna go with numbers,



05 AU: @@@ (4.0)



06 PJ: she thinks she's doing the REsults,

07 AU: @ (1.0)

08 PJ: <@ tops fou_↑`our;

09 jackson fi_↑`ive; @>

10 AU: @ (1.0)



4.1.1. Interactional organisation guides conceptualisation processes

The sequence begins with GL guessing the name of the band whose song (or rather the first beats of the introduction) has just been performed by her fellow panellists. Seemingly by accident, GL suggests two consecutive bands with a number in their name. Due the conventions of the discourse setting, which explicitly requires (adversarial) joking, RG's 'joke-first practice'⁵⁹ (Schegloff, 1987) is a relevant next

⁵⁹ Schegloff (1987:212) defines the 'joke first' practice as "before providing the serious 'next' which is sequentially relevant, the participant provides a joke first."

action to GL's first-pair part (FPP, lines 01-03). Yet his joke, which is acknowledged by the audience via laughter and which is deemed a relevant second pair-part (SPP) to a joke (ibid), digresses from and disrupts the topical turn-by-turn talk for several turns, and leads the conversation in another direction. PJ appropriates the central component of RG's joke, which profiles the numbers in the band names offered by GL, and uses it as a springboard, so to speak, to deliver an impersonation; this draws attention to an aspect of shared knowledge about GL that had thus far not been oriented to: her 'day job' as a sports presenter. The jokes therefore complement each other, and the impersonation is "triggered" (Jefferson, 1978:220) in the unfolding of turn-by-turn talk. The impersonation itself is, in other words, *locally occasioned*, and constitutes a relevant next action to a concluded (i.e. FPP followed by SPP) joke sequence.

Again, laughter is a relevant next action to the humorous impersonation, and, as none of the other participants contributes another joke, topical talk is resumed by means of RG delivering the SPP to GL's original FPP. As such, the sequential organisation of the sequence in which the impersonation is embedded follows the structure outlined below:

FPP₁
 FPP₂: Joke₁ → figure-ground reversal, salience imbalance
 SPP₂: @
 FPP₃: Joke₂ → impersonation integrates Joke₁ with non-relevant aspects of GL's role
 as guest panellist
 SPP₃: @
 SPP₁

RG's first quip (line 04) constitutes a case of figure-ground reversal, which is a commonly utilised mechanism in humour (cf. Brône, 2010; Veale, Feyaerts, & Brône, 2006). More precisely, the joke's success is based on a salience imbalance (Giora, 1997, 1999, 2003): the context of GL's utterances profiles the salient meaning of the band names only. The numbers are part of said band names, and their individual meanings are therefore suppressed in favour of the conventionalised band names which are stored – and in FPP₁ oriented to – as a whole. Yet by semantically analysing, or rather "hyper-understanding" (cf. Veale et al., 2006), the lexical meanings of the band names given by GL and profiling the numerical dimension they both share, RG redirects the focus of attention, and activates their 'core' lexical meaning (something which had thus far been suppressed). As a result, on the

pragmatic level RG ascribes an anomalous strategy to GL: rather than correctly guessing the band name based on shared knowledge of pop music, she ‘lists’ numbers. Figure 15 illustrates the Current Discourse Space (CDS) for these turns.

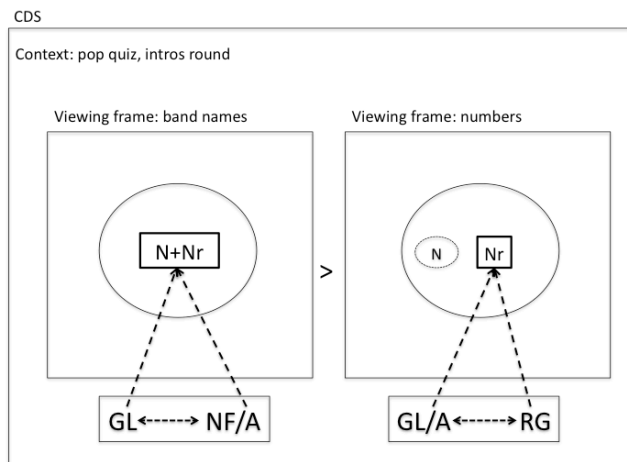


Figure 15: CDS for the Sports Commentator, FPP₁ - SPP₂

The box in the focus of the viewing frame represents the profiled meaning. In FPP₁, the band names are presented as conceptual entities. The lexical item and number are a semantic whole, and designate one concept (represented in the visualisation in the bold box), namely the bands Four Tops and Jackson 5. This meaning is negotiated between GL and her primary addressees, i.e. her fellow panellists with team captain NF. This addressee relation becomes evident in the visual grammar of the sequence, which presents images of the panellists whenever GL pauses during her turn. Her construal is not challenged by the panellists, and even RG returns to GL's construal after the joking sequences (in SPP₁), thus confirming the greater salience of the original construal which did not semantically analyse the individual meanings of the lexical units contributing to the utterance.

Note that there is no integration of elements and spaces present as yet. This first joke relies solely on information already available within the CDS itself, and no new elements need to be inferred. It is only with PJ's contributions (line 06ff) that we find more complex, across-space mappings (or rather: elements). PJ accepts RG's construal in his FPP₃ (line 06 in the transcript), and builds on the profiling of the numbers. However, his quip is more intricate, as it changes the activity type by analogy from a pop music quiz to the announcement of sports results and, in consequence, profiles GL's professional (GL_p) rather than 'celebrity' persona. The incongruity resulting from this clever alteration in activity type, which contrasts GL's

alleged views about the current activity type with the salient ‘actual’ (i.e. interactionally negotiated) activity type, might, in addition to the pleasure received in recognising the reference to GL’s professional persona, lead to a humorous effect. The incongruity projected in the impersonations (i.e. supposedly presenting sports results in a comedy panel game show) is resolved through recognition of her ‘other’ persona, thereby leading to sudden congruity. Figure 16 on the left illustrates the CDS for this first part of FPP₃, which constitutes the *preface*⁶⁰ for the impersonation upon which the impersonation in FPP₃₋₂ (lines 08-09 in the transcript) is built and which is represented in the figure 16 on the right.

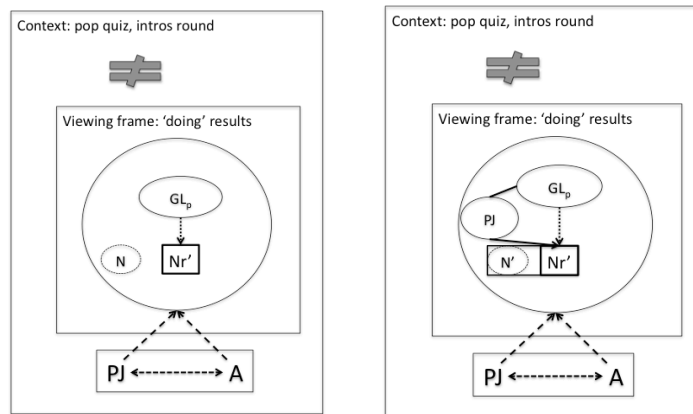


Figure 16: CDS for FPP₃₍₁₋₂₎

The impersonation preface which PJ sets up in line 06 consequently profiles the elements and activity types relevant for processing his reframing of GL’s original utterance. Instead of guesses in a pop music quiz, the listing of band names is reframed as the announcement of sports results in the impersonation that ensues. Element N of the band names is profiled once again, but is treated independently rather than as a constituent of the fixed form-meaning pairings originally profiled by GL in FPP₁. The band Four Tops consequently becomes “tops four” in the emergent

⁶⁰ The notion stems from research on conversational narratives, and describes the means by which storytellers manage participant alignment in story telling. In her study on storytelling in conversation, Jefferson (1978:219) found that stories in interaction may be preceded by a “story preface with which a teller projects a forthcoming story”. Such prefaces can affect the reception of the story by the other participants due to their promotion of a certain stance. In a similar manner to story prefaces, impersonations can be preceded by such prefaces that not only project an impersonation sequence and consequently reserve an extended number of turns-at-talk, but also create a sense of distance (cf. Gülich & Quasthoff, 1986) between the Current Discourse Space and the situation the narrative is set in. In short, prefaces provide the central elements (such as characters, the setting(s) and event structure) of the impersonations’ narrative world which participants in interaction require in order to process impersonations (cf. Sacks’ (1974) *joke prefaces*).

Sacks (1972, 1974) also discusses *pre-tellings* in his treatment of conversational narratives and joking. As this term explicitly emphasises the diegetic (telling) dimension of a narrative, I will refrain from using the concept, but instead adopt Jefferson’s (1978) slightly broader term, which allows for a greater tolerance for the mimetic dimension of impersonations.

structure of the impersonation; this echoes the presentation of sports results, in which the team name is stated before the results. GL's original $(Nr+N)_B$ and $(N+Nr)_B$ becomes N_T+Nr_R : the N component of the band names becomes a team name while the Nr component becomes a result. Figure 17 provides a detailed visualisation of the process.

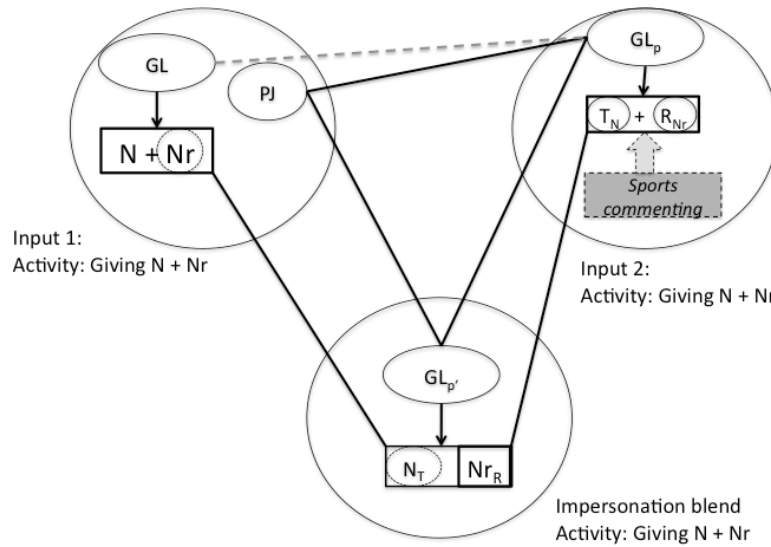


Figure 17: Detailed visualisation of “The Sports Commentator”

All inputs and elements contributing to the impersonation blend (PJ, GL_p and the generic activity of uttering $N+Nr$) are already present in the discourse scenario. Despite the *form* of the utterance being identical in all inputs (uttering generic $N+Nr$), the impersonation performance is of a marginally different activity *type*, as it constitutes the announcement of sports results. As such, the performative dimension differs, yet the form and event structure (listing numerous $N+Nr$) remains the same.⁶¹ In the blend, GL is consequently assigned an incongruent performance embedded in and contrasted with the normative activity of guessing during the intros round. The impersonation sequence (including the preface) thus exhibits a different pragmatic sub-structure, as it constitutes a double-voiced reworking of a previous utterance; this is to say that it attributes a different communicative goal to the original utterance. Yet this other activity type attributed to GL's FPP is presented as occurring *within* the current speech event (as is evident from PJ's “She thinks she’s doing” in line 06), the

⁶¹ Cf. Fauconnier & Turner's (2002) analysis of the regatta blend, which is their prototypical example of a mirror blend. Both yachts sail the same course, i.e. follow the same ‘form’ and event structure, yet their differing contexts and goals are not focused on in the linguistic realisations the blend gives rise to.

pop music quiz, and not another ‘world’. This is therefore contrasted with the normative activities associated with the current game in the current discourse. GL is construed as mixing up her ‘hats’, so to speak, habitually presenting (team) names in combination with numbers which, in the emergent structure of the blend, become results. The impersonation consequently profiles an aspect of her persona that has thus far not been oriented to, and attributes GL’s alleged views to what she is doing, but does *not* provide visits to a clearly distinct discourse scenario, such as a football match. The impersonation purely reworks elements that are already present in the current discourse scenario and that are subject to the same meta-frame structure of the current normative activity. Throughout the impersonation, all participants remain oriented to the relevant activity, namely guessing the correct band name and song title; in short, they all focus on the same discourse scenario (cf. the configuration of the CDS in figure 16). Furthermore, the form and event structure remain the same in all spaces. “The Sports Commentator” is hence a mirror impersonation in our blending typology.

4.1.2. Signalling the blend: Changing voices

It has been established above that this sequence requires complex mappings that lead to the high point of the two joking sequences, the impersonation. Although the impersonation heavily builds on the first joke (FPP₂), PJ’s impersonation preface in line 06 provides the central elements of the performance by outlining its protagonist and central activity (“doing the results”). Note the use of the progressive in the preface, which effectively ‘zooms in’ on the activity and puts the event structure into focus.

After the preface, PJ leaves space for audience appreciation signals (line 07), which signal the audience’s willingness to align themselves as the recipients of the impersonation (cf. Norrick, 2010a). During the laughter sequence, PJ breaks off eye contact with his co-participants, and delivers the impersonation sequence in a marked intonation pattern that accentuates the numbers. As PJ established the activity of ‘doing results’ in the preface, this intonation pattern can be taken to mirror the announcement of results by sports presenters, whose accentuation of the numbers reflect the importance placed on results. His orientation away from his co-participants during the impersonation is further accentuated by the visual grammar of the

interaction-as-broadcasted, which shows PJ together with his co-panellists during the preface, but in close-up focuses solely on him turning his head to his right at the onset of the voice change. Once his short performance is over, we are presented with an image of a laughing GL, thereby further accentuating the fact that she was the ‘butt’ of the joke, the highlight of which was the impersonation sequence.

The Sports Commentator is a brief mirror-blend impersonation that illustrates the complexities involved in setting up impersonation blends. Jointly negotiating blends involves the methodological step-by-step introduction of relevant epistemic domains. Such blends are never unmotivated, but result from, and are normatively relevant to, turn-by-turn talk and joint action. As such, this sequence clearly demonstrates that (shared assumption on) joint action both enables and constrains the meaning which emerges from blending.

4.2. The Proprietor: A Single-Scope Impersonation Blend

In single-scope impersonation blends, the discourse scenarios of the input spaces are clearly distinct from each other, yet only one space (i.e. that of the impersonated) provides the organising structure of the impersonation blend. In short, such single-scope impersonation blends facilitate visits to a discourse scenario that is construed as alien and distinct to the current discourse setting. Mirror impersonations, on the other hand, reframe actual previous discourse, and do not ‘switch’ discourse worlds. Other than double-scope impersonations, however, single-scope blends show no frame clashes with regard to the semiotic code used in meaning-making processes; this is due to the fact that a fictive or non-fictive human from a similar cultural and linguistics sphere is impersonated. “The Proprietor” discussed below is an excellent example of such a single-scope impersonation blend, as it clearly demonstrates the systematic introduction and careful negotiation of visits to other semantic ‘worlds’ that conceptual blending in interaction can give rise to. The impersonation occurs during the identity parade game, which involves identifying a former pop star from a line-up of five people.

NMTB SE19E06_3 04:15 (The Proprietor)

01 BB: i think number five is



02 the proprietor of a reMote (.)

03 GUEst house;



04 AU: @@@ (4.4)



05 BB: <@ <<len, nas> ~YE:s, (.)

06 there is ~r[oo::m in the east wing,



07 JJ: [hehe

08 BB: ↑ahu[huhu

09 AU: [@]@@ (4.1)



10 BB: would you like me



11 to run you a BATH,

12 AU: @@x@ (4.6)



13 JJ: wha [what what



14 BB: [breakfast is served
15 [between seven and seven o three;;



16 JJ: [ʔʰh
17 AU: @@@ (3.0)



18 BB: <<ff> ON MY (.) BAcK.> > (2.0)



19 <<ff> ↑hu[ohoha> @>
20 AU: [@@@



4.2.1. Belief space and narrative space in impersonation prefaces

As the task at hand is to identify a 1970s pop star from the members of the line-up, the impersonation needs to be relevant to previous discourse. To this end, BB sets up, in MSCi terms, a belief space (“I think”, line 01), in which No. 5 is the proprietor of a remote guesthouse. This belief space is set up relative to the ‘base space’, or the CDS in our terminology. Via the access principle (or ID principle), the already familiar

element *a* in the Current Discourse Space is connected to element *a*₁ in the Belief Space, as visualised in figure 18.

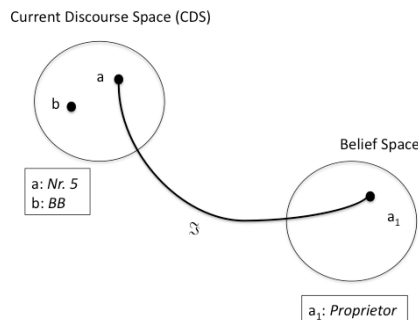


Figure 18: Belief Space set up in the preface to ‘The Proprietor’

The Belief Space is interactionally relevant to current discourse, as it first of all utilises an element of that discourse as its trigger. Secondly, engaging in guessing the right candidate is relevant to the activity currently underway. Furthermore, attributing this identity to No. 5 in the line-up is acceptable, as the comedy genre in which this panel quiz is set leads to greater acceptance of non-topical talk. In this impersonation preface (lines 01-03), BB not only attributes a novel identity to No. 5, he also frames the story he is about to introduce by providing basic information as to the setting and the protagonist. In short, BB establishes a *narrative space* (cf. Dancygier, 2008). He provides cues for the audience to further infer the details of this narrative world: the setting of his impersonation is, as BB explicitly remarks, a “remote (.) guest house”. The emphasis of the adjective ‘remote’ might be employed in order to draw a link between the cultural saliency of remoteness and scenes of horror and deviancy (cf. Davies, 1998; Davies, 2005). Remote hotels, motels and inns are, after all, a common setting for horror stories in Anglo-American culture, as is evident when considering Hitchcock’s *Psycho*, Mangold’s *Identity*, the splatter film *Hostel*, and Stephen King’s *1408*. In *Psycho* in particular, the initially innocent-looking proprietor of a motel is actually revealed to be a ruthless murderer. It is this widely shared knowledge, culturally speaking, which may be capitalised upon in BB’s preface. The preface is followed by audience laughter, a token of appreciation that signals understanding and approval, and licenses BB to further illustrate the idea he introduced and therefore the internal frame structure of the belief space.

The preface, and thus the belief space it sets up, yields a number of inferences

that float back to the Base Space (the Current Discourse Space) and alter it. No. 5 is attributed the identity of a guesthouse proprietor via the identity connector \mathfrak{S} . He is, furthermore, probably slightly deviant in one way or another, as, due to its marked intonation pattern, the adjective ‘remote’ has to be regarded as an ostensive stimulus that is intended to activate culturally shared frames of deviancy and horror.

4.2.2. From preface to performance

Although BB’s preface is acknowledged by the audience and his co-participants through appreciation tokens in the form of laughter, the belief space is not fully abandoned after the preface. Instead, BB reserves more turns-at-talk to further illustrate and explore the internal structure of the story space introduced in the preface. Interactionally, this is firstly signalled by BB’s slightly rising intonation pattern at the end of the Turn Constructional Unit (TCU)⁶², thereby projecting that he has not yet finished his turn-at-talk. The other participants do not claim the floor, thus showing that they accept the preface as leading to a further elaboration of the narrative world introduced. The belief space consequently becomes a narrative space, furnished with expectations as to the further interactional unfolding of the sequence (e.g. reservation of extended sequence of turns-at-talk).

Secondly, during the four seconds he leaves for the audience to deliver their appreciation signals BB prepares an impersonation blend by employing embodied disjunct makers that signal he is breaking with topical discourse. Before starting to speak as “The Proprietor”, he alters his body position and direction of gaze. He hunches his shoulders and stares straight ahead, away from his fellow panellists and with eyes wide open. Using these ostensive stimuli, BB shows that he is now oriented to a discourse scenario which is different to the CDS. As the proprietor role is highly salient in the personal common ground between BB and his audience, this character might be recruited from the very beginning of the bodily performance via the drama connector \mathfrak{D} (cf. Fauconnier, 1997); at the latest, this will occur in line five, when BB as “The Proprietor” talks of rooms available and the interaction-as-broadcasted presents an image of No. 5 (see line 09 in the transcript). BB as the proprietor is contrasted with an image of No. 5 at the end of each segment of the impersonation,

⁶² See chapter 3.2.2. for an overview of the turn-taking system. Chapter 5.3. will further focus on floor management in impersonations.

thereby using yet another semiotic channel to indicate that the mapping is still active. BB and the other participants in the interaction-as-broadcasted thus carefully introduce the blend and the mappings it requires, drawing on a number of semiotic resources.

In the four recycles of a transition place (lines 05-08, 10-11, 14-15, and 18-19) that follow (cf. Jefferson, 1978), BB further elaborates on the internal structure of the narrative space by playing on culturally salient horror scenes, such as the bath (cf. Hitchcock's *Psycho*), and deviant, incongruous behaviour, such as providing just three minutes in which to eat breakfast and which is served on the proprietor's back. Despite the fact that BB still 'inhabits' the body of The Proprietor, JJ attempts to claim the floor in line 13; however, he is interrupted by BB's continuation of his story. As JJ's attempt occurs after a rather long (5 seconds) space left by BB for signals of appreciation, this may constitute evidence that such pauses should be subject to a maximum length of four seconds (cf. chapter 5.3.).

The impersonation reaches its climax through prosodic markers such as micropauses after each word in line 18, and ends with highly dramatised post-utterance completion laughter that is characterised as abnormal and maniacal. The two final elements of the impersonation, the accentuated "on my back" and the 'maniacal' laughter, are thus produced in punch line mode. Norrick (2010a :75) has shown that the "recipients may also laugh at what they take to be the punch line of a joke". For this reason, a line delivered using the features associated with punch lines (increased loudness, words accentuated individually) is treated as the punch line that ends the impersonation, even though semantic features, i.e. the creation of incongruity (Kotthoff, 2006; Oring, 1989), are also met with other elements of the performance. However, Jefferson (1978) has found that for conversational storytelling, punch lines often summarise the central elements of a narrative. In this respect, these last two lines of the impersonation constitute a clear recapitulation of the madness theme that had already been alluded to in the preface via the explicit mentioning of "remote". All the other segments that contribute to the impersonation (lines 06-15 in the transcript) accentuate the madness-theme less overtly than these final two turns. Their prosody and semantic content are presented as far more deviant than in the previous recycles, and consequently echoes the structure projected in the preface far more effectively. This first of all illustrates that Jefferson's (1978) findings of punch lines "condensing" narratives might also apply for impersonations. Secondly, the sequence also

demonstrates that delivering punch lines may require a prosodic ‘mode’ that distinguishes the punch line from previous elements. Furthermore, such punch lines may constitute an exit device, signalling that the impersonation sequence has reached its climax and is now complete. This becomes evident when the participants in the interaction-as-broadcasted ‘leave’ the narrative space after the climactic laughter sequence in line 19, and return to topical talk in the CDS, which is introduced with an “anyway” by a fellow panellist.

4.2.3. Frame structuring in The Proprietor

The frame structure of the blended performance is provided solely by the narrative space, and is induced via pattern completion. BB exclusively supplies high points of a story that has to be co-constructed by the audience on the basis of shared common ground in the narrative space. The (potential) event structure surrounding the ‘best-of’ utterances presented in the performance, i.e. the script (cf. Schank & Abelson, 1977), and other elements of the internal space structure, such as the addressee(s) of the contributions, have to be inferred by virtue of frame induction. The narrative structure of the impersonation is consequently exclusively structured by the story space set up by BB in the preface. There are no major frame clashes regarding the semiotic code, which leads to the blend being single-scope in our typology.

However, the way in which BB chunks his contributions and the visual grammar (Holland, 2000) presented in the interaction-as-broadcasted illustrate that despite presenting a single-scope visit to another ‘world’, he remains primarily oriented to the discourse base space, interactionally speaking. Leaving enough space for appreciation tokens on the part of the audience after every TCU and concluding it with post-utterance completion laughter demonstrate this point. The images shown in the interaction-as-broadcasted during the spaces BB leaves for audience appreciation signals are medium shots of BB’s entire panel and close-ups of the ‘butt’ of the impersonation, No. 5. As a result, the visual grammar of the interaction-as-broadcasted further illustrates that the discourse base space is oriented to at these moments. During BB’s impersonation sequences, however, BB is shot in close-up, thus providing evidence that the narrative space is the only element in focus. Even so, both BB and the visual grammar of the interaction-as-broadcasted chunk the performance such as to continuously orient back to the discourse base space that contains topical

turn-by-turn talk. Although the narrative is entirely structured by the story space set up in the preface, it is therefore the pragmatic dimension that ultimately determines much of the performance of the narrative world presented.

Continuing in this vein, the impersonation has emergent pragmatic properties that are relevant to, and project back into, the discourse base space: No. 5 is ascribed deviant behaviour and constructed as a laughable. It also allows inferences about BB's qualities as a comedian, as well as his culturally shared values, knowledge and themes. In brief, all inferences are created from the perspective of the base, i.e. the space containing topical turn-by-turn talk. Figure 19 illustrates the temporal unfolding of the conceptual mapping process negotiated in the impersonation blend.

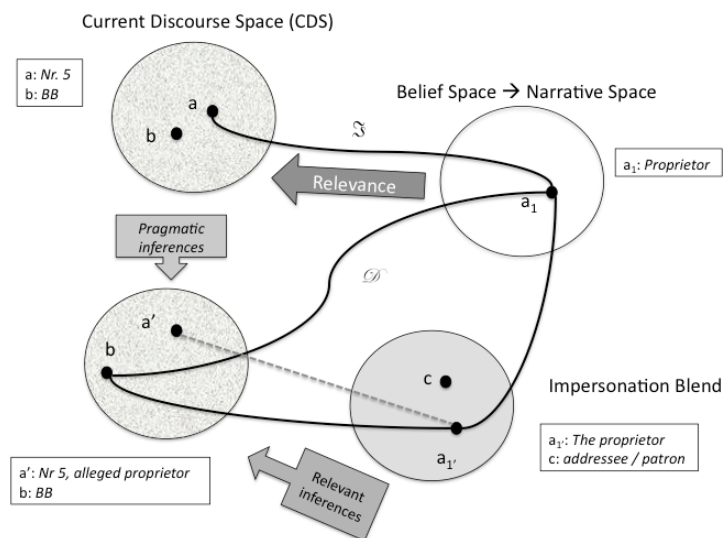


Figure 19: "The Proprietor" impersonation blend

It therefore becomes evident that the interactional, pragmatic dimension exceeds considerable power over both the 'selection' of frames that concepts can be blended with in impersonations and the delivery of the impersonation and inferences arising from this.

4.3. The Metal Detector: A Double-Scope Impersonation Blend

Double-scope impersonations are characterised by frame clashes with regard to both discourse spaces and semiotic codes. Impersonators perform a non-human character, subsequently rely on other signs for communication, and inhabit a world of that is clearly distinct from the current discourse scenario. The Metal Detector sequence discussed in this chapter is an excellent example of such an impersonation. The team members seated next to team captain NF are asked to guess which pop star designed their own brand of metal detector.

NMTB SE21E04_1_09:34 (The Metal Detector)

01 NF: are metal detectors made out of
metal,



02 cos that's always confused me.



03 AU: @ (1.0)

04 SA: good point=

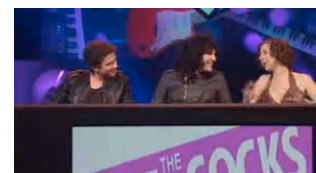


05 NF: they just detect themselves



06 KS: right they just beep

07 [they're CO:nstantly beeping=



08 AU: [@@@

09 NF: <@ ↓meep (.) me again (.)

10 AU: @@



11 NF: ↓meep (.)



12 ah that was me (.)

13 so:rry (-)

14 AU: @

15 NF: ↓meep

16 i can't HElp it, (.)



17 I'm MAI:nly MEtal. @>



18 AU: @@@@



19 NF: could we have a look at the clip.



4.3.1. Interactional organisation: Negotiating viewing frames

This sequence is not the first reaction to the FPP posed by host SA, i.e. the question regarding which pop star has their own brand of metal detector. NF's team has already provided an answer "expressed in dance" (further evidence for the joke-first strategy central to NMTB), which is considered "wrong" by SA. The sequence under scrutiny here is therefore the second diversion from providing the 'correct' answer and hence the projected SPP. The interactional organisation of the sequences is as follows:

FPP₁: Pop star – metal detector
 SPP₁₋₁: Dance
 SPP₁₋₂ (Assessment): Wrong
 FPP₂: Materiality (warranted)
 Acknowledgment token: Good question
 Further elaboration of FPP₂, including impersonation
 SPP₂: @@
 FPP₃: Clip

After SA deems the team's first answer as "wrong", NF moves away from the topical talk, and focuses instead on the materiality of the metal detector. This change in focus has to be accounted for (cf. Sacks, 1992), and is warranted by giving a reason for the diversion (line 02 in the transcript "cos that's always confused me"). If we consider the changes the CDS undergoes in this sequence, we find the following: FPP₂ now foregrounds the element Metal Detector (MD) in the viewing frame, whilst backgrounding elements P (pop star) and D (activity 'to design'). These elements are not relevant to the action proposed by FPP₂. Figure 20 illustrates the change in focus.

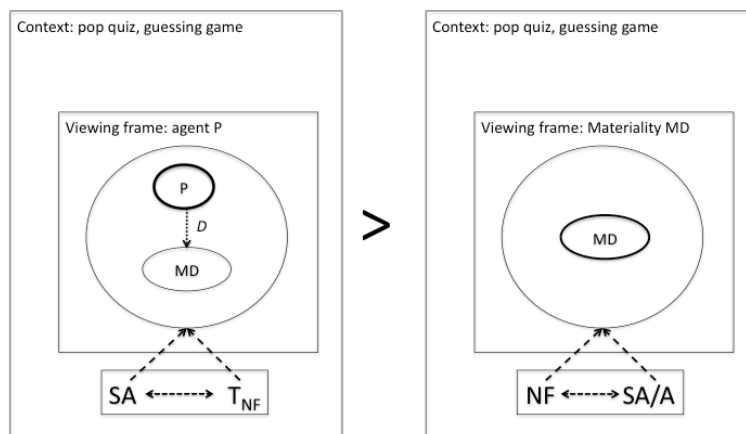


Figure 20: Changing viewing frames in "The Metal Detector" (lines 01-04)

4.3.2. 'Zooming in' on a narrative

SA, who is presented as the prime addressee for this question, provides an assessment and acknowledgment token ("good point"), which, together with the audience appreciation token in the form of laughter, licenses NF to further elaborate on the problem he has just presented. Both NF and his fellow panellist KS each provide one turn in which they expand on the central aspects of the problem posed in FPP₂, with NF providing the activity ("they just detect themselves", line 05) and KS offering the sonic dimension ("beep", lines 06-07) of the dilemma. In giving these further details,

they provide the event structure and protagonists, as well as the central conflict of the narrative. Rather than a factual discussion concerning the materiality of metal detectors, it is KS's contribution in particular which advances the 'discussion' to a more detailed and sensual level: in using progressive aspect (line 07, "they're constantly beeping"), she 'zooms in' even further on the conflict, presenting it as an ongoing process and from the "inside" (cf. Yule, 1998:64ff). As such, by building on the viewing frame proposed by NF, KS zeroes in on the narrative established. The viewing frame once again shifts, with the 'world' of an incessantly beeping, self-detecting metal detector coming into the spotlight. As figure 21 illustrates, this narrative space is now oriented to, and provides the ground for NF's impersonation. The central character, activity and conflict are already in focus, so NF starts his impersonation with the salient sonic feature introduced by KS during her previous turn, namely a rendering of the beep sound made by metal detectors when detecting metal. From the outset of this impersonation, NF shifts his gaze away from his co-participants (a detailed analysis of gaze and postural orientation of this sequence is offered in chapter 5.2.2.), signalling that he is not oriented to the discourse world inhabited by his fellow participants; instead, he now 'lives' in the narrative space introduced over the course of the previous turns. Employing a discursively dissonant sign, the "meep", is a further means of signalling the role change. The sonic dimension of the metal detector's world can be taken as salient in the common ground due to KS' prior description of the detector's activity. After a micro pause NF adds "me again", returning to a familiar code; nevertheless this utterance is by now firmly attributed to the metal detector. Figure 21 visualises the transitions the CDS undergoes over the course of these two turns.

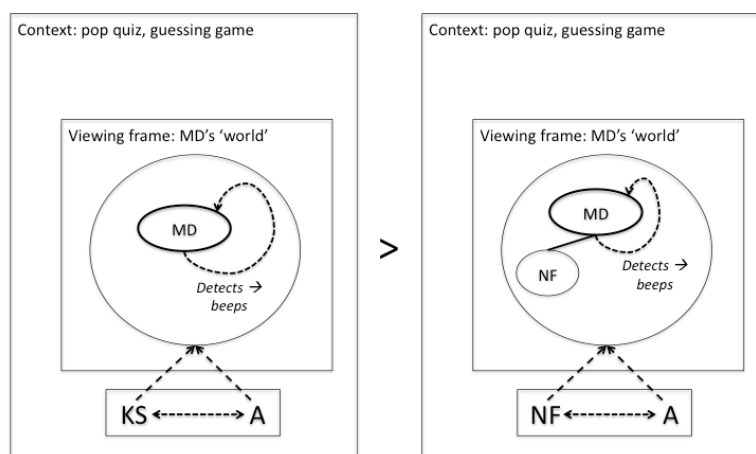


Figure 21: From narrative to impersonation (lines 06-07)

Over the next few recycles of a Transition Relevant Place (TRP), each of which are followed by appreciation tokens in the form of laughter, NF expands upon the conflict faced by the metal detector, ending with a climactic and highly dramatised explanation of the crisis it encounters; this can be considered the punch line of the impersonation (lines 16-17). The explanation offered in these lines is a repetition of the component(s) already introduced in the preface to the impersonation and, by repeating this, “condenses” (Jefferson, 1978:223) the narrative in the punch line once again. The impersonation comes ‘full-circle’, so to speak, and its contingency with previous talk is once more displayed. Once the audience has delivered their final appreciation of the impersonation sequence, NF formally returns to topical talk, asking to replay the clip which introduced the original quiz question posed by SA in FPP₁.

As is the case for the other impersonation sequences discussed in the previous sections, the metal detector impersonation seems irrelevant to topical talk, as it is not sequentially implicative; rather, it is “deleted” (Jefferson, 1978:229). Although introduced as relevant to prior topical discourse, the impersonation, it appears, is treated as a unit that is not germane to further turn-by-turn talk. Impersonations are for interactional play and entertainment, but are not sequentially implicative for ensuing talk. In other words, while an impersonation sequence projects a laughter sequence as a relevant next action, it is usually not followed by further related commentaries, as is commonly the case with stories.

4.3.3. Frame-structuring in double-scope impersonations

When we investigate the impersonation sequence in more detail, it becomes evident that different processes of frame integration are at work than the processes in the previous two examples. First of all, the impersonation is set in a discourse scenario marked as clearly distinct from the one in which the original interaction (FPP₁) is situated. Similar to single-scope impersonation blends, the impersonation allows for visits into the established narrative worlds. Other than mirror-scope blends, however, these impersonations are not accessed from the Current Discourse Space and the topical prior talk it ‘contains’; instead, they provide mimetic illustrations of narratives from within the narrative space. In focus is the (fictive) narrative world rather than the reframing of (actual) previous discourse scenarios.

Single-scope impersonation blends like “The Proprietor” differ, however, from more double-scope blends such as “The Metal Detector”. While “The Proprietor” relies on human semiotic codes in both input spaces, “The Metal Detector” faces a frame clash regarding not only the signs used in communication, but also human intentions and feelings attributed to the machine. The impersonation consequently encounters frame clashes both with regard to the discourse scenario and semiotic systems. The emergent structure of the blend for this reason presents a human-like machine which not only masters its ‘own’ semiotic codes (the ‘meep’, which is purely a human rendering of a non-human sound), but also those of humans (complete with the politeness system). Furthermore, in addition to being able to feel frustration over its materiality, it can also express this frustration in an apologetic manner.⁶³ Moreover, an addressee A needs to be inferred in the blended space via pattern completion.

Figure 22 provides a detailed illustration of the blending process behind this impersonation, and shows the double-frame clashes which the Metal Detector impersonation contains (see also figure 21, which shows the discursive unfolding of the impersonation).

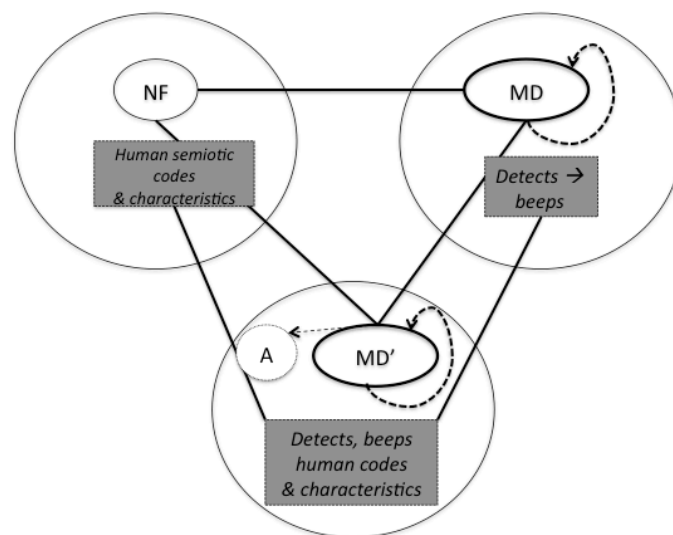


Figure 22: The Metal Detector impersonation blend

While the impersonation is accessed from the narrative world of the metal detector alone, it contains NF’s voice in an unmarked code (except for the “meep”), and

⁶³ Cf. other human-like machines in popular culture that master politeness, such as *Star Wars*’ C-3PO, which blend machines with human pragmatics.

consequently borrows structure from both input spaces. It is thus double-scope in nature.

4.4. The Way We Think, The Way We Act

The case studies provide hypotheses as to the way in which discursive conceptual blending is coordinated in interactions. Impersonation blends have to be relevant to previous actions (i.e. their relation to previous topical talk has to be displayed); they also have to be methodologically introduced, entered and exited. The relevance to joint action both constrains and licenses impersonation blends, and furthermore guides conceptual mappings, and, consequently, meaning-making. From a methodological point of view, this relevance is made transparent in interaction, and a prototypical impersonation structure can be observed. *Prefaces* relate narrative worlds to previous interaction, and ‘zoom in’ on the narrative space in which an impersonation is set. Aside from this, they project impersonation structure and evaluative stance, and provide the central frame elements, such as protagonists and event structure. Conversely, *Demarcation markers*, such as breaking off eye contact and postural orientation, signal actual voice changes, and make clear that all actions are now to be understood from within the narrative space rather than from the discourse space containing topical turn-by-turn talk. The *performance design* of impersonations is such as to secure the floor for an extended sequence of turns-at-talk whilst providing the audience with sufficient opportunity to signal appreciation and continuity. *Exit devices* such as TCUs delivered in punch line mode are employed to mark the end of a longer impersonation sequence, and ensure that the impersonation relates to the preface.

It appears, then, that impersonating another voice is first and foremost a social activity that needs to be carefully organised and negotiated. ‘Doing’ an impersonation is a social activity, an activity that is carried out for and with someone and which has a social purpose. For this reason, such blends cannot be delivered ‘wildly’; rather, novel concepts introduced need to be relevant to the action underway, they must be carefully introduced for the interactants to comprehend the pertinence of uniting these concepts, and they must be delivered in orientation to the action at hand. As a

consequence, it is the performative, interactional dimension that appears to structure the delivery of the impersonation, as well as making the impersonation blend relevant and legitimising it, i.e. visiting ‘other’ worlds and integrating semantically distinct concepts.

The joint, rule-dominated quest for shared meaning construction evident in interactional data consequently suggests that orientation to joint action guides conceptualisation processes and therefore meaning-making to a considerable degree; this is a factor which is often neglected in more canonical Cognitive Linguistics research. To further discuss this hypothesis, the following section(s) will describe in detail the interactional methods employed in organising impersonation blends of all three types and, in a second step, will discuss the extent to which the “level of performance” (Kotthoff, 2006) governs meaning construction in conceptual blending processes.

CHAPTER 5: INTERACTIONAL ORGANISATION OF IMPERSONATION BLENDS

Building on the results gathered and hypotheses generated in chapter 4, this section systematically discusses the interactional organisation of impersonation sequences and provides a typology of the methods employed to set up impersonations for each blending type. By carrying out a detailed analysis of the techniques used to negotiate the introduction, performance and ending of such impersonation sequences, the issue regarding whether interactants show an orientation to frame integration complexity in the joint quest for meaning-making is addressed. In other words, the analysis will discuss the question as to whether the various blending types proposed in MSCI actually reflect distinct discursive practices.

I hope to first of all trace these blending practices in the local occasioning of impersonation blends. To this end, chapter 5.1. describes the strategies employed by interactants to introduce impersonations of varying frame-integration complexity. How are the various narrative worlds introduced in which the impersonations are set? How is the tellability of impersonations warranted in interaction? The successful parsing of the voices present in impersonation sequences is of key importance to the performance. For this reason, chapter 5.2. focuses on the methods employed to distinguish between the voices speaking and acting in impersonations. This central chapter deals with both linguistic and bodily semiotics as a means of further discussing the role played by quotatives in setting apart voices, and observing the extent to which discontinuation of gaze and postural orientation might be utilised to signal semantic distance. Chapter 5.3. will then concentrate on the linguistic performance of impersonations. Are there varying strategies when it comes to floor management? How are turns-at-talk negotiated? Are there differences as to the chunking of the impersonations? In the final step of the analysis, I will describe the strategies used for ending such performances and return to topical turn-by-turn talk in chapter 5.4.. Are some impersonations sequentially deleted and others sequentially implicative? Do certain voices require more interactional work to be exited than others?

The methods described therefore work together to construct discursive practice, meaning that they do not themselves constitute actions. Thus, while the

mode of presentation might suggest that these methods are already meaningful when viewed in isolation, I would like to emphasise the fact that it is only the finely tuned orchestration of these methods that builds actions. For this reason, the last section, chapter 5.5., provides an overview of the methods identified and the actions which they serve to develop, and furthermore discusses their relation to frame integration complexity.

Each chapter will provide initial conclusions and consider the relevance which the findings obtained have on determining the extent to which conceptual blending complexity is reflected in discursive practices. The implications of my results are discussed in detail in chapter 6, which – based on the outcome of my analysis – proposes a refined model of meaning-making through conceptual blending.

5.1. Local occasioning and diegetic prefaces

In his *Republic* (Πολιτεία), Plato differentiates between *diegetic* and *mimetic* narration. While the first form refers to a narration in which poets do not explicitly hide their voices, but speak as ‘themselves’, mimetic narration refers to characters speaking in their own voice in the narration. This division has since been described as *telling* versus *showing* and *simple* versus *scenic presentation* (Friedman, 1975).

As the case studies in the previous section have shown, voice changes in impersonations are often dramatic illustrations of just such diegetically established storyworlds, providing the “narrative macro-structure” (Solomon, 2004:261), or an ‘abstract’ (Labov, 1972a; Norrick, 2000) of the mimetic impersonation. The case studies furthermore suggest that double-scope blends may require a more elaborate prefacing structure than single-scope and mirror blends; while “The Proprietor” and “Sports Commentator” are prefaced by just a single TCU uttered by the impersonator himself, “Metal Detector” shows an elaborate preface which is co-constructed over the course of a number of turns by two participants. Based on this observation, this section will compare preface structures across blending types to find further evidence which will support the hypothesis. While such prefaces already play an important role in securing an extended sequence of turns-at-talk, floor

management will be discussed in further detail in chapter 5.3.; this chapter will pay close attention to the function of prosody in turn organisation and discuss in detail the prosodic format of prefaces.

5.1.1. Local occasioning and preface formats in mirror impersonations

Mirror impersonations attribute identities, affective stance or tacit motives to co-participants. In the NMTB corpus, six formats for introducing mirror prefaces can be identified, with each serving to achieve marginally different pragmatic effects.

First of all, impersonators may offer their assessment of a prior speaker's action before restaging that action in an accentuated form. For this type, voice change is achieved with the help of the quotative *be like* (cf. chapter 5.2.1.). In formalised terms, prefaces follow the format

Assessment S – *be like* [IMP_{marked repeat}]

This format occurs in scenarios in which the impersonation illustrates the affective stance diegetically displayed by, and hyperbolically attributed to, a co-participant. Consider the following sequence:

NMTB SE24E11_1_12:00 (Bemused look)

01 NF: i've NEver seen anyone
02 look more bemused;
03 you're like-



The highly exaggerated re-enactment of a co-participant's affective stance in “Bemused look” ‘cites’ an element of the talk which precedes the impersonation sequence. It thus shows that the sequence, while not being directly relevant to topical talk per se, is occasioned by that talk and therefore results from monitoring it. The Assessment S component of this format (lines 01-02 in the transcript) provides the diegetic evaluation of the source (which in this case is ironic). As such, it frames the ensuing impersonation, providing cues as to how it is to be understood. The second

component (line 03), on the other hand, serves to introduce the mimetic quote through a quotative frame (cf. chapter 5.2.), while the mimetic impersonation itself is presented in the format referred to by Jefferson (1978) as ‘marked repeat’. The (in this case visual) repetition of a previous action performed by speaker (S) is hyperbolically accentuated in the impersonation. Apart from highlighting S, the ‘butt’ of the impersonation, as a subject of ridicule, this accentuated ‘quoting’ further contributes to warranting the impersonation’s tellability by showing its relation to topical talk, which is its source. The superlative in the preface (Assessment S component, lines 01-02) likewise points to the reason why the impersonation is deemed narratable and “newsworthy”, to borrow Norrick’s (2010a:105) term. Such superlatives constitute an effective tool for displaying the impersonation’s importance and consequently asserting the right to perform it.

In a similar vein, impersonators may re-describe a previous action (U), diegetically displaying their own evaluative stance on this prior action, before mimetically restaging the original quote in a manner that supports their assessment of the original token. This second format can be formally described as

Redescription_U [IMP_{marked repeat}]

and is illustrated in “Smug Jonas”.

NMTB SE20E05_2_02:37 (Smug Jonas)

```

01 SA:  yes jonas; (.)
02      i would have given you a point, (-)
03      but you were ↑a bit SMUG about it.
04 JA:  [<<ff> i !WAS!n't smug about it >]
05 AU:  [@@@                                     ]xx (2.4)
06 SA:  <@ i know it (0.4)
07      it's northern- (.)
08      the stone roses wi:th, (.)
09      blabla (.) !WRO!ng. @>
10 AU:  @@@ (5.6)
11 SA:  please try and do better next time.
```

The impersonation here serves as a mimetic illustration of the assessment (“smug”) provided in the diegetic redescription of a prior utterance in the preface (line 02-03). As JA challenges SA’s assessment in line 04, the impersonation is also a relevant next

action, providing evidence for the correctness of the assessment. Redescription_U [IMP_{marked repeat}] is thus a format that is recruited when doing impersonations in order to illustrate a purpose. The diegetic preface not only provides details as to the voice speaking and acting in the impersonation sequence, but also sets out the evaluative and affective macro-structure in which the ensuing illustration is to be understood and interpreted. The impersonation itself mirrors the source, but accentuates components that serve to prove the assessment provided in the preface. In “Smug Jonas”, the impersonator assumes a relaxed and self-assured body language throughout the impersonation sequence. The accentuated complacency of the marked repeat of JA’s original contribution is then contrasted with the accentuated !WRO!ng in line 09, which completes SA’s performance, and ties back in with prior topical turn-by-turn talk, as it once more provides the reason for not granting the point requested by JA. The impersonation thus serves to illustrate the macro-structure diegetically outlined in the preface.

Marked repeat such as the instances highlighted in “Bemused look” and “Smug Jonas” may also feature in mirror impersonations preceded by a disjunct marker (DM), such as “Oh”. The disjunct marker (also referred to as *change-of-state token* (Heritage, 1984) or *touch-off marker* (Schegloff, 1996)) signals a break from topical discourse, while the impersonation sequence [IMP] cites an element of the original utterance, but reframes it so as to allow affective stance to emerge. This third type of mirror impersonation format features in “Banana outfit”, and can be formalised as

Disjunct Marker [IMP(marked repeat_E)]

NMTB SE23E07_1_03:06 (Banana outfit)

- 01 MW: DAniel have you ever worn anything for a
 02 photo shoot that you thought, (.)
 03 oh dear. i shouldn't be doing this.
 04 DM: we:ll um no sea horses; i (.)
 05 i've worn a (.) uhm BANana outfit before;
 06 MW: <<acc> OH, (.) (h) <@ NO(h) (.)>
 07 ↑nothing too embarrassing-
 08 just a banana outfit. @> (-)
 09 AU: @[@
 10 DM: [yea.
 11 MW: ↑why?

In this example, the impersonation sequence is diverted away from topical talk via the disjunct marker (DM) “oh” (line 06). The DM signals that the impersonation provides a break with prior discourse, instead offering an accentuated rephrasing of the original contribution in the impersonation. The impersonation’s relevance to topical talk is nonetheless clearly displayed in the impersonation sequence itself. The trigger for the impersonation (lines 07-08) in “Banana outfit” is an element of prior talk, which becomes clear towards the very end, namely in the punch line of the impersonation (line 08), when the token is cited. Jefferson (1978) discusses the device DM + marked repeat in her study into conversational storytelling, and has been able to prove that such repeats may also occur in punch lines only. As in Jefferson’s stories, then, the relevance of the ‘Banana Outfit’ impersonation to topical talk is only warranted towards the end of the sequence. The impersonation preface format Disjunct Marker [IMP(marked repeat_E)] thus occurs in scenarios which reframe the original contribution rather than simply continuing it, as would be the case in zero-preface impersonations (see below). To set this reframing apart from the original contribution, a DM is used. Relevance to topical turn-by-turn talk is signalled through a marked repeat, which illustrates that the reframing has the prior utterance as its source.

The fourth format, too, provides a means of reframing a prior contribution by a co-participant, yet does so more explicitly through the use of a format that may be formalised as

S V-ing X [IMP(marked repeat_E)]

In this format, the impersonator in the preface openly states the activity V-ing X s/he believes speaker S to have engaged in when doing a prior action. This is the case in the “Sports Commentator” discussed chapter in 4.1., but also in the following sequence which occurs in the same episode:

NMTB SE23E02_3_00:52 (Sports Commentator II)

```

01  GL:  number one number five and number three (.)
02      have smiled; (.) quite a lot;
03      number three has stopped now;
04      cos i've said that;
05      but number five and one smile a lot.
06  RG:  GAbY; (.) i know=i know you're a commentator

```

07 but can you ju(h) @
 08 @@ (1.6)
 09 GL: <<f> i'm BEIng a DEtective;>
 10 some[one-
 11 RG: [no you're being a COMmentator,
 12 [<@ number one just smiled
 13 GL: [NO twO's just smiled cos HE thinks-
 14 RG: now number two is smiling
 15 number three is coming with a late smile,
 16 <<acc> there, he's coming up,
 17 number four
 18 number four comes in with a smile
 19 number five with a grimace;
 20 number five number five,> @>
 21 @@

Both “Sports Commentator” sequences are preceded by a preface that outlines the reframing of the original contribution in which the ensuing impersonation is to be understood. The preface provides co-participants with the voice about to speak (S) and the activity which S is understood to be engaged in (V-ing X). Special emphasis is placed on the action V-ing, as the progressive is used, thereby further focusing on the activity at hand and serving to highlight the impersonation’s relation to prior topical talk. This relevance to topical talk is additionally warranted through providing ‘reasons’ for the original speaker’s alleged failure to orient to the normative activity at hand (i.e. providing an answer to the quiz question). This alleged violation of the normative activity being undertaken is made explicit prior to the concise redescription of the original contribution through the S V-ing X [IMP(marked repeatE)] format. In lines 06-07 of “Sports Commentator II”, RG already indicates in a jocular fashion his interpretation of GL’s contribution (lines 01-05) as failing to provide the relevant SPP (i.e. the answer to the question posed in the quiz). The same applies for “Sports Commentator I”, where GL allegedly fails to offer a straightforward answer, and instead only guesses band names. The impersonations therefore demonstrate the purported reasons for the alleged failure of co-participants to adhere to normative standards. Both the alleged failure as well as the reframed redescription of the original activity are already provided in the diegetic preface. The impersonation itself simply provides an accentuated mimetic illustration of the reframing proposed in the turns preceding the performance. In both “Sports Commentator” sequences, GL’s attempt to provide the right answer is partially quoted in the impersonations. However,

impersonators adopt the speech pattern employed by sports commentators, and do not cite the original version word-for-word; rather, they rearrange and recontextualise the lexical items of GL's original contribution in a way that aligns with the assessment which was diegetically provided prior to the impersonation.

Such assessments as sources for impersonations can furthermore take the form of questions, and can be formalised as

Questioning [IMP(marked repeat_E)]

This fifth format is recruited in “Northern lass”:

NMTB S24E08_pt2_01:49 (Northern lass)

```

01 DV: do we have to stand yea?
02 PJ: yes we do=o
03 DV: [great
04 PJ: [diana vickers
05 DV: very excited
06 LM: what about STAndin?
07 DV: [ye(h)es
08 AU: [@@@ (1.9)
09 LM: <@ <<northern accent> i'm a simple northern girl
10      with a simple wishes
11 AU: @[@@
12 LM: [i wish i could go to london and stand, > @>
13 AU: @@[@ (1.6)
14 LM: [<<f> WEll tonight your drEAmS come true.>

```

When DV claims to be “very excited” about standing up (lines 05-07), this constitutes a jointly negotiated figure-ground reversal and, in turn, salience-imbalance (Giora, 2003). While DV's excitement would be deemed legitimate were it related to the upcoming task, LM's question for understanding (line 06) proposes a non-salient reading that assumes a completed prior sequence (lines 01-02). Through his question, LM not only claims to ‘hyperunderstand’ (cf. Brône, 2008) DV's assessment (lines 03 and 05), but also reads into her original contribution as a source of deviant behaviour. Being excited about such a plain activity as standing up is presented as violating norms. DV adopts LM's reading (line 07), and acknowledges through laughter particles her appreciation of the humorous figure-ground reversal.

Questions of the type posed in “Northern lass” therefore constitute a means of proposing hyperunderstanding and figure-ground reversal. The impersonation then illustrates and further accentuates the diegetically established figure-ground reversal. This recontextualisation and rephrasing of the original utterance is methodologically linked to prior topical talk via the question as well as a repeat of a lexical element in the performance itself, which again shows that prior topical talk represents the source of the impersonation and the figure-ground reversal it relies on.⁶⁴

Finally, the last format neither shows preface, disjunct marker nor marked repeat. I refer to this format as zero-preface. Impersonations showing such zero-preface contribute to the further visualisation of thoughts and feelings. Rather than repeating, they expand upon a co-participant’s utterance which occurred immediately prior to the impersonation. As such, impersonations lack a diegetic preface that outlines the macro-structure in which a mimetic illustration is set, and they heavily rely on alternate means of signalling voice-changes and referents, such as postural orientation and bodily mimesis, gaze direction or style switching. In “Frustration”, it seems as though impersonator PJ simply continues the line of thought pursued by the original speaker:

NMTB SE23E08_2_03:00 (Frustration)

01 SB: Oh this is frustrating; (.)
 02 VErY frustrating.
 03 PJ: <@ i had no idea
 04 a quiz would be like that
 05 [i really didn’t @>
 06 AU: [@@@
 07 NF: well you might as well guess.

As the original speaker, SB has already shared his emotional state, and the impersonation continues in this sentiment, meaning that no further diegetic renderings of the type discussed in the sequences above seem to be required. The impersonation does not constitute a break with topical talk, as it neither reframes nor recontextualises the original contribution.

In sum, then, mirror impersonations seem to employ preface formats that allow for the assessment and ensuing redescription of previous utterances. Marked

⁶⁴ As the prefacing formats which other blending types exhibit illustrate, questions in general are a means of drawing attention to aspects which the impersonator finds noteworthy.

repeat of elements of a co-participant's prior action is of central importance in order to signal the impersonation's relevance to topical turn-by-turn talk. However, one format does not rely on marked repeat: zero-prefaced impersonations do not break with prior topical talk, or offer recontextualisations of prior utterances, but are continuous with them. The impersonator simply persists with the action in which the impersonated is currently engaged.

5.1.2. Local occasioning and preface formats in single-scope impersonations

Aside from mirror impersonations, single-scope impersonations are set in discrete discourse scenarios that are set apart from the scenario oriented to by topical turn-by-turn talk. For this reason, prefaces diegetically signal the shift to other semantic spaces via the formats discussed below. While

$$X_{I1} \text{ be like } Y_{I2} [\text{IMP}]$$

constitutes a simile and discloses a direct point of comparison between similar activities occurring in the two discourse scenarios I_1 and I_2 , other 'story abstracts' simply transfer one element (E) from topical talk (I_1) to a discourse scenario (I_2) that is explicitly marked as disjunct from I_1 via temporal (t), spatial (l), epistemic (e), and modal (m) space builders. In addition, the activities occurring in the discourse scenario which has therefore been established need not correspond to those in I_1 . Yet while these space builders help the narrative world and the actions (A) occurring in that world to be differentiated from the world oriented to by the current topical talk, the prefaces display the extent to which the ensuing impersonation is locally occasioned by prior talk in that these prefaces make salient an aspect of said talk (E) which serves as a trigger for a 'joke first' (cf. Schegloff, 1987) performance. This format can consequently be formalised⁶⁵ as

$$(e/m)\text{-}E_{I1} \equiv / \models E'_{I2}\text{-}A\text{-}(t\text{-}l)[\text{IMP}]$$

⁶⁵ This formalisation serves as shorthand for describing a common impersonation preface. In consequence, usage of the symbol " \equiv " is not intended to denote a relative complement in this context, but is meant to simply make a distinction between the elements contributing to the preface format. Similarly, parentheses are used here to indicate the optional status that epistemic, modal and spatiotemporal space builders enjoy.

Before elucidating upon this format further, it is important to first consider the highly common X_{I1} *be like* Y_{I2} [IMP] format as occurs in “Brothers”. The impersonation in this sequence takes place after two co-participants briefly argue over the optimum way to perform a song. In comparing the activity being engaged in during prior topical talk to those of a fictional scenario, NF’s diegetic preface (line 01-02) clearly shows that his contribution is relevant to prior discourse, as this discourse is the source of his assessment. In addition, NF’s contribution is ‘noteworthy’ and interesting to his co-participants, as it displays the behaviour engaged in during the previous turns as somewhat deviant and laughable. The simile allows for an emphatic assessment of prior talk.

Once the audience has acknowledged the preface through tokens of appreciation in the form of laughter (line 03), NF is licensed to perform an impersonation sequence that further illustrates the point made in the preface.

NMTB SE23E10_2_03:00 (Brothers)

```

01  NF:  YOU twO are like (.) brothers on
02        BUnk beds arguing (-)
03  AU:  @@
04  NF:  <@ ↑that's no way dad's favourite record
05        ↓it IS you idiot;
06  AU:  @@@
07  NF:  <<f> don't make me drop lego on your head;
08        from [up high > @> (--))
09  AU:  [ @@@xx@
10  NF:  okay.

```

Comparing the activities underway in current topical talk to a distant discourse scenario seems to be especially apt for triggering impersonation sequences (see also for example “Alan Bennett” in chapter 5.2.), which serve to further exemplify the comparison and allow for inferences to emerge. Note that this format is likewise based on the assessment of prior turn-by-turn talk, and thus parallels the prefacing strategies identified for mirror impersonations.

The second format identified for single-scope impersonations does not aim to compare activities between disjunct discourse scenarios, but places elements of one discourse scenario in another. Bear in mind, however, that the likelihood thus construed relates to elements (e.g. characters) of topical talk and not activities per se.

These storyworlds are disengaged from topical turn-by-turn talk via a number of distance markers that do not explicitly reflect on the relationship which exists between the two scenarios contributing to the impersonation blend; on the contrary, they first and foremost mark these scenarios as disjunct. The worlds set up in such prefaces might be delineated from topical turn-by-turn talk in the following ways:

- a) *Temporal distance (t)*: Future markers as in “I look forward to hearing on CNN [IMP]” (S24E10_1_07:35) or the past tense as in “I hope she was at the zoo, and they all just went [IMP]” (NMTB SE23E02_1_04:31) serve to indicate temporal distance, and act as space builders in the construction of narrative spaces.
- b) *Spatial distance (l)*: Narrative worlds in single-scope impersonations might also be clearly segregated from current topical turn-by-turn talk via spatial distance markers, as is evidenced not only in the zoo example above, but also in NMTB SE23E08_1_03:13, in which the impersonator prefaces his impersonation with “And then she does a drunk Welsh tour guide in Brussels [IMP]”.
- c) *Epistemic distance (e)*: Such prefaces set up ‘belief spaces’, to recruit Fauconnier’s (1997) original conception. Through lexical items such as *I think* (as in NMTB SE23E02_3_04:10 “I think the bloke who put them together just went [IMP]” or NMTB SE23E11_3_03:13 “I think it’s the Lithuanian Boyzone [IMP]”), belief spaces are built that are distinct from, but nonetheless stand in relation to, current topical talk.
- d) *Modal distance (m)*: These prefaces construct what Fauconnier (1997) refers to as ‘possibility spaces’, and often recruit the lexical item *might* to indicate modal distance from topical turn-by-turn talk. NMTB SE24E08_1_05:45 illustrates this. Here, the impersonation is prefaced by “There might be a slight legal issue that you cannot just go and go [IMP].” Alternatively *would* or *could* are used to signal modal distance.

These markers demarcate the boundary between the narrative space in which a single-scope impersonation is set and topical turn-by-turn talk. Prefaces recruiting the markers and illustrating the macro-structure of the narrative spaces thus framed are (and this an issue which should not be overlooked), locally occasioned in the sense that the scenarios sketched are either directly relevant to the action currently underway

(e.g. serve to answer a quiz question) or fall under what Schegloff (1987) refers to as the ‘joke-first practice’, a feature which is licensed by the pop quiz setting. The jokes always have as their source an element E of topical talk, but briefly ‘detour’ from said topical talk in placing that element in a disjunct narrative. Consider “Mexican Building Society”, which occurs during the line-up game. The panellists are confronted with a line-up of five people dressed in stereotypically Mexican costumes, and are tasked with determining which member of the line-up is the former guitarist of 1990s band “Terrorvision”.

NMTB SE20E06_3_01:08 (Mexican Building Society)⁶⁶

01 PJ: i'm looking at number one and i'm thinking (.)
 02 it'd be very nice to open an account
 03 in a mexican building society.
 04 AU: @@
 05 PJ: <@how may pesetas would you like to deposi::t,
 06 AU: @@@
 07 PJ: chan i incherest you in a chigh interchest accou=ount,
 08 AU: @@@
 09 PJ: i chave to remind you you chave limited
 10 access to your money::, @>
 11 AU: @@
 12 PJ: do(h)n't you(h) wa(h)nt to take these?
 13 AB: wh-what you wanna do to them?
 14 PJ: (h)HOMe
 15 AU: @@

Like in “The Proprietor”, the impersonation sequence does not aim at providing a SPP to the quiz question, but constitutes a joke-first. However, the joking has as its source an element of this topical talk: a member of the line-up (E) that is currently in focus in topical talk (I₁) is assigned a deviant personality in a disjunct scenario (I₂). The mappings that link the current discourse scenario to the narrative space are explicitly set up. For joking purposes, the impersonator here construes a likelihood between an ‘element’ (namely Nr. 1, the member of the line-up) in I₁ with the protagonist of the narrative scenario he establishes as a means of entertainment. Such joking sequences are certainly do not occur at random, but are locally occasioned and methodologically linked to prior talk.

⁶⁶ PJ is here aiming to emulate a Spanish accent. What is here transcribed as “ch” is realised as the voiceless uvular fricative [χ].

As prefaces to single-scope impersonations provide the macro-structure of the storyworld in which the projected impersonation is set, they will usually supply the following:

- a) Obligatory: The voice(s) speaking and acting in the story space (I_2) and their relation to current discourse scenarios (I_1). Consider for example NMTB SE21E05_1_05:29 “She ($E \in I_1$) was like the Jackson Five’s father ($E \in I_2$) [IMP]” and NMTB SE19E06_3_04:15: “I think Nr. 5 ($E \in I_1$) is (\equiv) the proprietor ($E \in I_2$) of a remote guest house [IMP]”, which relate an element E of topical talk (I_1) to another element in a disjunct narrative space (I_2).
- b) Obligatory: The actions A performed by the character, which are either provided in the form of a quotative (*go, be like*) or made explicit, as, for example, in NMTB SE21E05_3_03:20: “It’s one of these scenes where the villain tries to pass himself off as the member of the team to escape the night club; with number two is it [IMP]”. Alternatively, the characters introduced might metonymically represent actions and be and based on shared cultural knowledge, as in “The Proprietor”, in which the introduction of the voice and the spatial setting inhabited by said voice seem to suffice in order to infer event structures.
- c) Optional: The spatiotemporal properties of the story space (e.g. remote guest house, zoo, night club). These are, however, often metonymically inferred in impersonations that feature well-established characters whose spatiotemporal surroundings may be metonymically inferred (the setting for “Alan Bennett” might thus, for example, most likely be a theatre).

It may be concluded, then, that single-scope impersonation prefaces provide brief abstracts of their narrative space, and display the joke’s local occasioning in prior topical talk. They either follow the simile format $X_{I_1} \text{ be like } Y_{I_2}$ [IMP] that explicitly reflects on the relationship between activities occurring in both input spaces, or adopt the format $(e/m)\text{-}E_{I_1} \equiv / \models E'_{I_2}\text{-}A\text{-}(t\text{-}l)$ [IMP], in which element E triggers the setting up of a narrative scenario for entertainment purposes. In doing so, these prefaces allow for the observation of elements of the current discourse scenario in relation to a disjunct one. The ensuing clash of semantic worlds allows for inferential qualitative

assessments, and, if successful, for humour to emerge. Impersonations following these prefaces do not, however, feature the marked repeat found in mirror impersonations, and their relevance to topical talk is consequently only displayed in their prefaces.

5.1.3. Local occasioning and preface formats in double-scope impersonations

The methods used to display the impersonation's relevance to topical talk and the impersonation preface formats found in double-scope impersonations strongly parallel those formats recognised in single-scope impersonations. However, prefacing sequences to double-scope impersonations tend to be more elaborate and developed over a greater number of turns than those observed in single-scope impersonations. Double-scope impersonations are often locally occasioned through (rhetoric) questions that help to ensure joint attention to be paid towards an aspect of topical talk deemed noteworthy by the impersonator.

Consider, first of all, the X_{I1} *be like* Y_{I2} [IMP] format also found in single-scope impersonations. In “Holiday horse”, impersonator NF likens a co-participant's puzzled look to that of a horse “that's been asked to book a holiday [IMP]”. He therefore provides a simile, effectively a direct comparison between the two activities, which allows for assessments to emerge. The format thus exactly parallels the one detected for single-scope impersonations, with the sole semantic difference of the latter involving non-human points of comparison.

Similarly, the format $(e/m)-E_{I1} \equiv / \models E'_{I2}-A-(t/l)$ [IMP] accounts for the remaining double-scope impersonation prefaces. Prior to the performance of double-scope impersonations, their protagonists (E), spatiotemporal setting (t/l) and, activity (A) have to be elucidated upon. This becomes especially evident in “Stevenson's rabbit” (NMTB SE19E05_1_08:11), in which the narrative scenario of the double-scope impersonation is built up over a number of turns.

NMTB SE19E05_1_08:11 (Stevenson's rabbit)

01 DB: do you think it is uh (.)
02 a little (.) steam-powered dildo

03 with little (.) chcoloate (.)
 04 uhm toffee apples on the side?

05 AU: @@@
 06 PJ: a steam-powered DILdo?
 07 AU: @@
 08 DB: yeah
 09 PJ: <@ <<f>oh if i'm pleasuring myself
 10 i have to do it in the victorian
 style; >
 (-)
 11 AU: @@
 12 PJ: <<f> where's my coa:l james?> @>(-)
 13 AU: @
 14 PJ: stevenson's rabbit;
 15 <<steam sounds in three intervalls>>
 (--)
 16 <<whistle sound, twice>> (--)
 17 AU: x@x@x
 18 RI: <@ (isenba:rt) what's under your hat?
 @>
 19 AU: @@
 20 RI: <@ i have QUIte a SURprise for you
 mylady. @>



First, a rhetorical question consisting of a marked repeat (“A steam-powered dildo?”) of a co-participant’s prior utterance directs attention to the trigger E of what is to become a series of impersonation sequences. This marked repeat is followed by a first impersonation, a mirror-impersonation, which introduces the relevant action A and the temporal setting *t*. However, the agent, and consequently the perspective being profiled, changes between this first impersonation and the double-scope impersonation that follows. For this reason, the impersonator has to explicitly introduce the new voice, which he does exclusively by providing its terminus technicus, so to speak (“Stevenson’s rabbit”). This term could itself be analysed as a double-scope blend, uniting spaces containing frame knowledge about steam engines and brand names of phallic sexual devices – however, this shall not be discussed at present. Of more immediate interest is the fact that this sequence illustrates the importance assigned to identifying the voice acting in a particular impersonation. The

activity in which this machine is engaged can be inferred metonymically and through inferencing strategies of the action presented. These strategies interact in meaning-making. The sounds and body movements displayed by the impersonator can only be understood in reference to the frames made available by the introduction of the voice, whilst the performance serves as a hint as to the activity and, in a second step, illustrates and brings this activity before one's very eyes.

This rather intricate setting-up of narrative spaces for double-scope impersonations is also evident in "Metal Detector", whose preface is constructed by two participants over an extended sequence of turns, providing the voice (E) as well as event structures and their sonic dimension (A). However, even if only one interactant sets up the diegetic preface, it is still introduced with great care, as "Babybel" illustrates. This example features fictivity markers ("I like to imagine") prior to providing information on the voice about to speak (E, Babybel), the activity (A) engaged in ("walking / getting pulled up in a net"), and the location ("a forest").

In essence, then, the local occasioning and prefacing formats of single and double-scope impersonations are almost identical. Both exhibit two main characteristics: simile, which compare agents and an activity in the current discourse scenario with those of a scenario that is marked as clearly discrete, and the linking of trigger element E in the current discourse scenario to a role and activity in a discrete scenario. Marginal differences as to the degree of elaboration can nevertheless be traced.

5.1.4. Discussion

This section has shown that apart from providing the macro-structures of the narrative world in which a given impersonation takes place, such prefaces fulfill a number of other interactional functions as well. First and foremost, they serve an important role in displaying the local occasioning of impersonations and the narratives in which they are set, and consequently serve to warrant tellability (Sacks, 1974, 1992). Telling stories and performing voices in said stories requires showing the proposed narration's relevance to previous topical talk and getting co-participants to be interested in the story (Norrick, 2000). After all, the narration of a story usually requires the negotiation of a slight "shift in focus" (Solomon, 2004:261), as well as claiming an extended sequence of turns-at-talk and the alignment of co-participants as

story-recipients. Prefaces may thus project multi-turn discourse units, and consequently suppress the turn-taking model valid for unmarked topical turn-by-turn talk. Such discourse units are consequently expected to be noteworthy, and to relate to the current theme of the interaction; this is why prefaces tend to explicate their relevance to topical talk through such fixed phrases as *this reminds me of* (i.e. the ‘Speaking of x’ format identified in Jefferson, 1978). Secondly, prefaces may already project the intended reception of the narrative under construction. As such, they may already display whether, for example, the proposed narrative is to be taken as humorous or not.

When comparing such techniques across blending types, it becomes evident that differences in local occasioning and prefacing formats between the various types of impersonation blends can indeed be traced. While mirror impersonations do not require space builders, but rather DMs such as “Oh” and the marked repeat of prior contributions, single and double-scope impersonation blends on the other hand set up discrete discourse scenarios through space builders such as “I imagine”. However, the dividing line between single and double-scope impersonations is not quite as clear-cut. Double-scope impersonations rely on the same formats as those identified for single-scope, but usually require more elaborate prefacing sequences. These findings are summarised in table 2.

Table 2: Overview of preface formats across blending types

	Distance markers	Anchoring in topical talk	Pragmatic function
Mirror impersonations	<ul style="list-style-type: none"> • Disjunct marker (e.g. “Oh”) • Zero 	<ul style="list-style-type: none"> • Trigger profiled through question • Marked repeat 	<ul style="list-style-type: none"> • Attribution of identities, affective stance or tacit motives to co-participant
	<ul style="list-style-type: none"> • Redescription of activity 	<ul style="list-style-type: none"> • Simile (“be like”) • (Re-)description of original activity 	<ul style="list-style-type: none"> • Assessment of prior utterance by co-participant • Attribution of identity or affective stance to co-participant
Single-scope impersonations	<ul style="list-style-type: none"> • Space builder (epistemic, modal, spatiotemporal) 	<ul style="list-style-type: none"> • Subject of topical talk transferred to distinct discourse scenario (trigger element may be profiled through question(s) and marked repeat(s)) 	<ul style="list-style-type: none"> • Illustration of joke-firsts
	<ul style="list-style-type: none"> • Redescription of activity in prior topical talk + space builders 	<ul style="list-style-type: none"> • Simile (“be like”) 	<ul style="list-style-type: none"> • Assessment; displaying theme of topical talk as comical

Double-scope impersonations	Cf. single-scope	Cf. single-scope	Cf. single-scope
-----------------------------	------------------	------------------	------------------

Can, then, these findings be understood to constitute the first evidence of cognisers' orientation to the typology put forward by Fauconnier & Turner (2002)? While the fact that mirror impersonations do not show the space builders found in more complex frame-integration scenarios initially points towards a validation of the difference between mirror and single or double-scope integration impersonations, a closer look actually reveals that the formats identified all relate to different actions, thereby further supporting the hypothesis outlined in chapter 4.4. Mirror impersonations exhibiting a preface or disjunct marker and single and double-scope impersonations recruiting the X_{I1} *be like* Y_{I2} [IMP] format all contribute to constructing an assessment of an element of prior topical talk. Single and double-scope impersonations recruiting the (e/m) - $E_{I1} \equiv / \models E'_{I2}$ - $A-(t-l)$ [IMP], on the other hand, usually serve to introduce illustrations of joke-firsts. These actions span across blending types, and it is these actions that interactants seem to orient to. However, further evidence is needed to support this hypothesis. The following sections will therefore investigate further methods employed in the local management of impersonation sequences in order to determine whether they do indeed serve to further legitimate the original blending typology or not.

5.2. Demarcation markers: Parsing the diegetic from the mimetic

The previous chapter outlined the mechanisms employed for creating story-worlds in which impersonations are set. While these prefaces constitute an impersonation sequence's diegetic aspect, the impersonation itself is a mimetic visit to the story-world. Despite this, parsing the diegetic from the mimetic requires careful and subtle changes in *footing* (Goffman, 1981). Footing, Goffman (1981:128) writes, involves "the alignment we take up to ourselves and the others present as expressed in the way we manage the production or reception of an utterance." In other words, impersonators have to demonstrate to their co-participants that they are now speaking in the voice of a character in their own story-world.

This section will explore the methods (i.e. demarcation markers) used to signal such changes in voice across blending types by first of all focusing on the use of *quotative frames* and secondly on the discontinuation of postural orientation and gaze in mimetic *oratio recta*. These two methods constitute the main demarcation markers identified in previous research. Performative semiotic resources, such as grimacing or code-switching, may also function as demarcation markers, and will be considered in chapter 5.3.

5.2.1. Quotative systems

Quotative systems have mainly been explored in studies on *oratio recta*, or direct speech⁶⁷, which has long been of interest to researchers from a multitude of disciplines. This has in turn lead to diverse sets of terminology to arise, all aiming to grasp the phenomenon of changing voices and thus restage or perform the dialogues of another. Tannen (1986:312), for example, talks of *constructed dialogue* to draw attention to the fact that such enacted quotes are not word-for-word copies of original quotes, but are consciously created “dramas”. In a similar vein, Macaulay (1987) talks of *quoted direct speech* to emphasise the polyphonic nature of such voice changes. The changes in footing in *oratio recta* are commonly signalled, previous research has found, through quotatives such as *say*, *be (like)*, *go* or *do*.

Similar to such constructed direct quotes in conversational narrative, impersonations may also be introduced by *imitation clauses*, to use Vandelanotte & Davidse’s (2009) more accurate terminology. Such quotative frames⁶⁸ signal the “syntactic bracketing” (Kohn & Franz, 2009:259) of the direct speech of a character other than one’s own, and thus shift the deictic focus (cf. Bückler, MS) from the impersonator and the current discourse scenario to the impersonator and the story-world inhabited by the character. It is clear, then, that they constitute a major means for signalling voice changes. Normally, quotatives include more neutral reporting

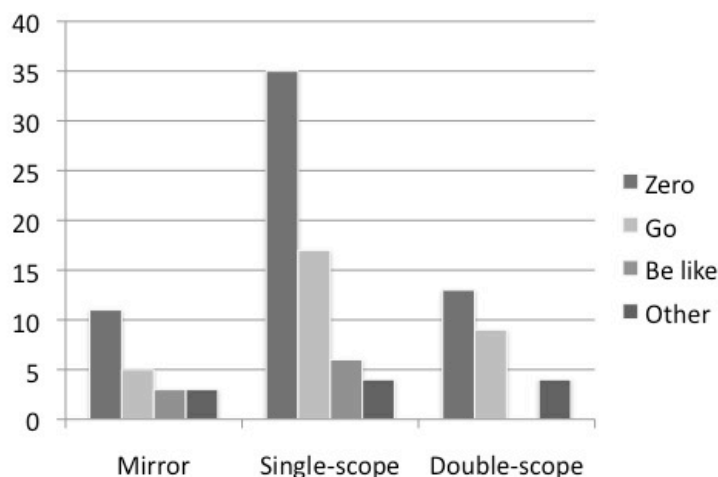
⁶⁷ Following Golato (2002:26), the term *direct speech* is here used to refer “to those kinds of reported discourse in which the speech and body behaviors of others or oneself are introduced as presently ongoing speech.” The term *reported speech* is not used in conjunction with the phenomenon, as Tannen (1986:311) suggests, it is a “misnomer”. Such utterances hardly ever accounts for that which has been uttered on another occasion, i.e. they do not constitute factual reports of actual events.

⁶⁸ *Overt introducers* in Ferrara & Bell’s (1995) terms, or *dialogue introducers* in Johnstone’s (1987) terminology. In the following, I will only use the term quotative, but will nevertheless take it to refer to imitation clauses, as only those quotatives that allow for embodied voice changes will be discussed.

verbs that do not exude a particular pragmatic effect (cf. Johnstone, 1987; Romaine & Lange, 1991; Tagliamonte & Hudson, 1999), such as *say* and *tell*, or verbs illustrating a certain voice quality, like *whisper* or *shout* (so-called *graphic verbs* in Labov's (1972a) terminology), a certain manner (*whine*), or emotion, for example *weep* (cf. Yule, 1998). The latter three are, however, not as common in oral narrative as they are in (more traditional) written narrative (cf. Macaulay, 2001; Yule, 1998). In addition, the English quotative system is not entirely stable, and has always been subject to 'fashion' trends, with quotatives such as *be all* being used exhaustively over a given period and then eventually discarded (Kohn & Franz, 2009).

As impersonations rely heavily on non-verbal aspects, such as sound-effects and body movements (gesturing and mimicking), traditional *verba dicendi* that focus on the act of speaking only, for example *say* or *tell*, are not found in the NMTB impersonation corpus.⁶⁹ *Verba dicendi* such as *say* would even be considered as ungrammatical in more mimetic quotes, Singler (2001) argues, as they do not account for their multimodal dimension. Instead, impersonations are introduced either through "semantically nonreportative" (Vandelanotte & Davidse, 2009:777) quotatives that allow for multimodal performances, such as *be*, *be like* and *go*, or – and this applies to the vast majority of cases – a zero-quotative.⁷⁰

Table 3: Distribution of quotatives across blending types (n=109)



⁶⁹ In NMTB, however, *say* is used as a quotative in more verbatim and less embodied reports of previous events that are presented as factual.

⁷⁰ Interestingly enough, these quotatives represent those "new quotatives" as classified by Yule (1998:287) in his grammar, thereby providing evidence for the degree to which non-standard language is used in the corpus.

As illustrated in Table 3, the most common quotatives in the NMTB corpus are zero-quotative, followed by *go* and *like*. Only single-scope impersonations show additional quotatives, such as *all that*, bare *be* and bare *like*.

The preference for zero-quotatives in impersonations in my corpus may have numerous causes, not least grammatical reasons. In this respect, Kohn & Franz (2009) have shown that a subject in the third-person singular prefers to use zero-quotatives. As most single and double-scope (and even mirror) impersonations are characterised by third-person singular subjects, this would be one way to explain its disproportionate popularity in my corpus. However, the zero-quotative is also used in instances in which a second-person singular subject is present, as is the case in prefaces of the “You’re V-ing + Impersonation” format. For this reason, the grammatical argument put forward by Kohn & Franz (2009) cannot account for all instances of zero-quotative usage. What seems to be of greater importance is the theatrical, mimetic dimension projected by a zero-quotative. Previous research (e.g. Mathis & Yule, 1994) has noted that zero-quotatives are especially common for introducing a quote that is set in a highly “dramatic scenario” (Yule, 1998:286), whereby the roles have already been established either in the preface or previous discourse. Further to this, zero-quotatives seem to have a propensity for introducing scenarios that do not necessarily have to be representations of (f)actual events, i.e. they may introduce even more fictional talk. According to this line of argumentation, zero-quotatives create a greater degree of drama in an already established story-world, and add a certain degree of immediacy to the impersonation. As they do not explicitly reflect on authenticity, they are more independent than impersonations introduced by a quotative: they are ‘as’ rather than ‘as if’. In short, the impersonated character is presented as more autonomous, as less ‘controlled’ by the impersonator than is the case for *oratio recta* introduced with a quotative. That a quotative that signals this much independence should also account for many of the mirror impersonations is thus surprising. After all, these mirror impersonations set out to explicitly attribute deviant behaviour to co-participants rather than letting independent voices act out (fictional) narratives. For this reason, a higher degree of framing by the impersonator could reasonably be expected.

The second-most frequent imitation quotative in my corpus is *go*. *Go* is considered to be a relatively young quotative, is believed to have originated in the US approximately forty years ago (Vandelanotte & Davidse, 2009), and has propagated

across the English-speaking world since. Macaulay (2001), for example, found it to be the most prominent quotative in his study on quotatives used in young Glaswegians' storytelling. This prominence was, however, preceded by a decline in usage in the 1990s, when *be like* became more popular (Buchstaller, 2006). My findings support Macaulay's (2001) and Buchstaller's (2006) claims regarding the renewed popularity of *go*, and show that *go* remains a key quotative in British varieties of English.

The etymology of the quotative *go* can be traced back to "imitative use[s] of *go*" (Vandelanotte & Davidse, 2009:798), which have been documented from the 18th century onwards in such utterances as "Tang goes the harpsichord" (1812, H. Smith & J. Smith, *Rejected Addr.* 117; OED online). When considering its origin as a quotative used to introduce onomatopoetic renderings, the main functions which *go* is associated with today are hardly surprising. Rühlemann (2008) finds it to occur primarily in conversational narrative, and Tagliamonte & Hudson (1999) claim that *go* may be used to introduce non-lexicalised sounds in said environment. Macaulay's (2001) findings support their results, whilst Yule (1998) adds that *go* shows a certain predilection for occurring in situations where non-factual scenarios are 'quoted'. In addition, Johnstone (1987) illustrates that *go* does not have a lexical meaning (i.e. it does not indicate a manner in which something was uttered). With this flexibility regarding its relationship to (f)actual events (cf. zero-quotatives) namely its tendency to occur prior to fictive, yet highly dramatised scenarios which include non-lexicalised sounds, it is little wonder that *go* should feature prominently in the corpus.

It is thus highly unexpected that the functionally similar quotative *be like* and variations thereof do not occur as frequently. Although *be like* does appear to be a relatively novel quotative (with first documented usage dating back merely to the 1960s), it has been particularly well-researched, and been subject to much attention from pragmatics, historical linguistics and grammaticalisation studies, corpus linguistics and contact linguistics (for references, see Kohn & Franz, 2009; Fox & Robles, 2010). Although the quotative originated in Californian American English, it is also well-documented in British varieties of English (cf. Macaulay, 2001; Tagliamonte & Hudson, 2008), and, with its usage being mainly recorded among young people⁷¹, could thus be expected to feature in the NMTB corpus.

⁷¹ The popularity of *like*, which has been described as an "explosion" (Fox & Robles, 2010:715) does not only concern its function as a quotative, but affects a variety of discourse marker functions, thereby fuelling further research interests in the different uses of *like* (e.g. Miller & Weinert, 1995). Its

As a quotative, *be like* is used to introduce performed, enacted quotes, as well as attitudes, evaluations, inner thoughts and feelings (Golato, 2000; Tannen, 1989; Yule, 1998; Fox & Robles, 2010; Streeck, 2002; Tagliamonte & Hudson, 1999). Furthermore, it does not necessarily attribute authenticity to the quote (cf. Kohn & Franz, 2009; Fox & Robles, 2010). One may therefore reasonably expect that it is even more suitable to be used in conjunction with impersonations, especially given that *be like* as a quotative originates from the etymologically oldest meaning of *like*, the comparative. This is still evident in its use as a quotative; it marks a comparison with an event, suggesting that which is to follow is not a precise reproduction of a speech event, but an approximation (Fox & Robles, 2010; Golato, 2000; Yule, 1998).

With its focus on the enacted, multimodal, evaluative and emotive as well as its tolerance for the fictive, *be like* would consequently be a prime candidate for framing impersonations. This preliminary hypothesis is further supported by Fox & Robles' (2010:716) recent study, which found that the much rarer variant of *be like*, namely *it's like*, is used to enact "the thought, feeling or attitude presented rather than describing it", thus adding "bodily displays" (ibid:725) to a quote. As a supplement to this, Streeck (2002:581) refers to *be like* quotes as "body quotations" due to their tendency to introduce embodied performance. However, variants of *be like* only introduce a small number (n=6) of impersonations within the corpus.

As the majority of speakers in the NMTB corpus are male, questions concerning gender naturally arise. Yet while the preference for *be like* is, in American varieties of English at least, linked to gender, *be like* in British English does not seem to be preferred by female speakers (Buchstaller, 2008) anymore, suggesting that gender might not play a role in the lack of use of the quotative *be like* in the corpus.

Due to the size of the data set, I will not be able to enquire further into the statistical distribution of quotatives, but will rather offer a qualitative analysis that aims to shed light on the functions performed by the various quotatives according to blending type. To this end, this chapter will explore the use of quotatives in each blending type in detail to determine whether blending complexity could have an impact on the use of quotatives in impersonations.

popularity as a quotative has been allocated to recent trends in US-American culture that feature "self-revelation as a preferred cultural mode" (Ferrara & Bell, 1995:283). A similar viewpoint that links the success enjoyed by *be like* as a quotative within a culture obsessed with performance and self-revelation is addressed by Streeck (2002) and Fox & Robles (2010).

5.2.1.1. Quotatives in mirror impersonations

In total, 61% (n=56) of all mirror impersonations are introduced without a quotative. As such, it will be useful to first consider zero-quotatives in mirror impersonations before comparing them to instances of *go* and *be like* in a second step.

When investigating the sequential environment in which zero-quotatives are embedded in mirror impersonations, it appears that they are especially prone to be followed by impersonations that restage, rephrase or further expand upon a previous utterance by a co-participant, with a focus on exposing the alleged ‘true’ motives and thoughts behind the original utterance – these are functions which previous studies (e.g. Ferrara & Bell, 1995) only attributed to the quotative *be like*. Therefore, zero-quotatives in mirror impersonations introduce performances that claim to expose internal aspects, inner monologues and motives, rather than externalised actual utterances. In doing so, these impersonations reframe the original utterance and attribute a distinctly fictional dimension to the original quote, projecting an incongruity between the actual, original utterance, and the inner thoughts opposing it. The original utterance is thus displayed as ridiculous. The impersonation, in consequence, does not recount a previous telling, but adds a novel fictional (and mostly derogatory) semantic aspect to it, claiming to uncover the actual, hidden motives behind the saying. Consider again “Northern lass”.

NMTB S24E08_pt2_01:49 (Northern lass)

```

01 DV: do we have to stand yea?
02 PJ: yes we do=o
03 DV: [great
04 PJ: [diana vickers
05 DV: very excited
06 LM: what about STAndin?
07 DV: [ye(h)es
08 AU: [@@@ (1.9)
09 LM: <@ <<northern accent> i'm a simple northern girl
10 with a simple wishes
11 AU: @[@@
12 LM: [i wish i could go to london and stand, > @>
13 AU: @@[@ (1.6)
14 LM: [<<f> WEll tonight your drEAmS come true.>

```

In this sequence, LM's impersonation of DV (lines 09, 10 and 12) expands on her original contributions within the framework of established stereotypes regarding the north of England, projecting that these are her actual motives behind expressing her approval of having to stand up. This focus on internal, psychological aspects behind utterances that seem to be projected by zero-quotatives is evident in a number of samples, such as "Travel Lodge":

NMTB S23E11_2_09:26 (Travel Lodge)

01 FB: i heard a story about you; (.) right;
 02 DJ: <<p> good>
 03 FB: it was a kiss and tell,
 04 DJ: right
 05 FB: by a lady who said, (.)
 06 DJ: yea
 07 FB: that YOU took her to a premier TRAvEl lodge,
 08 DJ: ↑ME?
 09 AU: @@
 11 FB: ye~a
 ((...))
 18 i'm not wrapped up i'm single=
 19 FB: yeah=
 20 DJ: so a single guy
 21 should be allowed to do what he wants
 22 FB: <@ yeah. if i'm not DAtin her- (.)
 23 it's the TRAvEl lodge. @>
 24 AU: @@

As in the previous example, this impersonation (lines 22-23) discerns a humorous 'true' intention from a prior contribution, thereby allegedly laying bare the more polite phrasing of the original utterance. These examples go to show that a zero-quotative in mirror impersonations is employed to reveal the hidden motives behind an actual utterance, thus creating incongruity and presenting the original as a laughable.

In contrast, consider the following sequence that recruits the quotative *go*.

NMTB S23E10_3_05:40 (Flirting)

01 DO: aston did you get her number?
 02 i LIterally saw her looking at you go,
 03 <@ HEy babe, @>
 04 and you went,

05 <@ ↑yeah it's been a long time

06 i, ((mimicks phone call)) @> (1.0)

07 like that.

08 AU: [@

09 DO: [did you do that,=

10 AM: it HAS been a long time, yeah.



Unlike in impersonations recruiting a zero-quotative, the voice-changes in this sequence do not showcase internal motives, but claim to restage actual utterances (lines 03, 05-06). Furthermore, characters in this sequence are not introduced in the preface. Neither is the event structure projected in the preface (line 01 – ‘getting phone number’) reflected in the two impersonation sets that follow. It thus appears that in the NMTB corpus, the quotative *go* is used in mirror impersonation blends in order to signal role changes in impersonation scenarios which

- a) focus on the externalisation of utterances, and
- b) whose character(s) and event structure have not fully been established in the turn(s) preceding the impersonation.

Impersonations preceded by *go*, additionally show somewhat more embodied aspects than those preceded by zero-quotatives. This might be due to zero-quotative impersonations focussing more on the internal while *go* introduces externalised utterances that were (allegedly) *uttered* rather than purely *thought*, which naturally allows for a greater degree of embodiment.

The same holds true for instances of *be like* in the mirror impersonations collection. In such cases, *be like* is used to introduce highly embodied ‘quotes’ that are often non-linguistic, i.e. non-lexicalised sounds or purely visual quotes (e.g. grimacing). Consider again “Bemused look”.

NMTB SE24E11_1_12:00 (Bemused look)

01 NF: i've NEver seen anyone
 02 look more bemused;
 03 you're like-



In this sequence, NF visually depicts the affective stance displayed by the impersonated in the previous turn. Similar to enactments preceded by *go*, those following *be like* thus claim to restage and, in doing so, accentuate and caricature actual previous actions by another character. Therefore, unlike impersonations using zero-quotatives, *be like* and *go* in mirror impersonations seem to introduce performances that display actions that were allegedly actually carried out in the turns preceding the impersonation. However, *go* seems to project verbal actions, whereas *be like* allows for the emotional or affective stance displayed in the original action to be focussed upon. Mirror impersonations preceded by zero-quotatives, on the other hand, tend to disclose internal motives that were not made transparent by the original action rather than bodily actions and verbalisations. In consequence, *go* and *be like* permit a greater degree of embodiment than found in zero-quotative impersonations.

In addition, restaged actions introduced through *be like* topicalise the emotional or affective stance displayed in the original action. Similar to those impersonations introduced via *go*, the content repeated and accentuated is not familiar to all participants in the interaction-as-broadcast. In “Flirting”, for example, the television audience was not able to bear witness to the original exchange. In “Bemused look”, the original negative stance taken by the impersonated is briefly shown in the interaction-as-broadcast, but is not emphasised. In effect, then, the impersonators make visible actions which had previously been hidden from view, meaning that these impersonations require the greater degree of framing provided by a quotative in order to ensure that all participants in the interaction-as-broadcast are able to detect the voice change and its referent. In contrast, impersonations introduced through a zero-quotative only restage previous actions which all participants in the interaction-as-broadcast had access to and which can be assumed to be part of the shared Common Ground. Table 4 summarises these findings.

Table 4: Overview of quotative functions in mirror impersonations

	Projected impersonation focus	Familiarity of participants with projected content	Embodiment
Zero-quotative	Ascribed tacit motives behind previously articulated utterances	The input utterance is known to all participants, but the ‘hidden motives’ attributed to the impersonated are not → lesser degree of framing required	Low
<i>Go</i>	Accentuated restaging of previous (largely linguistic) utterances by co-participant	The input utterances is not known to all participants → higher degree of framing required	Medium
<i>Be like</i>	Accentuated restaging of previous (largely bodily) actions that display affective stance	Not all participants are aware of the original action → higher degree of framing required	High

In all, then, the use of quotatives in mirror impersonations seems to differ from the functions identified in previous literature in two main ways: while *be like* indeed focuses on the display of inner states, it does so with inner states that had actually been expressed and hence made visible. Zero-quotatives, on the other hand, attribute hidden motives to previous actions. This stands in stark contrast to the claims made by Fox & Robles (2010) who found that *be like* may be used to display inner thoughts. In mirror impersonations in my corpus, this function is embodied by the zero-quotative.

Secondly, previous studies (e.g. Yule, 1998; Kohn & Franz 2009) found that *go* and *be like* are used to introduce more fictitious scenarios. In the NMTB corpus, however, both *go* and *be like* are found in impersonations that claim to accurately restage previous actions, thus accentuating certain aspects which the impersonator finds noteworthy. The following chapters will establish whether these discrepancies from previous findings might be due to frame integration type.

5.2.1.2. Quotatives in single-scope impersonations

As a zero-quotative is, again, the most common way to introduce an impersonation in single-scope impersonations, this chapter will first investigate the environments in which this type occurs before comparing it to *go* and *be like*.

Unlike in mirror impersonations, zero-quotatives in single-scope impersonations are not used to make visible hidden motives behind actual utterances, but are utilised as a means of introducing externalised actions that occur in highly fictive, but well-established story-worlds in which all characters and event structures have been clarified. “Polish band” illustrates this point:

NMTB S20E05_1_04:23 (Polish band)

01 PJ: five always had that vibe about 'em;
 02 like, (.)
 03 the band were something they did
 04 when they weren't ON Site. (-)
 05 AU: @@@
 06 PJ: five didn't break up
 07 they were just actually replaced
 08 by a chEaper Polish band;
 09 SA: ye[ah
 10 AU: [@@@
 11 PJ: <@ <<Polish accent> we come dance sexy
 12 for you (.)
 13 AU: @
 14 PJ: two pounds (.)
 15 AU: @
 16 PJ: cash in hand only (--)
 17 AU: @@@
 18 PJ: then we do conservatory. > @>
 19 AU: @@

In this sequence, PJ first establishes a fictive story-world in which the members of the 1990s boy band Five work as “brickies” rather than musicians, and are consequently replaced by “cheaper” Polish workers. Once he has set up the story-world with its characters and event structure, he does not need to further frame the impersonation through a quotative, but can provide a mimetic, theatrical visit to the discourses occurring in the alternative world. The same applies to “Cooking with villains”:

NMTB SE21E08_1_04:34 (Cooking with villains)

01 BB: yeah. (.) COOking with Villains. (h)>
 02 <@ i've just broken
 03 into this beautiful house; (.)
 04 AU: @
 05 BB: the alarms will go off
 06 in thrEE minutes; (.)
 07 AU: @@@
 08 BB: time for a LOvely salmon salad. @>

09 AU: @@@



This impersonation occurs in an exchange commenting on the recent boom in cookery shows in Great Britain. There is a low degree of authenticity (i.e. the impersonations provide mimetic visits to a world that is framed as fictive, yet alternative to the one shared via Common Ground) and the impersonations are largely verbal, with only limited recursions to embodied features, such as gesturing in “Cooking with villains”. As in “Polish band”, the impersonation occurs after a preface that provides the central character and event of the space set up, and provides a dramatic illustration of discourse that could occur in the narrative space. This discourse does not, however, add any new value to the development of the story itself. It merely illustrates the story, the event structure and characters of which were already outlined in the preface. This evocative quality of single-scope impersonations introduced through zero-quotative is the main factor that differentiates them from single-scope impersonations introduced through quotatives such as *go* or *be like*.

When investigating impersonation sequences introduced through *go*, it becomes evident that such sequences mainly occur in conversational storytelling when they are used to develop the narrative further. Instead of providing mimetic visits to a story-world already set up, then, these impersonations are crucial to the telling of events themselves. The following examples illustrate this point.

NMTB SE22E01_1_09:25 (Upset French)

01 PJ: I think- i think i think
 02 if there A:RE any french people
 03 watching they're gonna go,
 04 AU: <@ <<ff> AAH quelle est POp quiz



05 PJ: IMB!É!cile;
 06 AU: @[@@
 07 PJ: [((encore)) > @>
 08 AU: @@@

The preface of “Upset French” does not point to the event structure, the contents or the affective quality of the impersonation, as is the case for impersonations introduced with a zero-quotative. It does, however, introduce the voice speaking in the performance itself. The impersonation in this case is delivered in (rather rudimentary)

French, but due to the fact that it is highly embodied, the affective stance it projects should also become evident to co-participants who have little or no knowledge of the language. In addition, the story's main aspect – the rendering of previous discourse on the pronunciation of the French plural *-s* as “imbecile” – recruits a loanword (French: *imbécile*) which will subsequently be shared among the co-participants. Thus, the impersonation is not a visit to a story set up in the diegetic preface, but itself contributes to the event structure.

This phenomenon is also witnessed in longer impersonation sequences introduced through the quotative *go*. Consider the “Jason Orange” sequence discussed in further detail in chapter 5.2.2.2. Here, a lengthy impersonation of singer Jason Orange reflecting on nuclear power is introduced with *go*, and is preceded only by *but then again you see he'd go in and go*, (line 06). In short, the preface does not give any information as to the contents of the impersonation; rather, it purely provides the character about to perform an action. It is the impersonation itself that provides the narrative structure.

Furthermore, *go* is used to introduce shorter impersonation sequences which might, in addition, take place in dialogue with other voices performed by the same impersonator. As a result, *go* is used in the local management of voices that contribute to the narrative development of a story, as is evidenced by the following sequence.

NMTB SE22E01_1_11:07 (Kylie)

01 TM: i've got this image of her?
 02 cos she's she just seems so down to
 03 earth? you know? and she's Fortysomething
 04 and she's dressed
 05 in this roller skating outfit,
 06 i- i've got this image of (.)
 07 all theserecord producers
 08 coming up to her and going,
 09 <@ so kylie (.)



10 we're doing this new rave thing
 11 and it's like this eighties synthesizer,
 12 @>
 12 and her just going,
 13 <@ oh, okay, yea,



14 AU: @
 15 TM: aright? @> (-)
 16 <@ and you got to wear
 17 this futuristic roller skater outfit @>
 18 <@ oh, yea, aright, that sounds (-)
 19 AU: @@
 20 TM: that sounds fine, (.)
 21 how much am i getting paid for this? @>
 22 <@ TEn million dollars @>
 23 <@ Oh ten million dollars (.)
 24 i'll put that into my savings account.(-)
 25 AU: @@
 26 TM: oh i wish i had a baby; @>
 27 [@@@
 28 [you know, that sort of thing.

In this staged dialogue between Kylie Minogue and her record producers, *go* is used to introduce each voice. Once the two voices have been established, TM switches between them without quotative frames, relying solely on performative demarcation markers, such as pitch, accent and gestures. This again points to the fact that zero-quotatives may only be used in environments in which all voices have been firmly established. In sum, then, we find that *go* introduces impersonations that

- a) contribute to the unfolding of the narrative
- b) require local management of multiple voices and
- c) rely on embodied as well as linguistic features.

Single-scope impersonations introduced with *be like*, on the other hand, are first of all used to introduce impersonation sequences which transport attitudes rather than semantic content to further a narrative. Following on from this, *be like* secondly seems to allow for non-verbal items (e.g. screams, onomatopoeic renderings, gestures) to be implemented in the impersonation (cf. Fox & Robles' 2010 findings). The non-lexicalised sounds performed in "Umbrellas", for example, serve the display of the affective stance also evident in "Justin Bieber". This latter sequence occurs in an exchange on how a boy as "sweet" as Justin Bieber could possibly utter the offensive words "don't fucking touch me again". The semantic content of the impersonation does not, in consequence, add novel information to the narrative space in which it is embedded, but provides information as to the (alleged) attitude of the impersonated.

NMTB SE24E10_1_07:25 (Justin Bieber)

01 JG: he might be saying that in that picture
 0 he's like- (.)
 04 <@ <<playful voice> don't fucking touch
 05 me again; (--)



06 AU: @@
 07 JG: double dog dare you,> @>



08 AU: @@@

Be like can, in a similar manner to *go*, also be used to manage multiple voices within one impersonation sequence, as “Umbrellas” further illustrates.

NMTB SE23E09_1_05:37 (Umbrellas)

01 JL: People probably show up with umbrellas
 02 and then when she sings umbrella
 03 they're like,
 04 <@ <<f> AAhaa wh' aa !BRI!!lliant,>@>



05 and she's like,
 06 <@ not again. @>



07 AU: @@@

As the other quotatives found in single-scope impersonations do not occur in either mirror- or double-scope impersonations, I refrain from discussing them in greater detail. They include bare *like*, bare *be*, verb+*like* and speech act verbs (such as *propose*). Table 5 summarises the functions identified for the three most prominent quotatives in the corpus.

Table 5: Quotative frame functions in single-scope impersonations

	Projected impersonation focus	Familiarity of participants with projected content	Embodiment
Zero-quotative	Mimetic visitations into story-worlds already set up	Characters and event structure of the narrative are familiar to all co-	Low

		participants → lesser degree of framing required	
<i>Go</i>	Furthering of story; local management of multiple voices	The story has not been elaborated upon, not all voices have been introduced → higher degree of framing required	Medium
<i>Be like</i>	Staging of (often bodily) actions that display affective and evaluative stance; local management of multiple voices	Affective stance transmitted through the impersonation; not all characters introduced in preface → higher degree of framing required	High

These functions largely parallel those identified for mirror impersonations. However, differences are observed with regard to the focus of the impersonation. Whether or not they can be attributed to frame integration complexity will be discussed once the quotative system in which double-scope impersonations are embedded has been described.

5.2.1.3. Quotatives in double-scope impersonations

The double-scope impersonations in the NMTB corpus are either introduced through zero-quotatives or *go*. No variants of *be like* occur, which is surprising, as Golato (2000) has found *be like this* to introduce ‘quotes’ featuring non-human entities. It does, however, appear as if the function attributed by Golato (2000) to *be like this* can be fulfilled by zero-quotatives, given that the preface outlines characters and event structure. This becomes evident in “Holiday horse”, where NF impersonates a confused horse.

NMTB SE24E12_1_13:53 (Holiday horse)

```

01  NF:  my i thought that was you staring
02      blankly at the screen like a !HOR!se
03      that's been asked to book a HOLIDAY (h)
04  AU:  @@@
05  NF:  ((3 secs of mimicking horse;
        grimacing))

```



As in mirror and single-scope impersonations using zero-quotatives, the zero-quotative in double-scope impersonations requires a clear setting of the scene during the preface sequence. The impersonation itself is merely a mimetic visit, a ‘sneak-

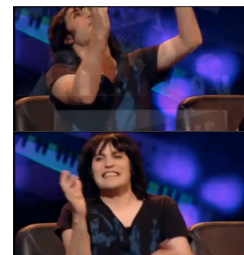
preview’ into the diegetically set-up story space. In “Holiday horse”, however, it has another function thus far only found in single-scope impersonations which are introduced through *be like*; in effect, the performance illustrates the affective stance of the impersonated character, and is highly embodied. This also applies to “Babybel”, where the impersonation accompanies the telling of a story, providing brief illustrations of both bodily action and the affective stance of the impersonated as the story unfolds. These illustrations are first and foremost gestural, with only a brief verbal element towards the end of the story. It is thus that such impersonations ‘bring before the eyes’ (cf. Aristotle, *Rhetorics*) a diegetically set-up narrative. They do not themselves contribute to narrative development, but merely demonstrate it visually through a high degree of embodiment and via sound bites which emphasise affective stance.

NMTB SE23E03_1_05:21 (Babybel)

01 NF: i like to imagine a babybel
 02 walking through a forest
 03 and all of a sudden
 04 getting pulled up into the air (-)

05 AU: [@@@
 NF: (((1 sec embodied impersonation of
 trapped babybel cheese))

06 NF: his little wax feet
 07 kicking through the netting,
 08 <@ ↑help me,
 ((2 sec embodied impersonation))



The ‘bringing-before-the-eyes’ quality of impersonation introduced by means of zero-quotative is, in my corpus, antithetical to those preceded by *go*. In these cases, the impersonation itself is central to the unfolding of the narrative. The narrative has not fully been set out diegetically, but is further driven by the characters’ actions. Consider the following sequence, which provides yet another instance of equine anthropomorphism.

NMTB SE24E02_1_03:15 (Displeased horse)

01 NF: see that horse's body language? (.)
 02 that horse is going,
 03 <@ get me out of this video. (-)
 04 AU: @@
 05 NF: this is the !WOR!st music EVer. (.)
 06 AU: @
 07 NF: i don't care how many carrots and sugar
 08 you gi(h)ve me(h); @> (-)
 09 AU: @@
 10 NF: you can lead a horse to a pop video
 11 but you can't ma(h)ke it like(h) it.

In “Displeased horse”, the voice of the horse contributes to the unfolding of the narrative. The diegetic preface does not provide any information as to the actions or attitudes that the horse might be offering. In contrast, double-scope impersonations introduced with zero-quotatives, such as “The Metal Detector”, take place in a fully established world. The metal detector’s problems and actions are established in the diegetic preface; all the impersonation adds is the Aristotelian “bringing-before-the-eyes” of an already established story. Impersonations introduced through *go*, on the other hand, are of key importance to narrative development.

As “Displeased horse” also illustrates, *go* is used for distinctly vocal and verbal elements rather than visual aspects. While zero-quotatives may also be used to introduce solely visual representations in double-scope impersonations, *go* seems to demand a sonic, if not verbal mimetic approximation.

Altogether, then, the use of quotatives in double-scope impersonations seems to be determined by

- a) the degree to which the story space has already been established in the preface, and
- b) the semiotic resources used in the impersonation.

Table 6 provides an overview of the functions identified for zero-quotatives and the quotative *go* in double-scope blends.

Table 6: Quotative functions in double-scope impersonations

	Projected impersonation focus	Familiarity of participants with projected content	Embodiment
Zero-quotative	Mimetic visits to story-worlds already set up	Characters and event structure of the narrative are familiar to all co-participants → lesser degree of framing required	High
<i>Go</i>	Furthering of story; mainly sonic / verbal	The story has not been elaborated upon, not all voices have been introduced → higher degree of framing required	Low to Medium

The functions therefore largely parallel the ones identified for single-scope and, to a lesser extent, mirror impersonations. The main difference here is the degree of embodiment for zero-quotatives, which was rather low for mirror and single-scope impersonations and was proven to be high in double-scopes. I will further discuss this point in the following section.

5.2.1.4. Discussion of findings

The high frequency of zero-quotatives and the low occurrence of *be like* in the NMTB corpus was unexpected considering the findings derived from previous research on *oratio recta*. Yet upon closer inspection it became evident that these “imitation clauses” (Vandelanotte & Davidse, 2009) fulfil clearly delineated functions and, in tight coordination, interplay with other phenomena, such as postural orientation and preface formats, as a means of building action. Quotatives contribute to signalling “This is what I am about to do”. These results further support Golato’s (2002) assessment of the context-sensitivity of quotatives. Their use is rule-governed and dependent on the action underway.

What is here referred to as ‘impersonation’ appears to in fact consist of a wide variety of actions used to achieve a number of goals. Through impersonations, participants may, for example, ascribe unfavourable motives or thoughts to a co-participant, restage past discourse, or ‘bring-before-the-eyes’ of co-participants the best-of instances of a narrative. For building these actions, impersonators have to first of all locally manage the voices contributing to the narrative, and secondly establish the epistemic ground upon which is built the greater narrative in which the impersonation is set. To this end, quotatives such as *go* and *be like* provide a means of

locally managing voices and the narrative unfolding of a story. Conversely, impersonations that attribute and reveal tacit motives or identities to a co-participant recruit zero-quotatives. However, quotatives are also orchestrated in conjunction with the semiotic resources which a given action employs. When visually illustrating a story through one's body, gestures and grimaces, *be like* or zero-quotative might be used. On the other hand, when the action is , intended to make a narrative more vivid through the staging of sonic or verbal aspects, *go* is the quotative of choice.

It thus appears that the use of quotatives is not dependent on frame-integration complexity, but more on the actions currently underway; as such, it parallels the strategies identified for diegetic prefaces. Some of these actions are only found in one blending type (e.g. displaying a co-participant's tacit motives) whilst others are found across blending types, such as mimetic visits to a story-world which had already been diegetically set-up in a preface. With regard to quotatives and prefaces, then, it is the actions that determine their selection, not frame-integration complexity.

5.2.2. Embodied demarcation markers: Gaze and postural orientation

Gaze, turning one's face and postural orientation all play a major role in organising interaction and participation frameworks (cf. Goodwin, 1981; Goodwin, 2003a; Kangasharju, 1996; Kendon, 1990). Depending on the situated activity in which it is embedded, gaze and body positioning can in this regard be employed by interactants to publicly indicate the focus of their orientation. Participants can, for example, position themselves as hearers of talk by orienting their bodies towards the speaker's. In this context, Goodwin (1981) has found that speakers will discontinue talk should hearers not visibly orient to them. The importance of bodily and visually orienting to communicative partners becomes evident in the NMTB corpus as well, when in "Listening", guest panellist NH explicitly addresses the host SA's visual disengagement with his story, questioning whether he was listening at all.

NMTB SE20E05_1_03:25 (Listening)

01 NH: uh (.) SEAn from five went to my
school.



- 02 FF: ↑oh
((SA heard shuffling through papers))
- 03 NH: WEnt to my school;
- 04 SA: so(rry)?
- 04 NH: a- are you LIStening?
- 05 SA: i'm [listening (--) so:rry
- 06 AU: [@@
- 07 NH: not sure you we=ere



This sequence illustrates the extent to which bodily orientation and gaze are central to the interactional organisation: NH interrupts his story-telling when he notices that one of his recipients, SA, is not posturally aligning himself as the recipient of the story. This goes to show that the discontinuation of bodily orientation is an accountable action and is crucial to the organisation of the participant framework.

Secondly, postural orientation and gaze can be employed in turn-taking to select addressees in interaction: I will normally turn to and, more importantly, look at the person I wish to address. In connection with this, turning away from an interactant can be employed to show distance and disengagement. Gaze, face, and body positioning are therefore crucial in displaying the focus of attention in interaction.

Impersonations exhibit unique features when it comes to gaze, face and body positioning. Previous research (e.g. Goodwin, 1984; Lee, Neidle, MacLaughlin, Bahan, & Kegl, 1997; McClave, 2000; Niemelä, 2010; Sidnell, 2006) has shown that

in the situated activity system ‘enactment’, gaze and postural alignment with co-participants is discontinued. During the enactment, the impersonator discontinues eye contact with co-participants, and further signals that s/he is oriented to another discourse scenario by orienting her/his body away from co-participants. Bodily and visual orientation with the main co-participants in interaction is resumed once the enactment is finished. Gaze and postural orientation have thus been identified as a key signalling device in parsing the diegetic from the mimetic.

While these findings largely concur with my data set, they do not, however, apply to all instances, as gaze, and more notably face and body positioning, are not presented as being fully discontinued in one type of impersonation. Mirror blending impersonations that ascribe an identity to a co-participant are, in the interaction-as-broadcast, presented as being delivered while posturally orienting towards the person that the impersonation is being assigned to. Furthermore, differences can be observed regarding the degree to which gaze and body positioning is discontinued; it seems that impersonations which more aggressively ascribe deviant behaviour to co-participants (i.e. face-threatening acts) can be delivered whilst gazing directly at the impersonated, as will be demonstrated below.

Video-based analysis of postural orientation: preliminary remarks

Analysing gaze and body position based on video, and more importantly, edited televised data runs the risk of being criticised for two reasons. Firstly, in all video data it is difficult to fully determine exactly what a person is actually looking at. Secondly, these visual elements are among the most notably edited in televised data: the final product that is aired, the interaction-as-broadcast, presents interaction not from one static angle, but is edited in such a way that shots are shown from a number of perspectives and distances. The interaction-as-broadcast appears as a unified whole. With this variety of shots ranging from close-ups to long shots, though, how is the effect of smooth interaction with people visually engaging with one another achieved? The “visual grammar” (Holland, 2000) of filming practices employs two key techniques. First of all, the *180-degree rule* holds that all people shown have to maintain screen directions, so as to avoid confusion as to their location in the set-up (similar to football games, which are broadcast from one side of the pitch only, as confusion would otherwise arise regarding which goal a team is attacking). NMTB employs three camera positions: long shots and medium long shots show the panel set-up from

the perspective of the studio audience (the studio audience is never visible, nor are the backs of the panellists), while ensuing close shots of teams and their members in interaction are shot in *shot, reverse-shot* structure. This second technique, whereby people in dialogue are filmed individually with one side facing right and the other facing left, allows the illusion of people interacting with each other rather than with the audience. When dialogues between the different teams or the host and the teams are presented in NMTB, the interactants are shown from the perspective of their respective interactants (known as an ‘over the shoulder’ shot). As this is standard video practice, the allusion of people visually engaging with each other is created when their gaze is level. The only time that people are filmed from a frontal perspective, consisting of gazing directly into the camera, is during sequences in which the host addresses the television audience rather than the panellists alone and during impersonation sequences (see below). The gaze direction of panellists is thus presented to the television audience as a means of visually engaging with another participant in interaction whilst employing the *180-degree rule* and *shot, reverse-shot* structure. In consequence, the gaze direction studied here is gaze as presented and constructed in the interaction-as-broadcast.

5.2.2.1. Gaze and postural positioning in mirror impersonations

In all instances of mirror impersonations in my data set, the postural and facial orientation of a speaker is not shifted consistently throughout the enactment to signal a change in *footing* (Goffman, 1981) or speaker roles, as the general gaze direction, level of gaze, and body positioning is maintained with only slight gaze shifts upwards or downwards and the reclining of upper bodies. There are even instances in which gaze is only interrupted at the beginning, but resumed during the impersonation. Consider “Narcissistic Barrowman”.

NMTB JB_pt2_02:35 (Narcissistic Barrowman)

01 SA: i uhm I’m putting my foot down.



02 YOU cheated there; bill bailey



03 JB: <<ff> ↑I didn't CHEAt,>



04 SA: not !YOU!; BArowman



05 JB: [right,



06 AU: [@@

07 SA: ↑ALways about you isn't it?



08 AU: [@@

09 JB: [<<smiley voice> Yea~eah>



10 SA: <@ <<f> I'll be on MAria (.)

11 I'll be on torchwood

12 I'll be on any !BLOOD!y show

13 that will have me> (1.1)

14 JB: @@@



15 SA: E(h)ven BU(h)zzcocks. @>



In this sample, SA impersonates JB, playing on the latter's alleged proclivity to appear on British TV shows. As is shown in figure 23, SA starts off his impersonation by reclining in his chair and audibly breathing in after the preface; his general postural orientation to his main addressee, JB, is nevertheless maintained. At the commencement of the enactment, SA briefly gazes upwards and to his left, but soon

levels his gaze directly at JB. By the end of the first TCU of the enactment, namely the prosodically marked **M**aria (line 10), he again fully orients both gaze and upper body towards the impersonated JB. His gaze and body retain the same orientation throughout the remaining two TCUs of the performance, with the sole exception of a brief lateral head shake towards his right on the accentuated **B**LOODY (line 12). McClave (2000) and others (e.g. Goodwin, 1980) have found that lateral movements of the head may be used as intensifiers in evaluations. As this head movement is synchronised with a prosodically marked word, it may therefore be used to further emphasise the word rather than act as a discontinuation of gaze. At the end of his performance, SA finally terminates his gaze, looking down at his desk. It thus appears as though in aggressive enactments that attribute a deviant identity to people present, the gazing practices differ. While the onset of the enactment is marked by discontinuing gaze, large chunks of the impersonation may be delivered whilst gazing and orienting towards the recipient. The end of the impersonation is again marked by the cessation of one's gaze. It is consequently different from 'ordinary' impersonations, where shared gaze is resumed once the impersonation is complete (as, for example, Niemelä (2010) contends).



Figure 23: Gaze and postural orientation in “Narcissistic Barrowman”

Interestingly enough, this ‘deviant’ enactment behaviour is counter to the practices employed in post-production, where impersonators are shown with frontal close-ups

during enactments to further add to the illusion that impersonators are oriented to to another discourse scenario rather than to their co-participants (which is usually shown using the *shot, reverse-shot* technique). In this example, however, SA remains oriented to his right whilst being presented from the front. This indicates that producers seem to place distinct importance on the discontinuation of body posture and gaze in their re-shaping of the interaction – more so than the original participants, who seem to accept orientation towards the impersonated during enactments.

In less aggressive or face-threatening mirror impersonations, however, gaze tends to be discontinued throughout the impersonation. Nonetheless, postural orientation towards the enacted is continued, contrary to claims made in previous research. Consider again the “Sports Commentator I” sequence discussed in chapter 4.1. In this sample, the impersonator, PJ, does an impression of a member of the other team. While PJ gazes at the host during the preface to his performance (i.e. to his left), he turns his gaze further to his right on the onset of the impersonation (cf. figure 24). As he is now sitting straight ahead, he is posturally positioned in a manner directly facing the ‘butt’ of his impersonation. However, his gaze direction in the impersonation-as-broadcast is presented as going towards his right rather than straight ahead. This means that he does not look at her during the impersonation, but is bodily oriented towards her.



Figure 24: Gaze and postural orientation in “The Sports Commentator I”

It may therefore be concluded that in mirror impersonations, gaze is mostly discontinued while a general bodily orientation to the person that is being impersonated is maintained – this stands in stark contrast to previous findings on this subject. Gaze may be resumed in especially aggressive impersonations once the voice change has been established. This type of enactment is discontinued by terminating one’s gaze, whereas less aggressive enactments are ended by re-establishing shared gaze again. The two subtypes notwithstanding, it appears that in mirror impersonations, participants show a strong orientation to the current discourse setting by maintaining a postural orientation to the co-participants with whom they were engaged during the preface.

5.2.2.2. Gaze and postural positioning in single-scope impersonations

While impersonators largely retain a postural impersonation towards the impersonated during impersonations that enact a person present in the discourse scenario, single-scope impersonations in which the impersonated inhabits a different discourse scenario show greater discontinuation of bodily and facial orientation. However, such instances are not always as clear-cut as standard literature would lead one to believe. While gaze is always discontinued at the start of an impersonation, it might –albeit briefly – be redirected towards the main communicative partner during the commencement of the Turn Constructional Units (TCUs) contributing to an enactment. Consider, for example, the “Jason Orange” sequence.

NMTB S20E04_1_07:05 (Jason Orange)⁷²

01 PJ: well one might imagine



02 that madonna isn’t really qualified to
03 solve such matters in this country.



⁷² The transcript contains IPA symbols (in square brackets) where marked deviations from the standard used by PJ occur. Jason Orange, a member of the boy band Take That, is from Manchester, and PJ mocks a northern accent in this impersonation sequence.

04 SA: but jason or˘ange;



05 PJ: he IS;



06 but then again you see he'd go in and go,

07 <@ ri::ght (-)

08 AU: @

09 PJ: nuclear power and tha';

10 AU: @

11 like-

12 S[u]m of it is good li::ke

13 @

14 Wh[o]t makes electricit[ə]

15 AU: @

16 PJ: B[u]t' (-)

17 AU: @@

18 PJ: S[u]m of it 's b[a]::d

19 AU: @@

20 PJ: like, (.) i wanna stop the bad stuff (-)

21 keep the good. @>

22 AU: @@@

As figure 25 shows, PJ is shown as being strongly oriented to SA during the interaction preceding the impersonation sequence. He turns further away from SA during the preface to his impersonation (albeit not crossing the 90-degrees axis), but glances back at SA three times during the first three TCUs of his performance: briefly on the lengthened vowel of *ri::ght* (line 07), then again at the onset of the second TCU on *nuclear* (line 09), and for the entire third TCU, namely *sum of it is good li::ke* (line 11). As his turning of the head is not synchronised with emphasised syllables, this glancing towards his left (and hence original communicative partner) could constitute a verification for acknowledgment or tokens of appreciation. However, this latter point unfortunately cannot be confirmed by my data set, as the interaction-as-broadcast does not show potential backchannelling signals by SA in this sequence, nor is this shown in some of the other instances (e.g. “Great Value”) in which this brief glance by the impersonator back to the host can be observed. There

is, though, evidence that such fleeting looks back to the host (who perceptibly holds most power in the studio set-up) are indeed backchannelling requests, as the interaction-as-broadcast shows a laughing host following such glances in other sequences (cf. analysis of “Metal Detector” below). This is also confirmed by previous research on head movements, which has indicated that brief lateral head movements may be used as nonverbal backchannelling requests (cf. McClave, 2000).



Figure 25: Bodily orientation in “Jason Orange”

The remaining five TCUs of the impersonation performance are delivered whilst fully oriented on the 90-degrees axis and hence away from SA. By the end of the enactment, PJ is reclining in his chair and crossing his arms, with his head even turned slightly more to his right than to his left. Over the course of the impersonation, then, he has continuously moved further away from his original interactional partner; from keeping a general orientation and even glancing back to SA during the first few turns, he crosses the 90-degrees axis and orienting more to his right. The longer PJ inhabits the role and the discourse world he has set up, the more he orients away from the original interaction.

For briefer single-scope impersonations, however, the impersonator does not *posturally* orient as far away from the original interactant, nor does s/he utilise gaze during the impersonation to check for tokens of acknowledgment. On the contrary,

gaze and facial alignment is ostensibly discontinued in short impersonation sequences, which is entirely different to topical talk. In “Northern Mac”, host LM mocks guest DV for the second time for her northern accent, and impersonates a fast food outlet customer from the north of England.

S24E08_2_11:18 (Northern Mac)

01 LM: <<singing> return of the mac>



02 PJ: <<singing> o=o u=u>

03 DV: I GEt uh burger uh mac?flurr?ic,



04 LM: MAc burger, you think'

05 ↑<<f> It's not an advert for big
mac,>



06 return of the mac,

07 <@oh I'll have fries with that>=



08 [return of the mac (.)

AU: [@@@

09 LM: we don't do ADverts

10 <<ff> we're trying to BREAK

11 this northern stereotype>

12 return of the mac,

13 <@ oh ri:te

14 i'll have cheeseburger and [fries

AU: [@@@



15 LM: and apple pie as well please @>

This sequence illustrates how in shorter impersonations that only last one or two turns, head movement and gaze play a key role in parsing the diegetic from the

mimetic. While LM fully orients to DV prior to the enactment as well as during the diegetic interlude, he aligns his upper body and gaze more at a ninety degree angle during the impersonations, gazing and gesturing straight ahead towards an implied conversational partner (see figure 26). As this is a brief and rapid sequence, he jolts his body towards the right and centre in order to signal the moment when he assumes the role of the stereotypical northern consumer of a fast food outlet and when he returns to addressing DV in topical talk. He does not, however, cross the 90-degree axis with his upper body during his impersonation, but retains a general orientation to DV by keeping his left shoulder directed forwards. This is in line with Goodwin's (2003b) findings, which suggest that the upper parts of the body are more flexible in shifting alignment, while the lower limbs are used to indicate general alignment. Brief engagement and disengagement is thus displayed by turning the head or shoulders, while general alignment is signalled through the posture of the lower body, such as direction of the legs when seated. As the "Northern Mac" sequence is quite short, the disengagement probably did not last long enough for the orientation of the lower limbs to change fully. In contrast, "Jason Orange" (cf. figure 25) constitutes a longer break of topical talk, and may thus lead to a further orientation away from the original interactants.



Figure 26: Postural orientation in "Northern Mac"

Although gaze may be briefly redirected towards previous interactants at the commencement of longer impersonations (presumably to check for tokens of acknowledgment that allow for further continuation of the impersonation), postural

alignment and discontinuing gaze appear to be crucial to single-scope impersonations. Other than in mirror impersonations, the body, face and gaze are aligned more to the 90-degrees axis (‘straight ahead’) than towards the original interactants. Whether this does indeed constitute a clear indicator for the division between mirror and single-scope impersonations will be discussed below.

5.2.2.3. Gaze and postural positioning in double-scope impersonations

The importance of discontinuing postural orientation and gaze with previous interactants is also evident in double-scope impersonations, in which impersonators assume the role of non-humans in other discourse scenarios. Consider “Metal Detector” once more. As shown in figure 27, NF orients to his main co-participants, host SA (to his right) and guest KS (to his left) during the co-constructed prefacing sequence, but then orients straight ahead from the onset of his impersonation of a metal detector.



Figure 27: Postural orientation in “Metal Detector”

The first question that provides the relevance for his later impersonation is directed at SA, and NF consequently bodily orients to the host. However, when KS further

elaborates upon the ‘problem’, NF turns to her, aligning himself as the recipient of her talk. Yet NF delivers his next turn oriented away from her, thus signalling that he is speaking in another voice. This is accepted by his co-participants, as they arrange themselves as recipients by showing postural orientation to NF and laughing at his performance without attempting to interrupt. Once NF’s impersonation has been established, his gaze can be interpreted as being directed by the metal detector, addressing co-imagined figures in space. This is, for example, evident in NF’s brief alteration in bodily posture to reflect an upward direction, presumably to the person using the metal detector, before delivering the apologetic *I can’t help it*. He, too, briefly gazes at host SA before offering the final TCU of his impersonation. Is this brief lateral head movement a request for backchannelling, or is it the metal detector addressing a general audience? As the next shot shows SA laughing, it can be inferred that in the interaction-as-broadcast, NF’s brief gaze to the right is indeed a way to verify recipient tokens. The interaction-as-broadcast consequently aligns host SA as the main recipient of the impersonation, i.e. the recipient oriented to by NF during his performance; this is despite the fact that the other co-participants also align themselves as recipients of his enactment via their strong postural orientation towards NF. As in single-scope impersonations, then, backchannelling requests are possible in the interval between an impersonation’s TCUs. Furthermore, double-scope impersonations also require an otherwise steady discontinuation of gaze and postural orientation.

5.2.2.4. Discussion

Overall, then, is there evidence for the usage of separate techniques in single and double-scope blends as there are in mirror and single-scope blends? When considering an example that features mirror, single-scope and double-scope impersonations in close succession, double-scope impersonations at first glance do indeed seem to require a clearer discontinuation of postural orientation. In “Stevenson’s rabbit”, impersonator PJ ascribes a tacit Victorian identity to a co-participant (i.e. mirror-blend; cf. discussion of zero-quotative mirror impersonations in chapter 5.2.1.) prior to enacting a ‘steam-powered dildo’. This double-scope impersonation of a machine is then followed by a single-scope impersonation performed by another co-participant, RI. Let us now consider this complex sequence in more detail.

NMTB SE19E05_1_08:11 (Stevenson's rabbit)⁷³

- 01 DB: do you think it is uh (.)
- 02 a little (.) steam-powered dildo
- 03 with little (.) chocolate (.)
- 04 uhm toffee apples on the side?
- 05 AU: @@@
- 06 PJ: a steam-powered DILdo?
- 07 AU: @@
- 08 DB: yeah
- 09 PJ: <@ <<f>oh if i'm pleasuring myself
- 10 i have to do it in the victorian
- style; >
- (-)
- 11 AU: @@
- 12 PJ: <<f> where's my coat james?> @>(-)
- 13 AU: @
- 14 PJ: stevenson's rabbit;
- 15 <<steam sounds in three intervals>>
- (--)
- 16 <<whistle sound, twice>> (--)
- 17 AU: x@x@x
- 18 RI: <@ (isenba:rt) what's under your hat?
- @>
- 19 AU: @@
- 20 RI: <@ i have quite a surprise for you
- mylady @>



The three impersonations are made relevant in a guessing game, during which one team member, DB, suggests to his fellow team members, PJ and RI, that the futuristic chocolate fountain in a video clip might actually be a “steam-powered dildo”. This is latched upon by PJ who, while strongly orienting to DB, repeats the key term before delivering his first impersonation, through which he describes a ‘Victorian’ attitude to DB. He briefly glances away at the start of the impersonation, but remains oriented to

⁷³ The transcript only shows frame grabs when they aid understanding of the sequence.

DB throughout most of the performance, and often looks at him directly. However, once he begins to impersonate the machine, he orients straight ahead, lowers his head and produces sound effects commonly associated with steam engines; these sounds are accompanied by gestures and body movements. He is not once presented as orienting to his co-participants during this sequence, as figure 28 shows.



Figure 28: Postural orientation in “Steam-Powered Dildo”

RI’s impersonation of yet another Victorian person (this time single-scope) follows PJ’s enactment, illustrating again that a subsequent joke is a relevant next action to a prior joke. However, RI shows crucial differences as regards bodily orientation. Whereas PJ, who is seated in the middle, positions his body to face straight ahead in his enactment of a machine, RI’s single-scope impersonation of a human is presented as keeping a postural orientation to his team members on his left. This postural orientation is much stronger at the beginning of the impersonation, and appears somewhat more discontinued towards the second TCU of his enactment. He gazes upwards and to his right, but from shoulder-level downwards maintains a general

orientation to his team (i.e. he does not cross the 90-degrees axis). This corresponds to the conclusion made above: discontinuation of postural orientation is gradual and progresses in impersonations that require ‘legitimacy’ from other co-participants. The impersonation is presented as something that is relevant to the joint action underway, as the impersonator aligns himself as a member of the interaction even during an impersonation.

To conclude, while PJ is oriented to his main co-participant DB during the impersonation of the Victorian human, he discontinues his gaze and is oriented at a ninety degree angle throughout his performance as the machine. RI’s single-scope impersonation is then delivered whilst looking away from his co-participants, although he maintains a general postural orientation towards them (i.e. he does not cross the 90-degree axis); this is consistent with previous examples of single-scope blending. Despite this, “Stevenson’s Rabbit” hardly proves that double-scope impersonations require more discontinuation than single-scope impersonations, as numerous impersonations recruiting novel discourse scenarios are delivered whilst continuously orienting to a co-participant. How, then, can these instances be accounted for?

Problem cases

A number of otherwise clearly single-scope impersonations are delivered with a particular gaze and maintaining postural orientation to a co-participant. In “Alan Bennett Monologue”, host FB does a single-scope impression of what he frames as a prototypical Alan Bennett monologue; he does so whilst being shown as clearly orienting to CV, who had the previous turn.

NMTB SE23E11_3_00:39 (Alan Bennett Monologue)

01 CV: YEAh but I DIId see it then cos the children
 02 were a bit younger then;
 03 so we used to STAY in on a saturday night
 04 <<p> and,>
 05 NF: but you’re an AMAzing mum
 06 AU: @@
 07 NF: and made them watch the [ex(h)fa(h)cto(h)r;
 08 AU: [<<p>@@>
 09 CV: NO=NO=NO
 10 cos they WANTED to watch the ex factor at that age.

- 11 NF: <pp> alright>=
 12 FB: THis is like a alan BENnett monologue now (-)
 13 AU: @@@
 14 FB: <@ i remember we had (raVEEtis) [that night; (--)
 15 AU: [@@@
 16 FB: [(there were)] the EXfactor
 17 AU: [@@
 18 FB: i went for the custard creams
 19 there was NO custard creams;
 20 HAR[vey; (.)
 21 AU: [@@
 22 FB: HARvey had taken the LAsT custard cream;
 23 FB: and in MAny ways
 24 THAT summed up OUr relationship. @>
 25 AU: @x@x@x



Figure 29: Postural orientation in “Alan Bennett Monologue”

As figure 29 illustrates above, FB briefly gazes to his right and away from his interactants at the beginning of his impersonation, but keeps looking back to his left and therefore back to his original interactants during the enactment sequence. He turns his gaze away very briefly at the onset of each TCU, but returns it towards the

end of each section. In addition to this, the interaction-as-broadcast presents CV as the addressee of this single-scope interaction, as she is shown laughing in close-up after the preface and at the end of the impersonation sequence. This parallels the strategies identified for mirror-blends.

How can this be accounted for? When examining the pragmatics of these performances, it transpires that in all these samples, the impersonation is employed to mock a co-participant for allegedly deviant behaviour. By maintaining a postural orientation towards and gazing at the interactant, impersonations can be used to effectively state ‘this is you’ – no matter whether they are mirror or single-scope (cf. assessment of mirror blends using zero-quotative).

In a similar vein, gazing at co-participants whilst embodying another role is also possible in joint, multi-party role-play, where the impersonator aligns himself as one character in the story space and the addressee another. This is, for example, the case in “Easy Robin”, where DM, PJ and BB address JA, who had previously been introduced as the actor playing Robin Hood in a recent BBC dramatisation, as “Robin”, assuming unnamed characters of the Robin Hood story space.

NMTB SE20E05_2_04:21 (Easy Robin)

```

01 DM: <@ EASy Robin leave it [<<f> LEAve it- >
02 JA:                                [<<p> a'right.>
03 AU:                                [@@
04 JA: I know I know.
05 PJ: <<f> we've A:ll had some meat>=
06 DM: [<<f> don't WOrk it robin,> (-)
07     [@@@
08 BB: NOt so MErry now are ya? @>
09 AU: @@@xxx

```

As shown in figure 30, discontinuation of gaze is not observed at the onset of the characters' role-swaps, showing that the enactors fully orient to the story space in a manner akin to the theatre. It is only BB's concluding contribution to the sequence that shows a clear discontinuation of gaze. However, it cannot be inferred as to whether BB is speaking as a character of the story-world set up in previous interactions or is actually speaking as 'himself', orienting again to the previous discourse space and signalling the break with on-going 'fictive' talk by discontinuing

gaze. Still the fact remains that DM and PJ assign a role to a co-participant, and address him as such in this story-world throughout their enactment. Their performances are characteristic of role-play and theatre-like, as they fully orient to the discourse-world set-up (and consequently the characters in said discourse-world) rather than topical turn-by-turn talk. Similar to the previous case, these samples attribute an identity to a character present through postural orientation, gaze, gesture and other strategies of address (such as naming).

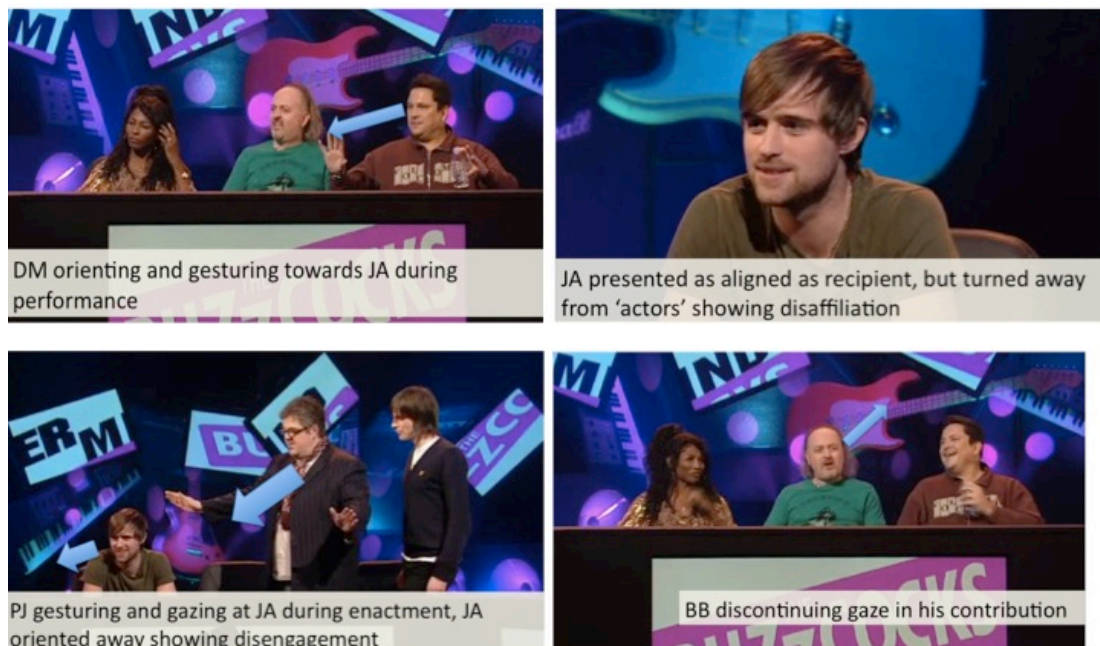


Figure 30: Postural alignment in “Easy Robin”

5.2.3. Conclusion

The analyses in this section confirm the hypothesis outlined in chapter 4. It is not the blending typology as proposed by Fauconnier & Turner (2002) that determines the quotative, gaze and postural orientation in impersonation blends, but the pragmatic dimension of a given impersonation. In impersonations that arise from spontaneous turn-by-turn talk and directly assign an identity to the co-participant currently being engaged, postural orientation is not discontinued, and gaze may resume during the performance. In addition, these impersonations recruit zero-quotatives. Impersonations that suggest potential scenarios in possible worlds for illustration purposes are, however, delivered while discontinuing gaze, regardless of how incompatible the input frames are, or how much structure from each input is

incorporated into the blended space. Their use of quotatives is dependent on a) the local management of talk (i.e. whether or not the story space has been fully elaborated upon or not), and b) the semiotic resources recruited in the performance (i.e. *be like* for solely visual or non-lexicalised sounds; *go* for largely verbal performances that themselves further narrative development). The demarcation markers employed in interactional impersonations thus depend on the action currently underway rather than on frame-integration complexity. Table 7 summarises the actions built through impersonations and the methods employed to achieve them.

Table 7: Overview of actions built through demarcation markers (quotatives, gaze, postural orientation)

Action	Quotative	Gaze	Postural orientation
Doing ‘this is you’	Ø	Discontinued with ‘butt’ at onset; may be taken up again during performance	Orientation to ‘butt’ maintained
Doing illustrations of spatiotemporally, modally and epistemically disjunct discourse-worlds	<i>Go</i> : impersonation furthers narrative development	Fully discontinued with co-participants	Fully discontinued with co-participants
	<i>Be like</i> : visual / onomatopoeic illustration of diegetically set-up narrative		
	Ø: illustration of diegetically set-up narrative		

Whilst sections 5.1. and 5.2. have largely focused on the introduction of impersonation sequences to topical talk, the following chapters will focus more on the internal structure of impersonation sequences. In the spotlight will be methods employed for floor management and exit devices which enable the return to topical turn-by-turn talk in order to discern whether these aspects are also related to the building of social actions that span across blending types or are rooted in frame integration complexity.

5.3. Designing performances: Floor management, chunking and listing

The previous chapters have already touched upon a central feature of impersonations: in most cases, they constitute multi-turn units (cf. Schegloff, 1996b) – or ‘discourse

units' in Houtkoop & Mazeland (1985) – and hence suppress the turn-taking model found in 'normal' turn-by-turn talk (cf. chapter 3.2.2. for an overview). Participants wishing to tell a joke, a story, or provide any longer turn thus need to a) secure the right to reserve an extended sequence of turns-at-talk through a TCU projecting subsequent TCUs (cf. chapter 5.1.), and b) project the end of their contribution.

The analyses in this section will thus focus on the question relating to how such multi-turn units are managed in the varying blending types. This will be accomplished by concentrating on the methods employed in securing further turns-at-talk during the performance of impersonation sequences. I will first focus on the structural level, and describe the ways in which the various impersonations are chunked and marked through "incompletion-signals" (Houtkoop & Mazeland, 1985:599) or "projecting components" (Selting, 2007:491) as "not-designed-to-be-ending" (Schegloff, 1996b:83). In a second step, I will closely examine the role that prosody and especially pitch play in floor management; this is due to the fact that it is "melodic phrasing" (Couper-Kuhlen, 2009:178) which is instrumental to signalling continuity and endings in longer projects (Selting, 2000, 2010).

The present section will furthermore concentrate on the ways in which the length of an impersonation performance is interactionally negotiated. Impersonation sequences that extend over a number of TCUs are, to this end, chunked such as to leave silent pauses in between the various segments of the performance. While these silences occur at what would in 'normal' conversation constitute a TRP, they are not treated as such by co-participants in longer projects (cf. Houtkoop & Mazeland, 1985). Instead, at these points co-participants in the interaction-as-broadcast signal that they are still aligned as recipients through tokens of appreciation in the form of laughter and clapping. Schegloff (1996b:96) points to the object-like nature of TRPs in general. They are thus shaped a particular way, and are subject to limitations regarding for example their temporal unfolding. This structural chunking of impersonations will be discussed with a specific focus on a) turn length and structure, b) the shape taken by such spaces at the end of the TCUs ('intervals' in my terminology) which comprise impersonations according to blending type, and c) the essential role played by prosodic design in signalling floor holding and yielding.

Prosodic analysis: Methodology

As this chapter will also investigate issues relating to prosody, a word on terminology and methodology is required.⁷⁴ While the terms intonation and prosody are sometimes taken to be synonymous (e.g. Wells, 2006), subsuming features such as pitch, stress and rhythm, I largely adhere to the terminology established in interactional linguistics. While prosody is seen as the “interplay of pitch, loudness, duration and voice quality” (Selting, 2010:5) employed as “communicative signals” (ibid), intonation principally describes the “melodic movement in pitch” (Szczepek Reed, 2006:4). As intonation (i.e. the relative changes occurring in pitch level over a particular segment of talk) centrally contributes to negotiating turn-taking (cf. Selting, 2005), the following chapters will recruit pitch analyses to draw conclusions as to the role that intonation plays in securing an extended sequence of turns-at-talk in impersonations. *Pitch* results from the “frequency of vibrations of the vocal folds” (Szczepek Reed, 2006:3), and is measured in hertz (Hz). Pitch contours (also referred to as *fundamental frequency (F₀) trace*) will be assessed by isolating the relevant TCUs and analysing the changes in fundamental frequency from pitch-peak to TCU-end with the help of the *praat* computer analysis package. These analyses serve to determine whether a declining, rising or level intonation pattern is visible, and indicates the extent to which such intonation patterns serve to achieve turn-holding or turn-yielding in multi-turn projects. Figure 31 shows one such frequency analysis of a female speaker ranging between 90 and 356 Hz.

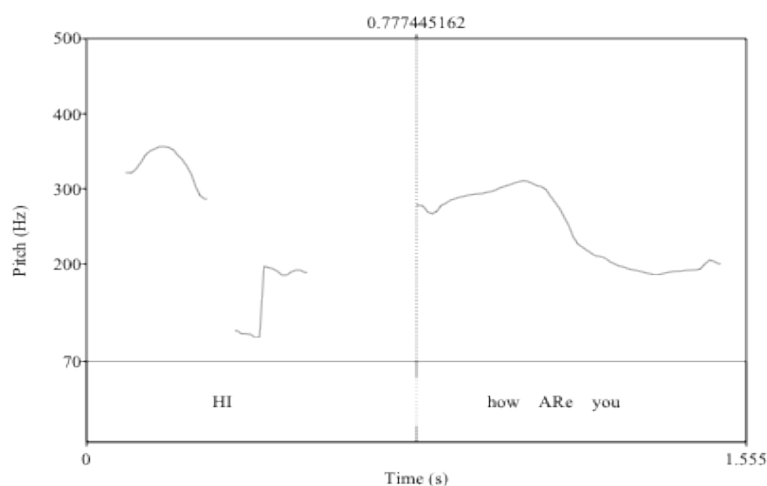


Figure 31: Intonation

⁷⁴ This project is not primarily concerned with issues in prosody, and will, for this reason, not be able to delve into the potential offered by prosodic analyses in detail.

The so-called ‘punch-line’ mode is a common technique for signalling the end of an impersonation. In order to determine the features contributing to the prosodic realisation of such punch-lines, *stress* and *accent* will be analysed. As loudness (i.e. intensity) plays a role in both realising and perceiving stress and accent⁷⁵ (ibid), changes in decibel (dB) and hertz (Hz) over the course of a TCU will be analysed, again with the help of *praat*. Acoustic waveforms (see figure 44) that illustrate loudness are further recruited to describe the negotiation of intervals (used for audience tokens of approval in the form of laughter and clapping) between TCUs contributing to an impersonation performance. Such ‘intervals’ and other pauses will, in addition, be calculated by isolating the interval space occurring between the end of a (linguistic) TCU and the beginning of the next word as a means of gaining information regarding their length and realisation.

5.3.1. Floor management and performance design in mirror impersonations

Chapter 5.1.1. illustrated the multitude of prefacing formats that mirror impersonations can assume depending on the action projected. This is also reflected in the way they are chunked. While an impersonation whose preface ends in a quotative (e.g. *be like* or *go*) or a disjunct marker (e.g. *oh*) is not followed by an interval during which co-participants align themselves as recipients of an impersonation, an impersonation prefaced by syntactically and pragmatically complete redescriptions and zero-quotatives do indeed leave such a space.

Consider the following examples. In “Banana Outfit”, the impersonation latches on to the previous TCU, and is indicated as distinct by the disjunct marker “Oh”. Similarly, after the preface recruiting the *be like* quotative, impersonator NF in “Bemused look” does not leave space for his co-participants to align themselves as the audience of his brief impersonation; rather, he appends it to the preface.

⁷⁵ *Stress* describes syllabic prominence achieved primarily through increased loudness, whereas (*pitch*) *accent* combines features of loudness and pitch (Szczepek Reed, 2006).

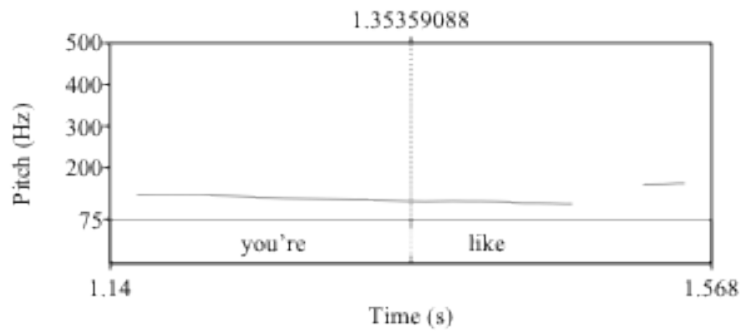


Figure 32: Pitch contour for the quotative *be like* in "Bemused look"

As figure 32 shows, the pitch contour of the quotative therefore does further project continuation and is turn-holding rather than turn-yielding via largely maintaining a mid-level plateau rather than falling.⁷⁶

There seems, however, to be a difference in pitch contour when it comes to quotatives. *Go* is usually found to have a higher rise in pitch than *be like*, as the following extract from "Flirting" (figure 33) illustrates.⁷⁷

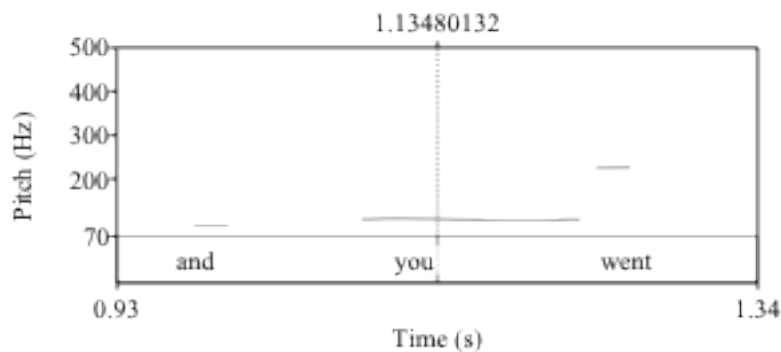


Figure 33: Pitch contour for the quotative *go* in "Flirting"

Nonetheless, both intonation contours, whether maintaining a mid-level plateau or showing a jump to a higher peak, contribute to securing the floor, and the impersonation latches directly onto foregoing talk. The same applies to impersonations that are displayed as being continuous with the previous turn and that will continue in the previous voice. They not leave an interval either; rather, they

⁷⁶ Most speakers in the corpus follow a Received Pronunciation Intonation (RPI) pattern. Speakers recruiting from other varieties showing different intonation patterns will be indicated.

⁷⁷ This difference in intonation between *be like* and *go* might be traced back to their respective etymological origin. Whereas the intonation contour found for *be like* points to its illustrating and comparative function and treats what follows more as an adverb rather than a distinct unit, *go* might function more like a traditional quotative, introducing the verbal externalisation of another person's words. Further research will be required to test this hypothesis.

adjoin themselves to prior talk. This applies to “Travel Lodge” and “Frustration”, which both claim to continue (yet accentuate) the affective stance displayed by the prior speaker.

In contrast, mirror impersonations that are prefaced through syntactically and pragmatically complete redescrptions of prior actions leave space for co-participants to align themselves as recipients of a longer unit. In “Sports Commentator I”, impersonator PJ offers a redescription of co-participant GL’s actions prior to starting his impersonation of her. This preface starts with the first syllable accentuated at a higher pitch; this is a technique which has been found to introduce “big packages” (Couper-Kuhlen, 2001, 2009). As figure 34 shows, after a sharp fall, the contour rises again towards the end of the TCU. This prefacing TCU is thus marked as incomplete and projects further elements (cf. Selting, 2010:8).

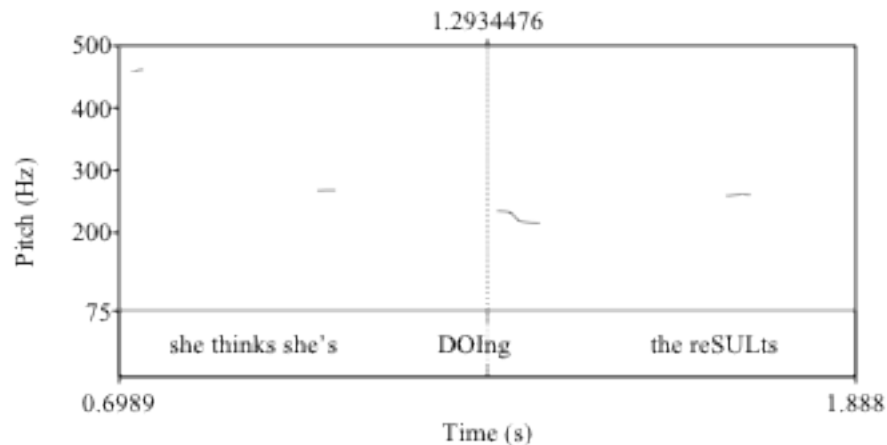


Figure 34: Rising pitch contour on "Sports Commentator I" preface

After the redescription in the preface, PJ leaves 1.8 seconds for the audience to provide tokens of appreciation and align itself as the recipient before initiating the impersonation he projected in the preface. This format is also present in “Northern lass”. In this sequence, impersonator LM leaves 1.9 seconds between his prefacing question the commencement of his zero-quotative impersonation. However, the duration of such an interval between preface and impersonation is jointly negotiated rather than pre-determined. The more approval displayed by the audience, the longer the pause may last. In “Smug Jonas”, audience laughter and clapping between the preface (lines 01-03) and the impersonation (lines 06-09) almost mutes JA’s dismissal (line 04) of the redescription proposed in impersonator SA’s preface, and is consequently ignored and sequentially deleted by his co-participants. The approval

sequence lasts a total of 3.8 seconds, 2.4 of which SA leaves for audience approval after JA's unsuccessful attempt at a rebuttal. The effect that more successful refutations of redescrptions have on the temporal limitation of intervals becomes evident in "Sports Commentator II", where the 'butt' of the joke, GL, strongly challenges the redescription proposed by impersonator RG. For this reason, RG interrupts GL's rebuttal (line 09), once more repeating his redescription, prior to latching the impersonation directly onto this repeat, with GL still attempting to gain the floor (line 13) and 'giving in' only once the impersonation has already started.

It therefore becomes evident that the property of the spaces between a prior turn and a mimetic impersonation are interactionally enacted and subject to the activity currently being negotiated. While those mirror impersonations that are prefaced through a TCU marked as syntactically, pragmatically and prosodically complete might leave approximately 2 seconds for an interval during which the audience aligns itself as the recipient of a humorous sequence, those impersonations which claim to continue a prior speaker's turn or are embedded in a TCU that cannot be projected as complete are appended directly onto diegetic talk.

Other than single and double-scope blends, mirror impersonations are rarely chunked in a way that provides further interval spaces during the performance, and only one instance of the otherwise prototypical 3+1 format (see below) is found in my corpus. The 3+1 impersonation format consists of a list of three elements (cf. Jefferson, 1990; Selting, 2007) which are delivered in close succession prior to a longer interval followed by a final element. As previous research has already unearthed, three appears to be a 'magic number' in interaction (e.g. Erickson, 1992; Jefferson, 1990; Lerner, 1994; Selting, 2007). When negotiating turn-taking, participants orient to such lists, projecting the third component as potentially turn-final. In impersonations, however, the prototypical format adds a fourth and final component to that list after a longer interval. The mirror impersonation "Narcissistic Barrowman" follows this format.

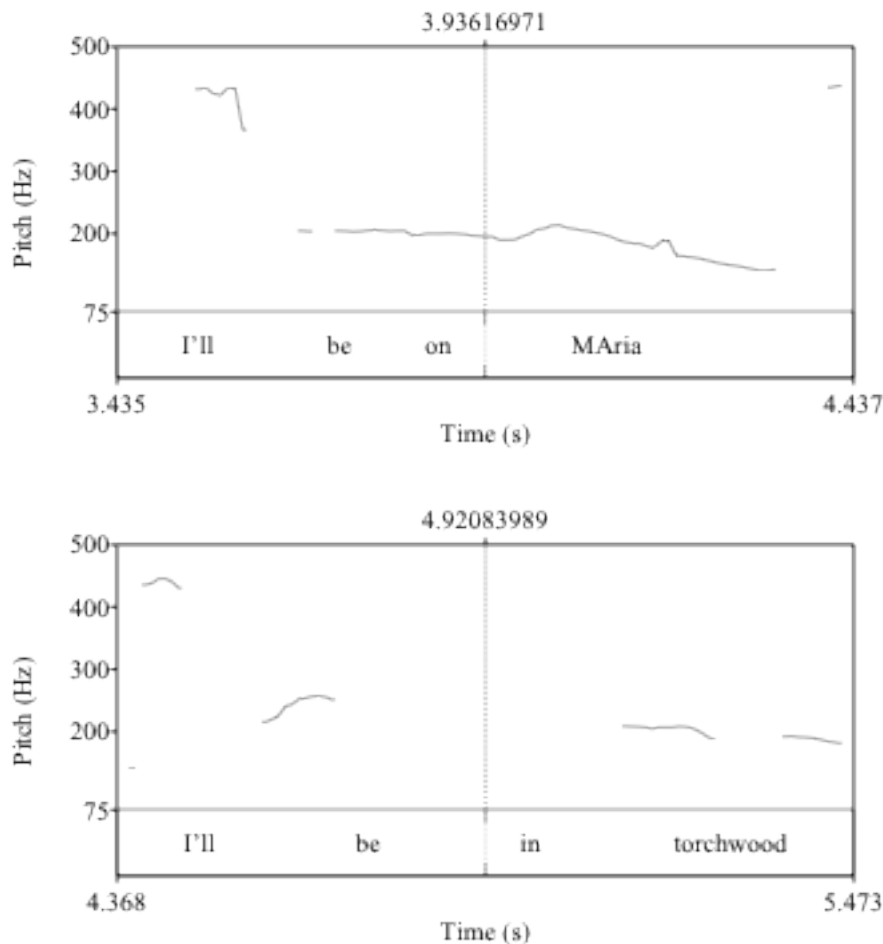
NMTB JB_2_02:35 (Narcissistic Barrowman)

10 SA: <@ <<f> I'll be on MAria (.)
 11 I'll be in torchwood
 12 I'll be on any !BLOOD!y show
 13 that will have me;>

14 AU: @@@ (1.1)

15 SA: E(h)ven BU(h)zzcocks. @>

SA's impersonation comprises a list that illustrates JB's seeming willingness to appear on British TV shows in three anaphoric blocks⁷⁸. These blocks are delivered in parallel intonation contours. They commence with high pitch on "I'll", which then drastically falls to a mid-level plateau prior to slightly falling towards the end of the TCU (see figure 35). The third unit is delivered in punch line mode (cf. chapter 5.4.2.), and is thus further marked as the climax of the impersonation.



⁷⁸ References to 'blocks' largely correspond to the *intonation unit* as described by Couper-Kuhlen (1986) and others. Intonation units are parts of speech that are produced in one coherent intonation contour (Szczepek Reed, 2006:27). Couper-Kuhlen does not consider intonation units to be on par with 'sense units', but introduces the notion of 'paratones' to describe topical organisation and coherence in spoken language. These paratones, she claims, are equivalent to "conceptual paragraphs" (1986:189f), constitute what Mukherjee (2001:41) calls "intonational paragraphs", and serve to prosodically structure longer and more topically coherent sequences of talk. As the concept has been extensively criticised as a functionally irrelevant concept that is not compatible with findings made in memory research (cf. *ibid*:42ff), I refrain from discussing impersonations in terms of paratones.

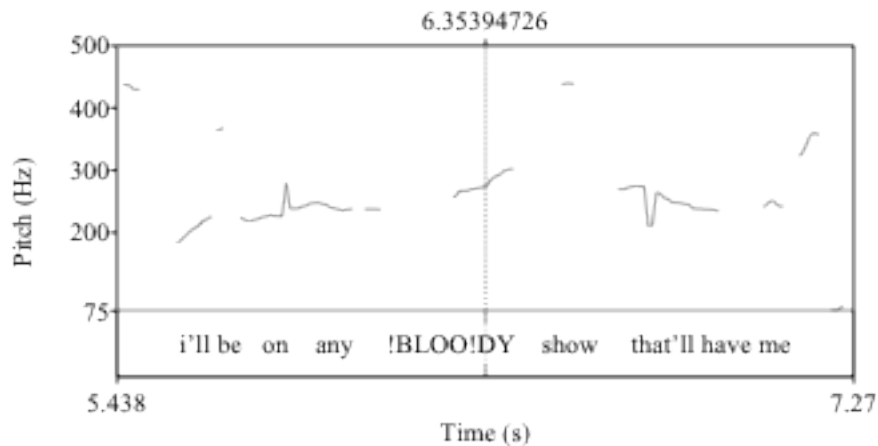


Figure 35: Pitch contour on "Narcissistic Barrowman"

The three-part list is followed by an interval of 1.1 seconds before the final element of the impersonation (depicted in figure 36) is delivered. This final element is marked as turn-final via laugh particles (cf. chapter 5.4.4.) and a relative falling, turn-final intonation contour.

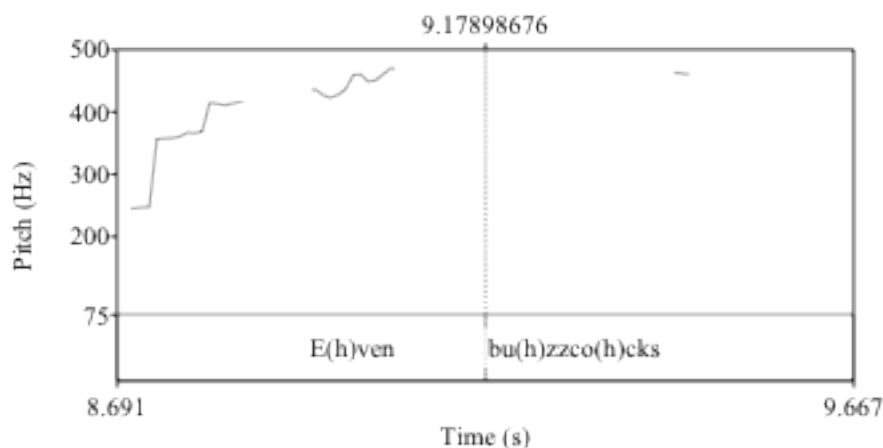


Figure 36: Pitch contour on Plus-1 in "Narcissistic Barrowman"

While such lists or other forms of chunking leave intervals in the performance, most other mirror impersonations contain a binary structure consisting of two TCUs with no prolonged interval spaces in between. This may be due to their restaging and accentuation of past dialogue, concisely pointing to perceived incongruities in prior talk. As 'normal' turn-by-turn topical talk prefers turns-at-talk to consist of one TCU only, the turns upon which the majority of these restagings are based are themselves relatively short, and, in a 1+1 format, illustrate the projected incongruity in their final element.

This binary structure is also reflected in the prosodic design of the 1+1 impersonation format. The first element in these samples is delivered in a way that ends on a mid-level plateau, while the second element ends in falling final pitch. Intonation contours therefore firstly play a role in marking continuity at the end of the first block and secondly on its completion and turn-yielding at the end of the second component. Consider again “Banana outfit”, which illustrates the 1+1 format and is visualised in figure 37.

Starting with a comparatively higher pitch, the first element of the impersonation (“nothing too embarrassing”) is delivered at a relative mid-level plateau, thus displaying turn-holding. The second element, which ends with the incongruous item “banana outfit”, then shows falling final pitch, allowing for turn-yielding.

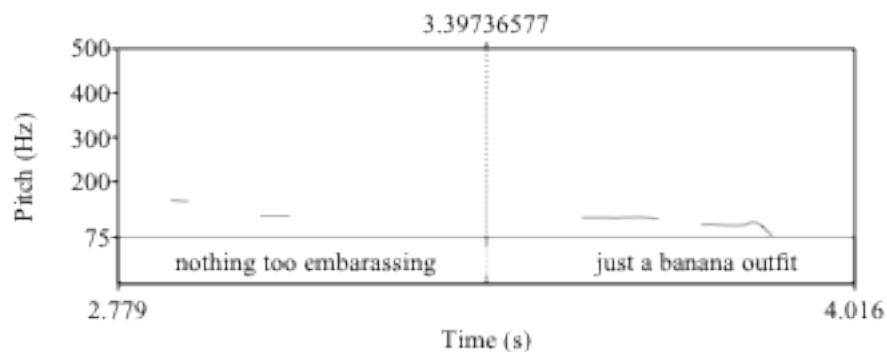


Figure 37: Pitch contour in 1+1 format ("Banana Outfit")

This aspect is further illustrated in “Travel Lodge” (figure 38), where the first element of the impersonation (“if I’m not datin’ her”) also ends in a relative mid-level plateau, while the second element that contains the incongruous moment displays turn-final intonation.⁷⁹

⁷⁹ Impersonator FB displays a Scottish accent, which is why the pitch contour does not mirror the one in “Banana Outfit” (uttered by a speaker using Received Pronunciation Intonation (RPI)). While speakers of RP usually display a falling intonation contour towards the end of declarative or neutral turns, speakers of northern British varieties may recruit Urban Northern British Intonation (UNBI), which tends to end turn-final declaratives in a low rise or rise slump (cf. Cruttenden, 2007), as displayed by FB in this sample.

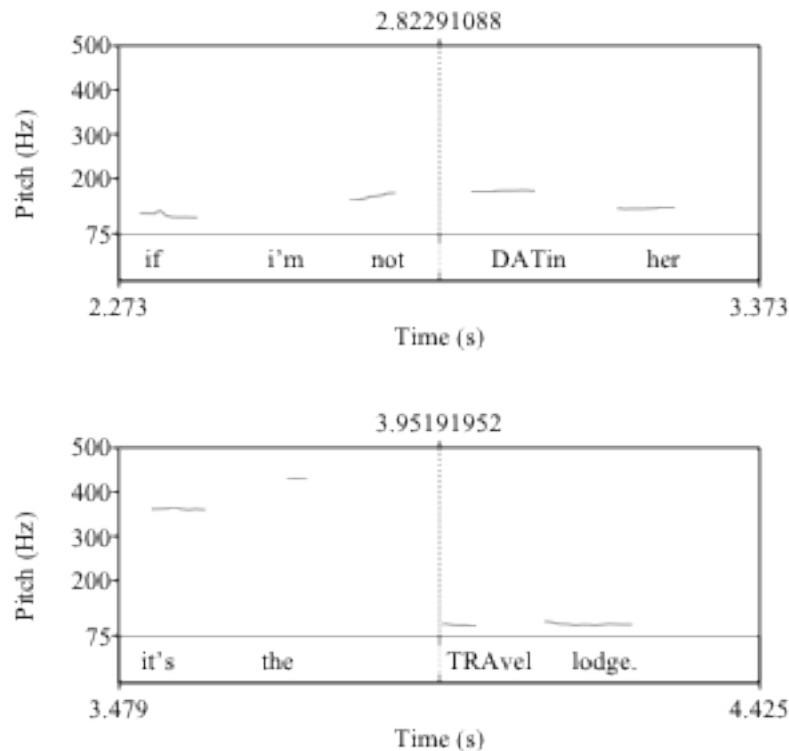


Figure 38: Pitch contour in 1+1 format ("Travel Lodge")

Due to their propensity to restage past dialogue and, in doing so, draw attention to perceived incongruities in prior talk by co-participants, mirror impersonations, then, tend to be short, consisting of just one or two TCUs. Only one instance of the otherwise common 3+1-format can be found in the corpus. Interval spaces for audience approval are only observed after prefaces that provide a diegetic redescription prior to its mimetic illustration. These prefaces show a rising intonation contour at the end, thus projecting further TCUs which contribute to the larger project. While the impersonations that ensue begin with a relatively high pitch, they show falling final pitch towards the end. Impersonations in the 1+1 format that accentuate allegedly incongruent behaviour deliver the first element in a manner that ends with a mid-level plateau intonation, while the second element that features the incongruity at the end is delivered in falling final pitch.

5.3.2. Floor management and performance design in single-scope impersonations

As has been proven, single-scope impersonations are either used to illustrate diegetically set-up story-worlds, or contribute to the narrative unfolding of these story-

worlds. Similar to mirror impersonations prefaced by a diegetic redescription, the first group is followed by an interval during which the audience delivers tokens of appreciation and aligns itself as the recipient of the ensuing performance. The diegetic prefaces to such performances are delivered as syntactically, pragmatically and prosodically complete. Their intonation contours shows unit-final falling pitch, as the prefaces of “Polish band” (figure 39) and “The Proprietor” (figure 40) illustrate.

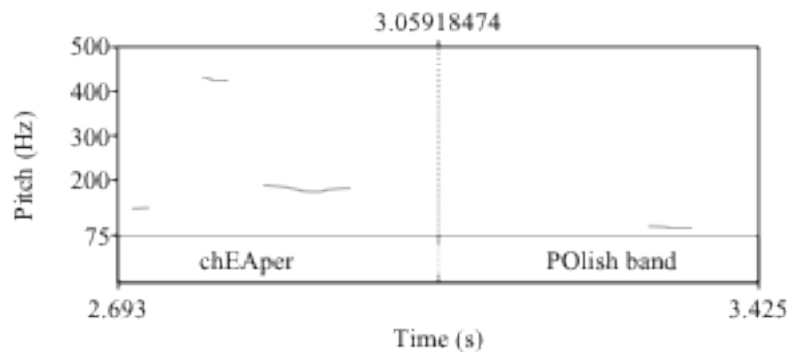


Figure 39: Pitch contour on single-scope diegetic preface ("Polish band")

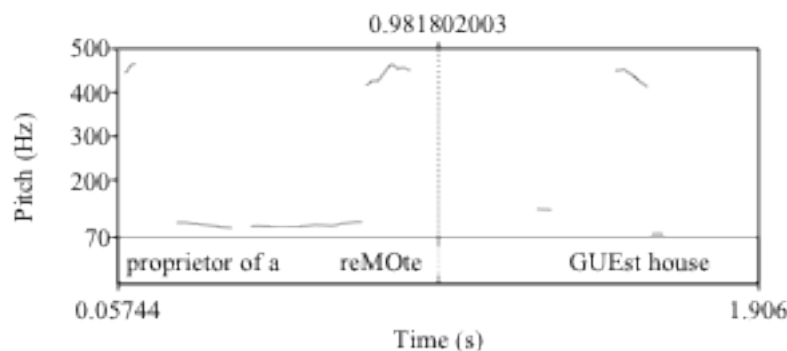


Figure 40: Pitch contour on single-scope diegetic preface ("The Proprietor")

These prefaces consequently differ from those used in mirror impersonations, which redescribe past actions by co-participants. As outlined in the previous section, these prefaces are delivered with a rising intonation pattern.

Those single-scope impersonations introduced with a quotative, however, mirror the intonation contour identified for mirror impersonations, namely a rise to a high peak in the final accented syllable for the quotative *go* and a mid-level plateau for *be like*. The following diagrams (fig. 41-42) illustrate the intonation contours for the two formats.

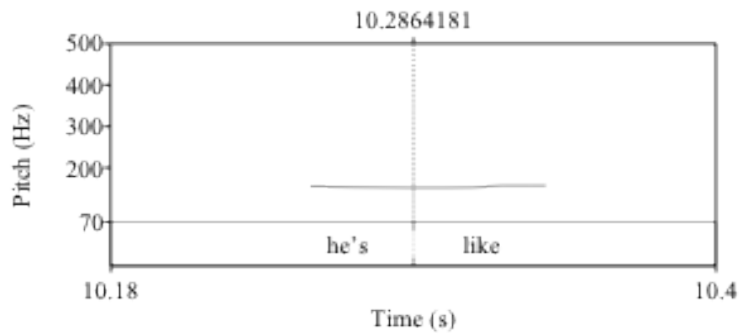


Figure 41: Pitch contour on quotative *be like* in single-scope impersonation (“Justin Bieber”)

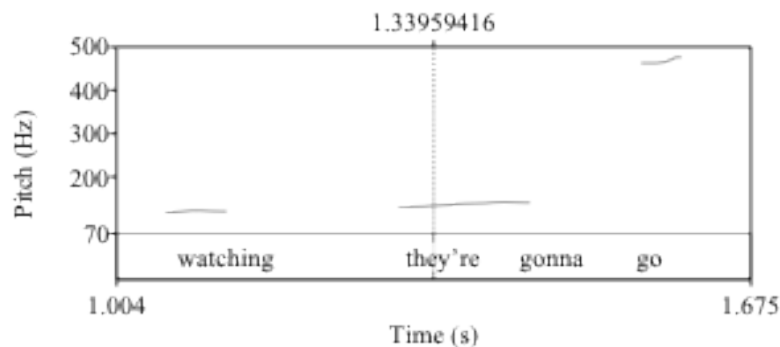


Figure 42: Pitch contour on quotative *go* in "Upset French"

The impersonations following such introductory dialogue elements directly latch onto previous talk with no prolonged interval between them. Due to their non-final intonation contour, they further contribute to projecting an extended sequence of turns-at-talk.

The third group constitutes those impersonations which occur after a TCU that is pragmatically, syntactically and prosodically marked as complete and has begun to illustrate a humorous narrative. These impersonations are followed by an interval during which the audience has the opportunity to display minimal responses and continuers, such as laughter. The duration of these intervals depends on the quality of the audience appreciation signals. If, for example, laughter starts to subside, a following TCU has to be delivered by the impersonator, as otherwise co-participants will have the right to claim the floor. In the NMTB corpus, the maximum length for such an interval seems to be around four seconds before co-participants attempt to take the floor. This becomes evident in “The Proprietor”. Impersonator BB leaves an interval space of 4.4 seconds after the preface, during which he ‘slides’ into the role of the impersonated, aligning his body posture and facial expression accordingly. The audience shows appreciation through clapping and laughter, thereby licensing him to further inhabit the scenario set up. His performance follows the 3+1-pattern already

introduced for mirror impersonations. The first contributions mark the construction of the 'list', as they are delivered in rising intonation at the end of each unit, thus projecting further contributions to the performance. Figure 43 visualises this intonation contour.

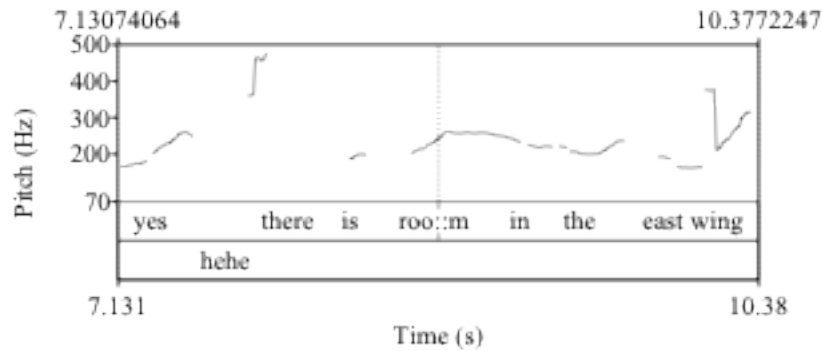


Figure 43: TCU-final rising intonation in list construction ("The Proprietor")

After each chunk, he leaves an interval consisting of approximately four seconds for the audience members to signal appreciation and their status as recipients. After the second chunk, however (lines 10-11), BB leaves a space of 4.6 seconds. As the sound visualised in figure 44 illustrates below, audience laughter is already fading when his co-participant JJ attempts to take the floor in line 13. Apparently, BB has left too much time for this interval, but, as he continues his performance almost simultaneously to JJ's attempt, he shows sensitivity to the issue.

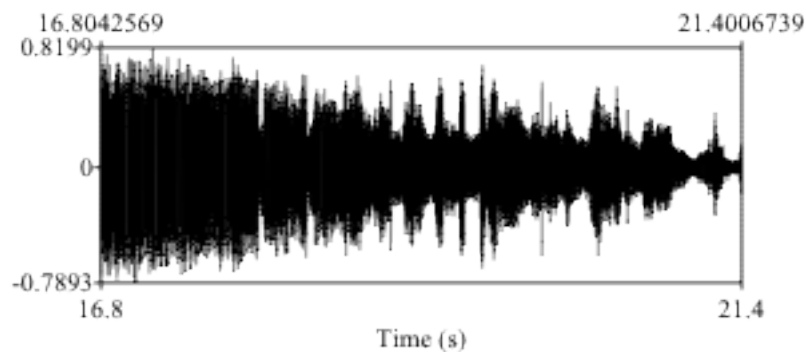


Figure 44: Co-constructing intervals in the chunking of "The Proprietor"

After delivering the third element of his performance (lines 14-15) in a slightly falling intonation contour, he leaves another interval space of 3.2 seconds before delivering the plus-1 in punch line mode (cf. chapter 5.4.2) with a falling, turn-final intonation

contour. A highly stylised post-utterance completion laughter sequence finally ends the performance.

This 3+1-format is a common way of doing impersonations, and can be schematically illustrated as follows:

- 1) Diegetic preface
- 2) Interval for audience appreciation and alignment
- 3) Chunk 1: turn-holding, rising intonation
Optional: Interval
- 4) Chunk 2: turn-holding, rising intonation
Optional: Interval
- 5) Chunk 3: turn-holding, intonation falling to mid-level
- 6) Interval
- 7) Plus-1: turn-yielding, turn-final intonation contour
Optional: punch line-mode, laugh particles as end markers

Depending on local needs, further chunks may be added (as can be observed in “Jason Orange”). The plus-1 is nevertheless always delivered in turn-final intonation (i.e. falling), and illustrates a perceived incongruity.

In essence, then, single-scope impersonations differ from mirror impersonations in one regard: while the prefaces to mirror impersonations tend to be delivered in a turn-holding (rising) intonation contour, prefaces to single-scope impersonations are delivered in falling (i.e.) turn-final intonation. In addition, mirror impersonations are likely to be shorter, and do not provide interval spaces as liberally. Single-scope impersonations, on the other hand, are usually chunked in such a way as to provide jointly negotiated spaces for audience appreciation signals during the performance.

5.3.3. Floor management and performance design in double-scope impersonations

Double-scope impersonations differ from mirror and single-scope impersonations in one major respect: in not one case in the corpus can an interval between the preface and the performance be detected. Even impersonations that are prefaced by

prosodically, pragmatically and syntactically complete TCUs are not followed by an interval; rather, the impersonation latches on without leaving further space.

There may be two reasons for this lack of intervals after prefaces. First of all, such double-scope impersonations of non-human entities occur less frequently than impersonations of fellow humans. They are, in consequence, less established as a discursive practice. To secure the floor for their performances, then, impersonators may need to capitalise on their preface more quickly, as co-participants cannot be expected to orient to a possible follow-up in the form of an impersonation. If this is the case, then this lack of interval supports the interactional relevance of the blending typology as proposed by Fauconnier & Turner (2002). On the other hand, however, double-scope impersonations are often non-linguistic, or are initiated with a non-verbal component, such as a grimace. Entirely non-verbal embodiments, such as in “Holiday horse” or “Stevenson’s rabbit”, begin immediately after the diegetic preface, and last for approximately three seconds. They therefore parallel non-verbal mirror impersonations, such as in “Bemused look”, which also start immediately after the preface. Although “Bemused look” is introduced by the quotative *be like* (i.e. it would therefore not be followed by an interval in any case) and the double-scope examples show zero-quotatives, this already points to the pre-determining function embodied by other semiotic channels, such as grimacing or body postures. All impersonations which recruit such channels in addition to the verbal layer ‘slide’ into the impersonated body prior to assuming its voice. “The Proprietor” here, too, constitutes a typical example: immediately after the preface and during the interval that follows, impersonator BB already assumes the impersonated character’s posture and facial expression, leading to further laughter coming from the audience (which might to some extent explain the unusually long duration of the interval). It thus transpires that non-verbal elements of impersonations may be delivered straight after the preface, and, in cases that also contain a verbal layer, serve as an ‘interlude’ which allows the silences between verbal elements of impersonations to be prolonged.

Those double-scope impersonations that feature a linguistic component in their performance follow the techniques described above for mirror and single-scope impersonations. Once impersonations containing verbal elements have been established, they may be chunked in such a way as to allow for intervals to occur. They also orient to lists of three and the 3+1-format discussed above. In “Metal Detector”, for example, impersonator NF delivers three blocks, all of which are

introduced with the non-lexical rendering of the sound made by metal detectors (“meep”). He leaves intervals for audience approval between the blocks, and delivers the third and final component in punch line mode and falling, turn-final intonation contour. “Displeased horse”, on the other hand, demonstrates the 3+1-format: the impersonation is delivered in three units, while the plus one is a joking meta-comment on the horse which has just been impersonated.

5.3.4. Discussion

When studying the performance of impersonation, then, single-scope impersonations appear to be the most typical. In general, they are longer projects delivered with more interval spaces which are longer in duration, thus pointing to their well-established nature. Impersonators doing single-scope impersonations here seem to be able to rely on their co-participants’ orientation to a salient communicative practice. Mirror and double-scope impersonations are, all in all, shorter, and leave less space for such intervals. In addition, double-scope impersonations are more often purely non-linguistic, relying on other semiotic channels such as facial expression and gesturing.

While these aspects point towards the validity of the typology put forward by Fauconnier & Turner (2002), it should, however, be kept in mind that these aspects also contribute to the building of action and frame integration complexity is not relevant to action. The action accomplished via non-verbal gestures in double-scope blends is indicative of a diegetically produced story – an action to which single-scope blends such as “The Proprietor” also contribute. Similarly, purely non-verbal impersonations are used to build the action of detecting deviant or noteworthy behaviour from co-participants, as the mirror impersonation “Bemused look” illustrates. The extent to which single-scope impersonations are celebrated in contrast to mirror and double-scope impersonations may furthermore have cultural reasons: a mirror impersonation often constitutes a potentially face-threatening act (cf. Brown & Levinson, 1986), and is hence interactionally problematic. Single-scope impersonations which allow for the dismissal of a non-present human are, on the other hand, a useful tool for creating solidarity. By mocking and jointly laughing at a non-present other, in-group status may be celebrated. This means that single-scope impersonations are also a means for creating group solidarity. According to this line of

argument, then, marking non-human entities as outsiders may be less attractive and for this reason less common.

In short, the methods identified in this section contribute to building social action which occurs in situated cultural scenarios and spans blending types. On this basis, they cannot be used as evidence for validating that different blending types constitute separate discursive practices.

5.4. Ending and exiting impersonation sequences

While the previous section focused on the largely prosodic methods used to secure an extended sequence of turns-at-talk, this chapter will closely examine on further techniques for indicating the end of the multi-turn unit ‘impersonation’. Such practices are crucial to this situated activity system, as they are of central importance to managing turn-taking. They display a discourse unit’s (Houtkoop & Mazeland, 1985) conclusion and consequently the point at which the standard turn-taking model deployed in spontaneous ‘normal’ conversation (cf. Sacks, Schegloff, & Jefferson, 1974) may be resumed.

In this section I will therefore discuss the various methods employed for indicating designed closure by the impersonator (cf. Schegloff, 2009), and the techniques used for re-engaging with topical talk and the shift in participant alignment that this requires. The previous sections further support the hypothesis that the techniques employed in the interactional negotiation of impersonations seem to be guided by and simultaneously inform the social action in which such sequences are embedded rather than orient towards the blending typology advanced in MSCI. As such, this section will not discuss the techniques grouped according to blending types, but will start with the methods themselves and discuss the extent to which they are found in the varying types against the backdrop of previous findings. Furthermore, the chapter will only discuss techniques specific to impersonations, and will not touch upon other markers that project a possible Transition Relevant Place (TRP), such as syntactic or pragmatic completion (cf. Sacks et al., 1974).

5.4.1. Repeats

Previous research has established that a common practice for ending longer units in talk-in-interaction (such as stories or answers) are so-called ‘repeats’ (Schegloff, 2009, 2011), which cite an element from the beginning of the unit and thus mark its conclusion. Such repeats are, however, not exceptionally common in the NMTB corpus, and can only be found in mirror impersonation and a single double-scope impersonation.

Mirror impersonations, which have as their source prior talk, are naturally inclined to signal conclusion via such repeats. Amongst other mirror impersonations, “Northern lass” features such a repeat. The lexeme *stand* is focused on as early as the preface introducing the impersonation sequence (“What about standin’?”), and also concludes the performance (“I wish I could go to London and stand”). The same pertains to “Banana outfit” and “Travel Lodge”, in which the impersonators conclude their performances by repeating the central element of the original contribution. This central component also constitutes the incongruous element that clashes with assumptions created in the first TCU. The cited element in turn-final position therefore draws attention to the allegedly incongruous behaviour of co-participants. In all sequences, the repeated lexeme stands in TCU final position and follows the format *x*[IMP_x]. The lexical item repeated thus provides a bracketing of the impersonation sequence, and at the same time relates back to topical talk, thereby signalling that the performance sequence has reached a conclusion.

Conversely, single-scope impersonations seem not to favour this technique, as no repeats occur in this type. They do, however, appear again in one double-scope impersonation, namely “Metal Detector”. This impersonation occurs after a question sequence that focuses on the materiality of metal detectors and ends with a repeat of the incongruous ‘problem’ dwelt upon in the preface. The central component of the preface, “Are metal detectors made out of metal?”, is also the concluding element of the impersonation’s final TCU (“I’m mainly metal”, line 19). As such, the adverb ‘metal’ brackets the impersonation. Still, this is the only sequence in which such a repeat contributes to completing an impersonation sequence. A possible reason for this distribution of the ‘repeat’ exit device may be that such returns to lexemes used in the turn preceding the performance are especially prone to occur in impersonations that have a verbal component of topical talk as their trigger, as is the case for most

mirror impersonations (except for those continuing in the voice of a co-participant without disjunct markers) and “Metal Detector”. In contrast, impersonations that do not reflect on or further illustrate a verbal component of prior turn-by-turn talk resort to other means of indicating their potential completion.

5.4.2. Punch line-mode

Punch lines are usually the final element of a joking sequence, providing both its climax and humorous incongruity. In verbal humour, such punch lines are delivered in a special ‘mode’ (Norrick, 2010) that involves the emphasis of each lexical⁸⁰ item as well as increased overall loudness. The marked performance of punch lines is of paramount importance, as participants will orient to it and, according to Norrick (*ibid*), laugh at what they consider to be the highpoint of a joking sequence (based on performative evidence). Punch lines (or, rather: units delivered in prototypical punch line mode) are thus a major means of signalling the end of an impersonation performance. In addition, Jefferson (1978) asserts that punch lines summarise narratives. This feature relates to other ending devices discussed in this section, such as codas and postscripts.

Based on these findings, such punch lines can be assumed to constitute a major means of concluding impersonation sequences. Impersonations delivered in punch line mode are found across blending types. However, as mirror impersonations largely rely on the quoting of previous utterances, they are not prone to showing such climactic prosodic framing in their performance.⁸¹ They might nevertheless feature a punch line which follows the impersonation sequence. “Northern lass”, for example, ends with a comment delivered in punch line mode. As this punch line is not attributed to the previous speaker, it is not part of the impersonation proper, but as part of a coda (see chapter 5.4.4.) delivers a meta-comment on the impersonation in

⁸⁰ The definition of ‘stress’ is, of course, a notorious problem, as it is not just loudness, but also pitch, duration, and ‘secondary cues’ such as glottal stops which contribute to perceiving a given syllable as stressed (Couper-Kuhlen, 1986). My analysis here concentrates largely on loudness as the key contributor to realising and perceiving a given syllable as stressed. Accent, on the other hand, is achieved via loudness and changes in pitch (cf. Szczepek Reed, 2006:6).

⁸¹ Mirror impersonations that recruit ‘climactic’ speech patterns such as the ones employed by race commentators are found in the corpus. I would, however, refrain from treating them as punch lines, as they first and foremost copy a speech genre rather than serve the local management of talk, and secondly do not lead to a conclusion that condenses a narrative, as is evidenced by both “Sports Commentator” sequences.

accentuated punch line mode. Figure 45 visualises the intensity with which the punch line is delivered.

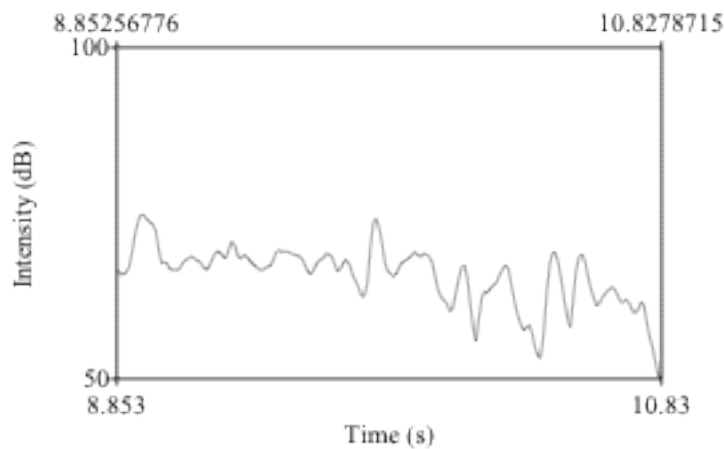


Figure 45: Intensity for "Northern lass" punch line-attachment

Impersonator LM leaves 1.6 seconds in the interaction-as-broadcast for audience laughter after the impersonation performance and prior to delivering the punch line. This final line is instigated with the stressed discourse marker “well” at 74.84 dB, followed by a strongly accentuated utterance of the word “dreams” at 73.99 dB, and then succeeded by a less intense “true”. The average intensity for this impersonation attachment is 66.46 dB; this is noticeably louder than the impersonation performance itself, which featured an average intensity of 62 dB.

“The Proprietor”, a single-scope impersonation, illustrates even more clearly the concluding function fulfilled by the punch line mode. After a long pause of 3 seconds, line 18 (“on my back”) is in this impersonation delivered in a husky voice. All words receive equal stress (cf. figure 46), with a micro-pause before “back”. Of all the items in the entire impersonation sequence, “back” is the one which is uttered in the lowest pitch (cf. figure 47), thus further proving the concluding function of this TCU.

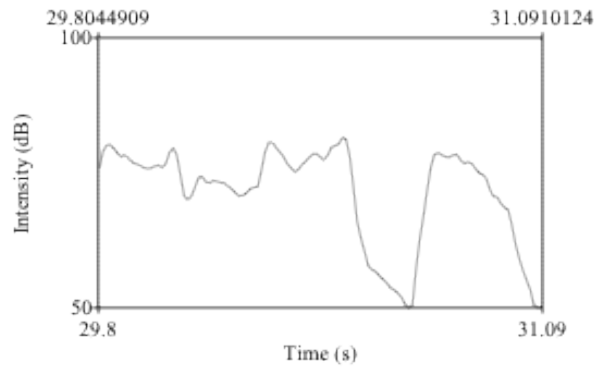


Figure 46: Intensity for punch line in "The Proprietor"

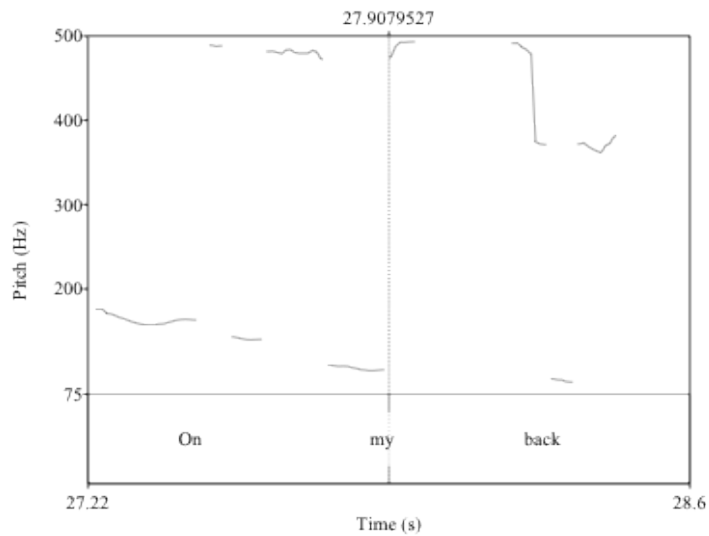


Figure 47: Pitch contour for "The Proprietor" punch line

Finally, consider once more “Metal Detector” and the punch line (line 17 in the transcript) which condenses its narrative structure. As in “The Proprietor”, all lexical items of “I’m mainly metal” are approximately equally stressed, as is shown in figure 48.

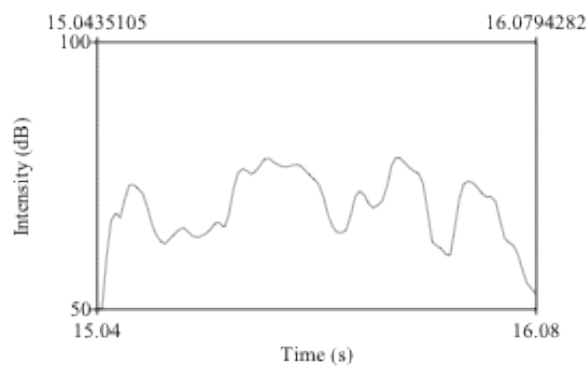


Figure 48: Intensity for punch line in "The Metal Detector"

It therefore transpires that punch lines – performed after a longer silence consisting of between 1.5. and 3 seconds long in a louder and highly accentuated mode – are a means of ‘doing’ endings of multi-turn unit impersonations. While they are more typical for single- and double-scope impersonations, mirror impersonations might also feature prosodic framings akin to the frame employed in punch line-mode.

5.4.3. Coda

In their influential treatment of conversational narrative, Labov & Waletzky (1967) introduce the term *coda* to denote the final element of a story that relates back to the moment of speaking. A similar concept is found in impersonations. Once an impersonation has been delivered, impersonators might provide a brief meta-commentary upon completion, recruiting a non-marked register and body positioning and, in doing so, ‘return’ to the current discourse scenario characterised by topical turn-by-turn talk. Such codas are found across blending types, and thus cannot be accepted as a reflection on frame-integration complexity. In mirror blends, “Flirting” shows such a postscript, when, in line 07 in unmarked voice for example, impersonator DO utters “like that”, directly attaching itself the impersonation sequence. DO thus points to the illustration of the activity offered by the impersonation, but also introduces a trajectory that allows him to return to topical turn-by-turn talk when, in line 09, he repeats the question which he posed to AM prior to the impersonation sequence.

Similarly, TM uses such a meta-commentary in “Kylie”, a single-scope impersonation, when he utters “you know, that sort of thing”, again directly appending itself the impersonation. Labov & Waletzky (1967) have already pointed to the use of the deictic marker *that* commonly found in such codas, arguing that *that* points to the sequence recently concluded as distant and accessed from the perspective of topical talk. In using a coda with the obviate deictic marker *that*, impersonators display their impersonation sequences as complete and distant from the current discourse, and invite co-participants to re-engage in topical turn-by-turn talk following their multi-turn performance.

Another form of coda for impersonation sequences are what I refer to as ‘summaries’. In “Displeased horse”, a double-scope enactment, impersonator NF follows up on his impersonation with just such a summary, and evaluates the

impersonated horse's actions from a distant perspective, that of the discourse scenario containing topical turn-by-turn talk. In lines 09-10, he subsequently concludes that "you can lead a horse to a pop video, but you can't make it like it". In contrast to the previous two examples, however, this evaluation concerns the impersonated horse rather than the performance as such (as the "like that" of the sequences discussed above do), and is not delivered in a manner that directly 'latches onto' the impersonation, so to speak, but occurs after a silence of about 1.5 seconds. The summary, while delivered through NF's 'own' voice rather than the horse's, thus constitutes a continuation of the joking, i.e. a top-up joke, rather than a return to topical turn-by-turn talk. It does, however, contribute to marking the impersonation sequence itself as complete, and thus parallels the functions fulfilled by the coda in "Northern lass", a mirror blend ("Well tonight your dreams come true"). Both impersonation sequences receive a top-up joke after a break of approximately 1.5 seconds delivered in unmarked voice, but are still oriented to the joking activity rather than topical talk.

It thus appears that such attachments contribute to the conclusion of impersonation sequences in two ways: they firstly provide a means of exiting a voice and distant discourse scenario by pointing to the illustrating function of the impersonation and can secondly serve as a means of returning to the original voice, delivering a follow-up joke prior to returning to topical turn-by-turn talk. These two functions differ with regard to their temporal distance to the impersonation. While the first is delivered in a manner that directly attaches itself to the impersonation, the latter is uttered after a brief silence of about 1.5 seconds (cf. chunking). For all this, though, no difference as to blending type can be detected, as impersonations across all sets employ codas.

5.4.4. Laugh particle transitions

Across all impersonation types, laugh particles in the final TCU of an impersonation are used to signal the transition from the impersonated voice to the unmarked voice and, consequently, the return to topical turn-by-turn talk. Apart from marking this transition, laugh particles also provide a meta-commentary on the uptake expected from recipients (cf. Jefferson, 1979), and thus further contribute to the humorous framings of the performance. Finally, such laugh particles frequently occur towards

the climax of a humorous narrative (cf. Goodwin, 1984), and in doing so contribute to signalling the conclusion of an impersonation.

Consider, first of all, “Narcissistic Barrowman”. In this mirror impersonation, SA concludes his performance with “E(h)ven BU(h)zzcocks.” in line 15. The insertion of laugh particles into this concluding TCU serves to further frame the performance as pragmatically complete; in effect, a humorous illustration of JB’s allegedly deviant behaviour has been delivered, and impersonator SA invites his recipients to laugh along with him. In this manner, however, SA also tones down the face-threatening act (FTA) (cf. Brown & Levinson, 1987) inherent in the mirror impersonation. Marking it as humorous, i.e. non-serious, mitigates the FTA. This mitigation is further enhanced by the fact that the laugh particles in this segment also serve to frame a case of self-deprecating humour, as they contribute to the presentation of SA’s own show as worthy of ridicule.

Such laugh particles only occur in final TCUs of impersonation sequences, as becomes clear in the double-scope impersonation “Displeased horse”. In this sequence, two joking sequences centre on the topic of horses in pop videos. The first (lines 05-07) is an impersonation of the horse, whilst the second (line 09-10) constitutes a comment on the topic displayed in the form of a coda (see above). Both joking sequences end in TCUs containing laugh particles. While the impersonation is concluded with “gi(h)ve me(h)” (line 07), the ensuing comment ends with “ma(h)ke it li(h)ke it.” (line 10). The laugh particles therefore serve to ‘do’ endings of distinct joking sequences and to invite recipients to laugh as well.

Apart from such laugh particles occurring within speech, Jefferson (1979) finds that post-utterance completion laughter may serve similar goals. Only one instance of such post-utterance completion laughter occurs in the NMTB corpus. In “The Proprietor”, however, it is the impersonated character himself that delivers a highly stylised laughter that ends his ‘appearance’ rather than that of the impersonator. It nonetheless serves to fulfil a similar objective to the embedded laugh particles found in the examples above. Laugh particles project designed (i.e. projected) closure of humorous impersonations, and thus provide a means of transitioning from the impersonated voice back to the original. In addition, they invite recipients to laugh along as a relevant next action before topical turn-by-turn talk is continued. As laugh particles in unit-final position occur across blending types, they do not hint at frame integration complexity either.

5.4.5. Practices for reengaging with topical talk

Impersonation sequences project laughter as their relevant next action. Once the audience has delivered this token of appreciation, topical talk is resumed at the point at which it was discontinued through the impersonation. These impersonations are, in other words, treated as non-implicative for topical talk. Impersonations may, however, have further implications for topical talk, and can be followed by codas (see above) in the form of “postscripts” (Jefferson, 1978:229), assessments, or top-up jokes.

As a starting point, consider those impersonations which are treated as irrelevant to topical talk. In these sequences, topical talk is resumed from the point where it stopped prior to the impersonation. The techniques employed for turn-taking management in this case are either

- a) co-participant self-selects, or
- b) current speaker continues after the token of appreciation.

In a), co-participants have a variety of techniques at their disposal with which to return to topical discourse. Firstly, they may recruit discourse markers such as *anyway* as a means of signalling the conclusion of the ‘detour’ made by the impersonation sequence and the recommencement of prior topical talk. Previous research (e.g. Bublitz, 1988; Ferrara, 1997; Lenk, 1998) has identified *anyway* as the paradigmatic discourse marker for concluding ‘digressions’. In such environments, *anyway* is used to return to prior topical talk. The use of this discourse marker by co-participants rather than current speakers to return to topical talk is thus relatively dramatic, as by doing so they mark the impersonation sequence as non-implicative for topical talk. However, co-participants only resort to such apparently drastic measures after the implementation of impersonations that provide illustrations of playfully established story-worlds. Treating them as conversational ‘décor’, which is irrelevant to the progression of topical talk, is, for this reason, not face-threatening. After all, the relevant SPP, namely laughter, has occurred prior to the co-participant forcing the resumption of topical talk. This becomes evident in the single-scope impersonations “The Proprietor” and “Cooking with villains”.

The Proprietor (NMTB SE19E06_3_04:15)

18 BB: <<f> ON MY (.) BAck.> > (2.0)
 19 <<ff> ↑hu[ohoha> @>
 20 AU: [@@@ (1.5)
 21 AB: anyway.
 22 there's a bit of a division between,
 23 some of them got sparkly jackets,
 24 some of them (.) jackets.

In the first example, impersonator BB uses numerous techniques (punch line-mode, falling intonation contour, stylised post-utterance completion laughter) to show designed closure of his multi-turn impersonation unit. Once the audience has delivered their tokens of appreciation, co-participant AB proposes a return to topical talk in line 21, again referring back to the guessing game.

NMTB SE21E08_1_04:34 (Cooking with villains)

08 BB: time for a LOvely salmon salad. @>
 09 AU: @@@ (2.0)
 10 SA: anyway.
 11 we're very pleased about the cooking show.
 12 JE: brilliant.

Similarly, in “Cooking with villains”, host SA resumes topical talk after the laughter sequence, addressing his previous primary conversational partner, JE, and redirecting talk to the theme of their prior interaction.

However, co-participants self-selecting do not need to recruit discourse markers to signal conclusion of the impersonation sequence or a return to topical talk. They might also draw attention to an element of prior topical talk by either citing this element or posing a question on that topical talk.

NMTB SE23E02_2 03:30 (The Sports Commentator I)

08 PJ: <@ tops fou_↑`our;
 09 jackson fi_↑`ive; @>
 10 AU: @ (1.0)
 11 RG: you were on to four tops
 12 i give you a clue
 13 you were right.

NMTB SE22E01_1_09:25 (Upset French)

04 AU: <@ <<ff> AAH quelle est POp quiz
 05 PJ: IMB!É!cile;
 06 AU: @[@@
 07 PJ: [((encore)) > @>
 08 AU: @@@ (1.0)
 09 MR: and there was, (.) what was CEE?

In the mirror impersonation “Sports Commentator I”, host RG summarises – after the audience’s token of appreciation for the impersonation – GL’s contributions to topical talk prior to PJ’s impersonation, thereby leading back to the original activity. In the single-scope impersonation “Upset French”, team captain MR returns to topical talk by asking a question about the activity. In both cases, the impersonation sequence itself is not treated as relevant to the furthering of topical talk, but as a digression that served amusement.

Impersonators themselves may return to topical talk without further reference to the performance just delivered – thus illustrating that treating impersonations as non-implicative for further topical talk does not constitute a potential FTA, but rather common practice. In single-scope blend “Northern Mac”, impersonator LM, after the obligatory laughter sequence, provides the correct answer to the quiz question posed. In the double-scope impersonation “Metal Detector”, impersonator NF asks to again see the video clip upon which the quiz question posed to his team is based.

Secondly, a group of impersonations can be identified that is followed by further postscripts and commentaries in the form of a coda rather than a straight return to topical talk. These impersonations are not treated purely as conversational ‘décor’ that may be concluded merely by the bestowal of a token of appreciation, but are followed by a closing sequence that reflects on a) the impersonation’s relevance by providing a concise ‘morale’, or b) its aggressive nature prior to returning to topical talk.

After this concluding sequence involving the impersonator and the co-participants negotiating the emerging implications of the impersonation, topical talk may be resumed using the same strategies identified for impersonations that are not followed by such a sequence. Through questions and repeats, which are sometimes introduced with a discourse marker such as *anyway* or *well*, joint attention is redirected to topical talk. In “Travel Lodge”, for example, it is the impersonator who uses the

discourse marker *anyway* to mark closure of the impersonation sequence (“teller-triggered self-digression management” in Ferrara’s (1997:343) terms) prior to providing a meta-commentary on the aggressive nature of his mirror impersonation (lines 25-26). Once the ‘butt’ of his impersonation, DJ, has accepted this instance of mitigation (line 27) and impersonator FB has, too, confirmed this ‘ceasefire’ (line 28), FB reverts back to topical talk by starting to repeat the quiz question posed to the team prior to the digression proposed by FB (line 29).

NMTB S23E11_2_09:26 (Travel Lodge)

22 FB: <@ yeah. if i’m not datin her; (.)
 23 it’s the TRaVel lodge. @>
 24 AU: @@
 25 FB: anyway i don’t know why (.)
 26 i’ve decided to start a beef with you,
 27 DJ: we’ oh we’re cool we’re cool [he he
 28 FB: [we’re cool man. (-)
 29 erm any ideas, er=
 30 CV: !! have

Impersonations, whether mirror, single-scope or double-scope in nature, are, in conclusion, followed by similar strategies for reengaging with topical turn-by-turn talk. Topical talk might be resumed immediately after the token of appreciation in the form of audience laughter has been delivered through repeats of prior segments or questions leading back to prior talk. These strategies may be introduced through discourse markers such as *anyway*, which further serve to mark the multi-turn unit as complete. It is either the current speaker that returns to topical talk by employing these strategies, or a co-participant that self-selects. Impersonations might, however, be followed by a postscript (coda) that serves to negotiate the emerging implications of an impersonation between performer and co-participants. Once this postscript sequence has been concluded, topical talk recommences following the same strategies as for impersonations which are not followed by such postscripts.

5.4.6. Discussion

While mirror impersonations – due to their focus on prior talk by co-participants – seem to show a certain propensity to be concluded via repeats that illustrate alleged

incongruity in turn-final position rather than punch line-mode, other exit strategies, such as coda and laugh particles in unit-final position, occur across all impersonation types. The same applies to strategies identified for returning to topical talk. All blending types rely on the same array of techniques, including using concluding discourse markers prior to repeating an element of prior talk or asking a question concerning the previous activity.

These findings once again point to the hypothesis that interactants in meaning-making processes primarily orient to joint action rather than the cognitive tools employed to build this action. Meaning-making, then, emerges from and is guided by recursive beliefs regarding the action currently being negotiated rather than one's ability to blend distinct semantic frames. Blending may indeed be a tool used in meaning-making, but it does not appear to be *the* way that we make meaning. This hypothesis is discussed further in chapter 6, which will propose an organic model for meaning-making that integrates cognitive tools, such as conceptual blending, with the locus of meaning – intentional social action.

5.5. An action-based typology of impersonation blends

'Doing' an impersonation is a social, jointly negotiated activity that serves to achieve two main goals. First, impersonations may be used to build assessments of prior contributions. These stance-taking impersonations follow three distinct formats. Firstly, they may offer an accentuated restaging of a prior contribution (e.g. "Banana outfit"), as visualised in figure 49. This form of impersonation is not diegetically prefaced, and voices are set apart through disjunct markers such as "oh". Shared gaze and postural orientation is discontinued in these impersonations. Secondly, impersonations-as-assessments may be prefaced with a reframing of the prior activity, which is followed by a (re)staging of the original contribution. In impersonations that offer a reframed restaging (e.g. "Sports Commentator", "Northern lass"), shared gaze and postural orientation are discontinued, and quotatives may or may not be used depending on the degree of bodily display projected. In this respect, *be like* is, for instance, used to introduce more bodily restagings (cf. "Bemused look", "Holiday horse"). On the other hand, shared gaze and postural orientation is maintained in

impersonations that offer adapted stagings of diegetically presented assessments of allegedly negative character traits displayed by co-participants (e.g. “Narcissistic Barrowman”, “Smug Jonas”). Furthermore, quotatives are not used in this type. These more aggressive impersonations effectively state, ‘this is you’. Finally, assessment of prior talk through impersonations may be delivered by continuing in the voice of the prior speaker without reframing the contribution (e.g. “Frustration”). These impersonations latch directly onto prior talk, and gaze and postural orientation are maintained.

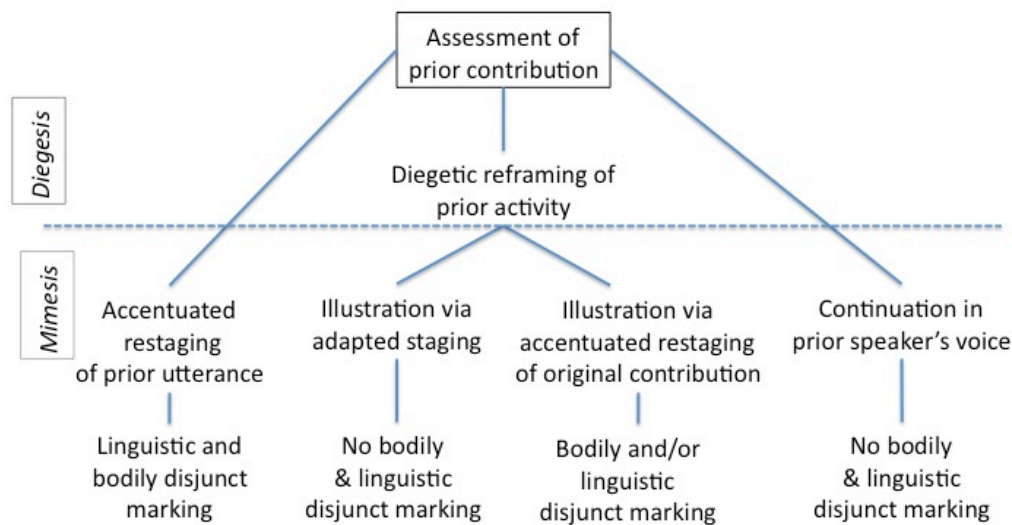


Figure 49: Typology of impersonations for assessments

The second group of impersonations subsumes narrative joke-firsts that are set in distinctly separate discourse scenarios. Two subgroups can be identified (cf. figure 50). First of all, such joke-first narrations may be diegetically presented; the preface outlines the spatiotemporal macro-structure of the story-world, while the impersonation provides a mimetic visit to that world, and illustrates ‘best-ofs’ resulting from the narrative presented (e.g. “The Proprietor”, “Cooking with villains”, “Polish band”). These impersonations are introduced with zero-quotative, discontinued gaze and postural orientation. They are highly dramatised, and tend to be chunked in the 3+1-format, showing a propensity to end in punch lines. Secondly, such narrative impersonations may not constitute brief visits to a diegetically set-up story-world, but actually contribute to the construction and narrative advancement of the story-world. They are thus polyphonic narratives, with short voice changes contributing to the

unfolding of a story (e.g. “Kylie”). These impersonations are introduced through quotatives (*go* is largely used for verbal elements), and shared gaze and postural orientation is discontinued during the performance, but may be adopted again during ‘intervals’.

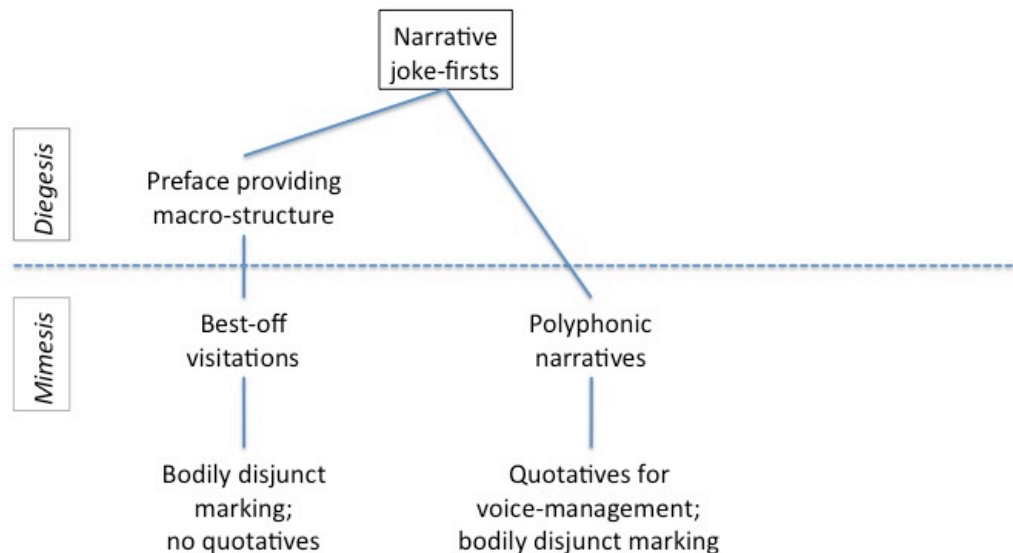


Figure 50: Typology for narrative joke-first impersonations

These greater actions – assessment and story-telling via and in conjunction with impersonations – span across all blending types. While mirror blends are more prone to contribute to assessments, single and double-scope blends may serve to build assessments (e.g. “Brothers” or “Holiday horse”), but also to develop joke-first narratives and illustrations. Participants in the interaction-as-broadcasted in their joint negotiation of meaning-making orient to these types, as is evident in “Sports Commentator II”, in which the ‘butt’, GL, explicitly challenges the redescription of her contribution to topical talk. GL’s attempts to rebuff the assessment proposed by the impersonator illustrate the fact that she is not treating the impersonation as a non-face threatening joke-first narrative, but as an explicit (negative) assessment of her contribution.

‘Doing’ an impersonation is thus a jointly negotiated activity that is triggered by and methodologically displays its relevance to topical talk. However, the analyses have evinced that an orientation to disjunct discursive scenarios is reflected in interaction via the discontinuation of postural orientation or the use of grammatical disjunct markers and space builders when orienting to non-topical scenarios. Yet it is the orientation to the ‘greater’ action achieved through the impersonation performance that guides both the intersubjectively negotiated construction of Mental

Spaces and the mappings occurring between them. The Mental Spaces recruited to the blends are systematically introduced with an eye toward recipient design, and the meaning emerging from the integration of distinct frames is dependent on and constrained by recursive beliefs regarding the joint action currently being negotiated. For Mental Spaces and Conceptual Integration Theory to be a comprehensive theory which comprises thinking and meaning-making, it will need to incorporate the cooperative and goal-oriented nature which is evident in jointly negotiated online-meaning-making. The following section will propose one such model.

CHAPTER 6: THE PRIMACY OF PRAGMATICS

The analyses in the previous chapters have shown that neglecting the interactional dimension of meaning-making processes does not hold up to online meaning construction in interaction. “Context matters”, in the words of Fauconnier (2006:657). The structure emerging from the integration of distinct frames, as well as the mapping processes occurring between input spaces is, it turns out, largely guided by a quest to arrive at meanings that are experienced as interactionally relevant, as Fauconnier (2006) also acknowledges. For this reason, blending is not something that happens wildly during interaction, but something that achieves shared meaning from and is made relevant (in the broad sense of the term as employed by Grice (1975), in and Sperber & Wilson’s (1986) expanded notion thereof) by being tied to a given social action. Although Fauconnier & Turner (2002:333f) mention relevance, among numerous others, to be one of the “governing principles” behind blending, and claim that “context” will determine the “happiness” or “equilibrium” of a blend (*ibid*:44), they do not seem to regard (pragmatic) relevance as one of the more important ones. They first of all devote only a single page out of the 396 which comprises their magnum opus to the principle, and secondly focus on “Network Relevance”, which “can be satisfied for an element in the blend if it can be successfully taken as a prompt for unpacking” (Fauconnier & Turner, 2002:334). With this focus on Network Relevance, Fauconnier & Turner concentrate largely on relevance in relation to internal aspects of the blend and the outer-space relations compressed in the blend. The blend is here conceptualised as ‘bigger’ than action, and might actually influence action due to the fact that agents ‘live’ in a blend.

The more pragmatic notion of relevance as a blending constraint that I would like to propose, however, starts at the other end, and regards joint action as ‘bigger’ than blending. Relevance to joint action is, I argue, not only the major constraint exerted upon blending; it also prompts blending processes, and determines their internal structuring, such as vital relations (*cf.* Brandt & Brandt, 2005; Schmid, 2011 for similar assessments).

Considering the issues raised in chapters two and three, and building on the results gathered in chapters four and five, in the following I propose a model of MSCI

that answers to the centrality of relevance and joint action in meaning-making processes.

6.1. Conceptual blending in interaction: a discursively grounded model of in-situ meaning coordination

As was shown above, the meanings emerging from creative blends are relevant to and prompted by jointly coordinated beliefs regarding social action, such as joking or commenting on a mutual friend's new appearance. Following on from these basic findings, the inputs into and mappings occurring in blending-in-interaction must be interactionally relevant, systematically introduced, sequentially organised, recipient designed, and jointly produced; furthermore, the meaning that emerges is grounded in assumptions concerning joint action and shared Common Ground, and is intersubjectively shared. Otherwise, repair mechanisms are employed. Figure 51 incorporates these results and illustrates a general blending model⁸² that hopes to be able to “explicitly model [the] specific features of local context” (Coulson & Oakley, 2005:1507) that impede on meaning construction. Other than Brandt & Brandt's (2005) model (see chapter 3.1.1. for a full description), which posits a “virtual space” that gives rise to a “meaning space” and formally differentiates between semantics and pragmatics, this model does not propose an apparently ‘meaningless’ space that precedes the one containing meaning. Rather, it regards all processes as guided by orientation to joint actions and a quest for interactionally relevant meaning. For this reason, no “virtual spaces” are required in fully contextualised meaning-making processes. In addition, my approach departs from Brandt & Brandt (2005) in that it explicitly considers Theory of Mind and interactionally coordinated beliefs with regard to the Common Ground existing between interactants. It furthermore makes allowance for the role of signs and stimuli. In doing so, this model aims to provide an integrated approach that considers both interaction and cognition, thereby providing suggestions for advancing MSCI within a socio-cognitive framework as suggested by

⁸² This model is not able to account for simplex blends; rather, it is only able to justify conceptual integration which recruits theatrically distinct scenarios, thereby resulting in dissonance on the discourse level (see chapter 3.1.3. and 3.4.1.).

scholars such as Croft (2009). In the following, I will describe the model in more detail, illustrating it with examples discussed in the previous chapters.

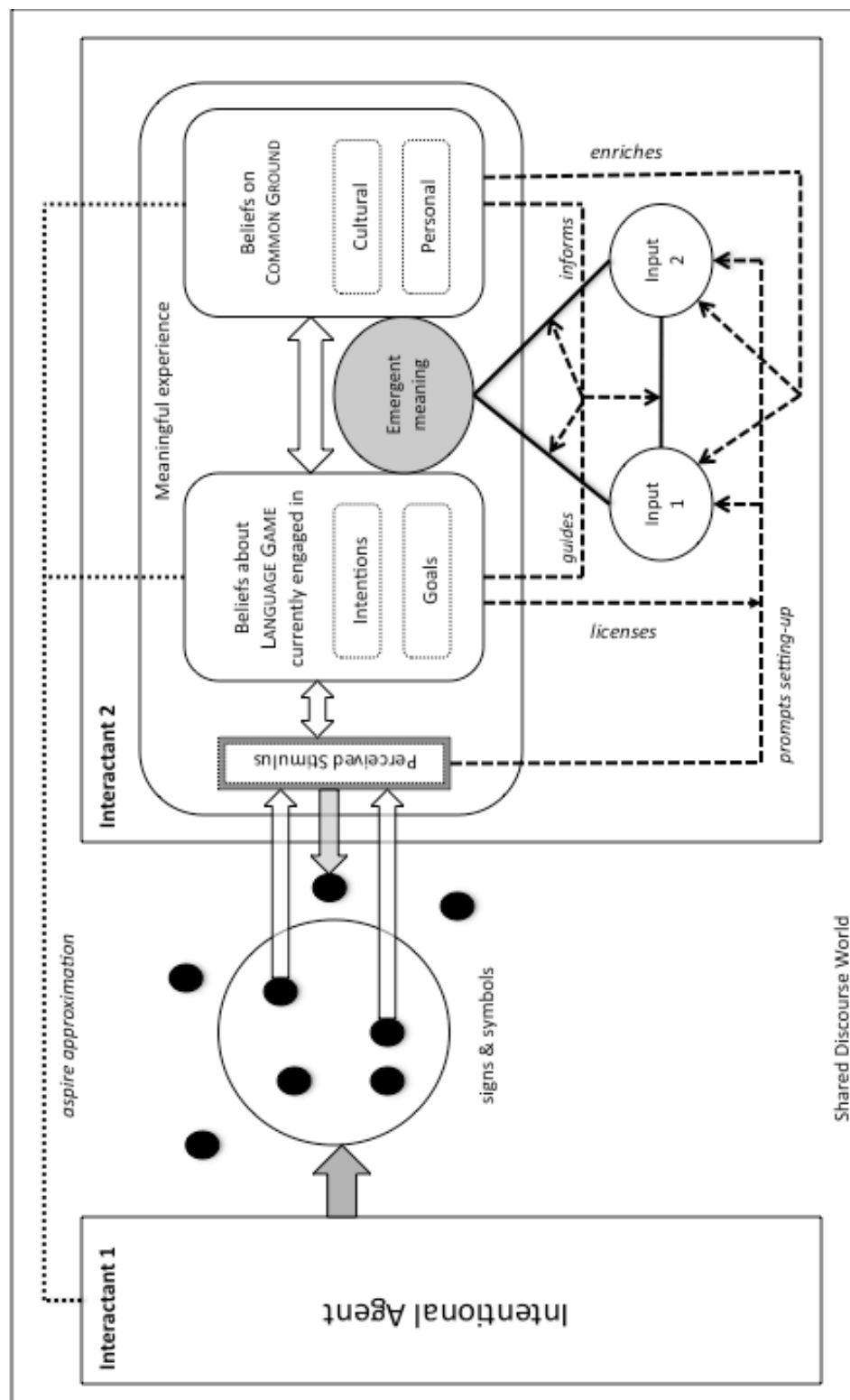


Figure 51: A discursively grounded model of in-situ meaning coordination

This model represents the processes understood to be in force on the side of the recipient (Interactant 2). Assuming their partners in interaction to be intentional agents capable of conducting similar cognitive operations, interactants aim at approximation in construal. The intersubjective coordination of meaning is conducted with the help of signs. For this reason, the signs and symbols⁸³ used in communication are represented in between the interactants. These include words, morphemes, prosody, gestures, and material objects (to name merely the most apparent ones discussed in chapter five). Not all signs, i.e. activities interactionally treated as signs are produced intentionally. Communicators designing their contributions also emit signs unintentionally. Blushing is a case in point. Furthermore, the blowing of one's nose' during a conversation may either be treated as a sign or simply ignored in our quest for joint meaning coordination. For this reason, the model represents some signs and symbols within a circle and some outside the circle. The latter are meant to illustrate that such signs do not necessarily need to be intended as signs and, secondly, that not all signs consciously emitted by one interactant need to be accepted by the other. Failing to notice an irony marker, such as a wink, would constitute such a 'neglected' sign.

Nevertheless, we assume our interactional partners to be intentional beings. We assume that the acts our interactional partners implement, the symbols they employ, must somehow refer to either their affective or mental states and/or elements which hold a particular degree of relevance or meaning for us (cf. Grice's (1975:46) Maxim of Relation that subsumes the axiom 'be relevant'). Why else would they employ them if not for others to notice and infer meaning? The major "evidence of [their] intention to convey a certain meaning" (Wilson & Sperber, 2006: 607) is, after all, the signs utilised by our communicative partners.

On what basis are the signs a) selected as potentially meaning-bearing, and b) attributed a 'referent' (i.e. meaning)? In the model I would like to propose, the network of signs we are confronted with is observed through a *filter* informed by beliefs about signs that can be expected to occur in the current interaction (and, vice versa, the signs inform beliefs as to what can legitimately be expected next). Consider a situation in which two conversational partners are communicating in English and the

⁸³ Under the term 'sign', Peirce subsumes indices, icons and symbols. Symbols are of importance to this model, as they refer to those signs "whose relation to their objects is an imputed character" (Peirce, 1868: Sec.14) and which are consequently subject to an inferencing process.

first partner abruptly switches to conversing in French. If this instance of code switching had not been expected, the second partner might be rather difficult to parse the linguistic layer, as the second partner was simply not prepared to hear another language within that particular conversation. The same applies to non-linguistic signs. It is only through this filter that signs become *stimuli*, to employ Sperber & Wilson's (1986) terminology.

Signs that were selected to become stimuli can serve to set-up of Mental Spaces, such as the spatial demarcation markers employed by impersonator BB in the preface to "The Proprietor". As interactants aim at arriving at joint construals, the stimuli might be produced in a more *ostensive* (ibid) fashion in order to coordinate the setting up of Mental Spaces. In "The Proprietor", BB stresses the adjective 'remote', and leaves a micro-pause before stating the location, the 'guest house'; this therefore places specific emphasis on the location and, in consequence, guides the audience to set up a new Mental Space. The 'content' of this Mental Space is continually evolving, and, for all participants, is informed by beliefs about shared cultural and personal Common Ground. The Mental Space is further constrained by beliefs about the *language game* currently being negotiated – the preface to BB's impersonation in "The Proprietor" would, after all, have awoken somewhat different expectations had it been uttered in, say, a business meeting and attributed to the CEO of a company.

In order to denote the social actions that guide localised meaning making, I employ the Wittgensteinian (1953) notion of 'language games' rather than social action, as the first has certain advantages over the latter. The term allows the following to be accounted for: a) conventionality and abstraction evident in the linguistic system, b) the rule-governed nature of language use, and c) the intrinsic link of meaning to action. Wittgenstein (1953:§7) defines language games "as the whole, consisting of language and the actions into which it is woven." In doing so, he places language games in between what de Saussure (1916) refers to as *langue* and *parole*, and thus manages to transcend this time-honoured dichotomy. Wittgenstein first of all considers the level of abstraction and convention in language usage – *langue*, in De Saussure's conception. After all, the original German text refers to *Sprachspiel* rather than *Sprechspiel* (Helmuth Feilke, personal communication). *Sprache* points to the systemic component of language, and the level of abstraction that leads to

conventionalised meanings associated with a given sign and structure.⁸⁴ Secondly, the term *game*⁸⁵ takes into account the rule-governed nature of language use, and emphasises the practices in which language is embedded. In this fashion, Wittgenstein ‘weaves’ the conventionalised meanings of language into communicative practices, in situ social action, and the larger situation in which it occurs. The “textile metaphor” (Friedrich, 2012) of *weaving* as employed by Wittgenstein in this respect places specific emphasis on ‘context’, i.e. the situatedness of linguistic systems in social action. The Saussurian *langue* and *parole* are not separate layers that can be investigated in isolation, but are interdependent and integrated – in short, they are interwoven. Pragmatics is inherent to our knowledge of the language system (cf. Mukherjee, 2004) and language is, for this reason, interlinked with action in language games. Employing the term *language game* rather than action alone consequently allows for the level of convention and abstraction underlying language use to be considered whilst simultaneously foregrounding the essential role that (beliefs on joint) action plays in meaning-making.

What, though, is ‘action’? Defining the term is notoriously difficult. The literature seems to agree that ‘proper’ action needs to be intentional and goal-oriented. In addition, some form of meta-awareness is required. For this reason, agents need to have a certain degree of control over their activities, and deliver them with a purpose in mind. Intention and goals are thus central components of action. While much research concerning the philosophy of action has focused on the individual, Bratman (1992) concentrates on cooperative activity and emphasises the shared nature of action, which fundamentally relies on “mutual responsiveness” (ibid:327). Agents here share intentions and coordinate their activities to achieve a joint goal. The coordination of activities contributing to the building of joint action depends on our ability to acknowledge others as intentional, thinking beings.⁸⁶

⁸⁴ cf. also Bühler’s (1934) notion of *Sprachgebilde* vs. *Sprechakt* and Chomsky’s (1965) rigid differentiation into *competence* and *performance*.

⁸⁵ The German term *Spiel* is broader than the English *game*, as it also incorporates aspects that would be subsumed under *play* in English. As Stern (2004:90) points out, the playful nature alluded to by Wittgenstein in his metaphor is, for example, lost in the English translation.

⁸⁶ This view departs slightly from canonical MSCI when Turner (2000:6) states that “[c]onceptual integration is path dependent: blends are put together and unpacked using what the understander already has in his conceptual structures.” Although Turner (2000) mentions in passing that the ‘sender’ might have a certain meaning in mind which is recipient-design (though he employs different terminology), he seems to identify the major cognitive load involved in meaning-making processes as lying on the ‘receiver’ rather than on all of the parties involved in the interaction.

As Croft (2009) points out, this highly complex process is facilitated by a number of devices. For one, joint attention, as for example established through linguistic devices or shared gaze and pointing, can act as a coordination device. Similarly, joint salience (i.e. features of context that are salient by way of human perceptual and cognitive capacities or shared cultural heritage) facilitates coordination between interactants. Thirdly, conventions help the individual to infer the intentions of the partner. Conventions comprise regularities in behaviour that are made use of recurrently in a community as a means of addressing repeated coordination problems. Meaning construction is facilitated by these coordination devices, yet starts with the basic assumption that interactants try to accommodate their communicative partners' needs (cf. Grice's (1975) *cooperative principle*).

This ability (cf. Theory of Mind) allows us to form predictions about the type of behaviour that can be expected from our interactants. Beliefs concerning the language game, i.e. the language system and the action it is integrated in consequently give rise to expectations both with regard to stimuli to be expected and the rules on what is normatively acceptable in a given discourse setting. It is hard to imagine that construing a blend such as "The Proprietor" would be acceptable in a more formal language game, such as, say, a court hearing. Knowledge of appropriateness is thus embedded in the language game.

Apart from licensing blending operations, beliefs about the language game furthermore *guide* mapping processes between the inputs set up by the stimuli and enriched via beliefs on Common Ground as well as shared world knowledge. The mappings are guided by coordinated beliefs with regard to relevance to the joint action currently being negotiated. In impersonations that are not prefaced in particular, the mappings are based solely on shared beliefs concerning joint action. In "Smug Jonas" and other mirror blends that mockingly reframe an interactant's prior contribution, correspondences are largely directed by beliefs about relevance. The impersonator does not explicitly name the role s/he is taking on, and the mapping is guided by beliefs about the language game of mocking ("What could s/he be doing now?"), while stimuli (such as the marked repeat of the prior speaker's utterance) serve as further evidence to those beliefs, which consequently strengthens the mappings. Mapping processes between input spaces and from the input to the blended space are thus subject to beliefs regarding the degree of relevance to a given language game, and are dependent on the interplay of certain stimuli and expectations.

Finally, the meaning emerging from conceptually blending the inputs is coordinated as being *relevant* to the language game and, within that language game, leads to what Wilson & Sperber (2006) term a “positive cognitive effect”, i.e. “a worthwhile difference to the individual’s representation of the world” (ibid:608). In many of the impersonations discussed in the previous chapters, this cognitive effect concerned inferences which provided assessments, or simply a pleasure in mimesis (i.e. the recognition of a referent and the ensuing celebration of shared knowledge and in-group status).

For this reason, the Blended Space is couched in between the fields denoting beliefs about the language game and beliefs regarding personal and cultural Common Ground. The meaning that intersubjectively emerges from interactional blending operations heavily depends on those two factors. Without considering the action currently being negotiated or the knowledge that is assumed to be shared between interactants, MSCI cannot faithfully claim to be a theory that a) transcends the semantic-pragmatic distinction, and b) is able to account for “emergent meaning” (Fauconnier & Turner, 2002:37). For this reason, “the way we think” is not conceptual blending, but rather the basic quest for meaning which is relevant to the social action currently being negotiated. This overarching cooperative quest for interactionally relevant meaning needs to be incorporated into cognitive semantic research.

6.2. Project summary

Conceptual Blending and Mental Spaces Theory has been proven to be a highly appealing theory for describing the cognitive processes which give rise to creative language usage. By bringing distinct semantic scenarios into mutual interaction, novel structure emerges which is fundamentally more than the sum of its parts and contains elements that are found in neither of the input spaces. MSCI is thus able to reconcile long-standing debates in metaphor theory and problems in semantics, such as indirect reference, presupposition and counterfactuals. Despite its intuitive attractiveness and promising ideas, MSCI has been criticised for a number of reasons. Most prominently, it has been accused of losing sight of the phenomenon, neglecting the

historicity, material and social dimension of situated cognising, and therefore not being able to account for contextualised meaning construction. MSCI thus appears to be largely rooted in ideational theories of meaning which neglect the social and action-based dimension of meaning-making (as emphasised for example by Wittgenstein (1953), Austin (1961), and Searle (1969)). MSCI is furthermore considered to be simply too complex, proposing artificial, post-hoc constructs on seemingly isolated minds. A case in point is the highly complex blending typology that lies at the heart of MSCI. This typology proposes a continuum of frame-integration complexity which cognisers have at their disposal; this ranges from simplex blends to double-scope blends. The authors claim that double-scope frame integration is the very factor which makes human achievements such as mathematics, arts and religion possible. Yet the typology proposed has so far not been tested on actual data, and has rightly been accused of constituting a post-hoc analytic category. MSCI's focus thus seems to lie on *competence* rather than *performance* despite aspiring to defy the semantic-pragmatic distinction (cf., e.g. Fauconnier, 2006). A usage-based approach to language – and Cognitive Linguistics in general claims to be rooted in the tradition (cf. e.g. Evans & Green, 2006; Geeraerts, 2005) – should, however, be able to overcome this “rigid Chomskyan dichotomy” (Mukherjee, 2004:96) between theoretical ability and genuine behaviour, and propose a “model of language cognition [...] able to account for actual usage” (ibid:87).

For this reason, this study aimed to investigate blending-in-interaction in order to

- a) trace the way in which conceptual blending is negotiated by interactants in situ, and
- b) determine the extent to which participants-in-interaction orient to the allegedly post-hoc blending typology proposed by MSCI's founding fathers.

In order to avoid the ‘ubiquity problem’ and the often-problematic delineation of Mental Spaces, one phenomenon that involves two semiotically distinct semantic scenarios was studied in detail. Impersonation humour requires the appropriation of voices other than one's own, and may therefore be considered to be an excellent example of conceptual blending that occurs on various levels of frame integration complexity from mirror- to double-scope blends.

When investigating blending-in-interaction, theoretical preconceptions that equip blending for the discursive approach taken here have to be developed. As such, it was first of all suggested that Mental Spaces in an interactional framework cannot be regarded as isolated conceptual packages in an individual's mind, but rather as interactionally displayed and locally negotiated in interaction. Secondly, it is assumed that new Mental Spaces are set up relative to a discourse base space that shares features with Langacker's (2001) Current Discourse Space; this Space subsumes assumptions concerning shared knowledge and beliefs about the current discourse setting and history. Thirdly, mapping processes in conceptual blending are regarded as interactionally coordinated.

Blending-in-interaction was studied by drawing on the methodological framework suggested by Ethnomethodological Conversation Analysis (EMCA). EMCA offers an apt array of tools to firstly discern the methods employed by interactants to jointly negotiate meaning which emerges from the integration of distinct semantic scenarios, and secondly to determine the extent to which participants in their interactional projects orient to the blending typology advanced in MSCl. The material was drawn from the British comedy panel game show *Never Mind the Buzzcocks* (NMTB). Impersonation sequences identified in shows aired between 2006 and 2009 were transcribed according to the standards set out by GAT. As the sequences stem from televised data, they constitute interaction-as-broadcasted (cf. Mondada, 2009), and were analysed with an eye on their constructed nature given that camera operators and editors serve as additional participants on the show.

The analyses primarily showed that an orientation to semantic worlds is evident in interaction. Gaze and postural orientation are, for example, discontinued when presenting a character that inhabits a distinct discourse scenario which is presented as separate from topical talk with regard to spatio-temporal, modal and epistemic properties. Interactants signal the distinct input scenarios contributing to impersonation blends in a methodological manner; however, in this respect they do not orient to frame-integration complexity. Single-scope and double-scope blends both are signalled drawing on the same interactional methods. In addition, postural orientation is maintained when attributing a role or properties of a character to a co-participant. For this reason, the orientation to the action currently being negotiated (e.g. attributing a deviant identity to a co-participant) appears to exert more influence

on performance than semantic distance. Although conceptual blending would indeed seem to be a cognitive tool, it is not a discursive practice in its own right.

It therefore became apparent that the performance of actions that rely on tools of conceptual blending is not guided by an orientation to frame blending, but by the language game (Wittgenstein, 1953) currently being coordinated. As a result, frame integration complexity is not a category that future studies should rely on. Rather, a distinct focus on the ‘doings’ which employ blending should be considered.

In my corpus, the language games impersonations contribute to in the corpus comprise the following:

- a) Assessment of prior contribution by co-participant to topical talk
 - a. Attribution of deviant identity to co-participant in interaction
 - b. Illustration of diegetically presented assessment of prior action
- b) Illustrations of joke-first narrations
 - a. Best-off visitations into diegetically set-up narrative scenarios
 - b. Polyphonically narrated joke-first sequences

For the reasons outlined, MSCI cannot be considered a theory of meaning-making that transcends the time-honoured division into semantics and pragmatics. Despite claiming that MSCI is able to account for “how meaning is constructed and how language prompts for meaning” (Fauconnier & Turner, 2002:147), the meanings resulting from the integration of distinct semantic scenarios do not relate to the process of conceptual integration. The meanings emerging from the joint coordination of frame integration are primarily guided by jointly negotiated assumptions regarding the language game currently being negotiated, and by the relevance which a given sign is taken to hold in relation to that particular language game. Conceptual blending therefore does not constitute the way in which we think in order to arrive at a meaningful experience, but is *a* tool that may be jointly consulted in our quest for interactionally relevant meanings. For MSCI to be a theory that transcends the semantics-pragmatics division and that is able to account for fully contextualised meanings, it will need to incorporate notions of Common Ground, language games (including social action), intention and relevance. It will also need to explicitly acknowledge the historicity inherent in meaning-making processes. One such model was suggested in the previous section. The model takes a serious stance as

to the key role that beliefs play regarding contextual relevance (i.e. relevance to jointly negotiated projects) in online, interactional meaning-construction. By incorporating notions of signs, stimuli, relevance, language games and Common Ground and regarding conceptual blending as a cognitive tool, it hopes to be able to account for the emergence of fully-contextualised meanings arising from jointly negotiated frame integration. The model thus accounts for the pivotal role which the ‘pragmatic’ (i.e. joint, cooperative action) plays in meaning-making. Thus, that which been deemed irrelevant (cf. competence – performance distinction) in previous research in cognitive sciences is actually revealed to exert considerable influence on cognition.

6.3. Prospects for further research

While this project proposed a model that “locate[s] meaning in communicating minds” (Brandt, 2010:495), it nevertheless only focuses on one cognitive semantic theory. Despite commendable attempts by eminent researchers such as Langacker (2001) and Croft (2009), other approaches within the discipline as a whole have thus far neglected the central role played by beliefs on joint action in meaning construction. For this reason, it would doubtless be interesting to develop a more overarching project that – in addition to emphasising the role of embodiment – also incorporated the inherently social nature of meaning-making in other cognitive semantic research enterprises. Such projects could investigate the way in which our social nature gives rise to image schemas, or how it influences issues in categorisation and polysemy. In this respect, further useful links to phenomenology could be developed.

Continuing in this vein, Cognitive Linguistics as a whole and Cognitive Semantics in particular would greatly benefit from a more explicit philosophy of language that goes beyond the level of embodiment, and takes into account prior debates and the historicity of the concepts discussed. Rather than focusing on an attempt to delineate Cognitive Linguistics by using analytic approaches to philosophy of language, the development of a distinctly cognitive-linguistic philosophy that is able to account for both the bodily and the social nature of language would certainly be worthy of merit. This cognitive-linguistic philosophy of language should furthermore

engage with prior and current debates in philosophy of language, such as those conducted in metaphorology, or studies on emergence and nature of meaning.

Within MSCI, further empirical and experiential studies would be required that strip the theory of its convoluted, highly-detailed, top-down principles and terminology, and focus on questions such as the following:

- *Psycholinguistic studies*
 - Development studies: does the ability to consciously integrate distinctly theatrical scenarios bear any relation to Theory of Mind development? Which role does the development of second-order cognition play in the ability to parse and produce blends that recruit semiotically distinct frames and rely on mappings occurring between them?
 - Clinical experiments: how and to what extent does metacognition impairment affect the ability to process and produce novel blends? To what extent does Theory of Mind impairment (through for example a unilateral frontal lobe lesion (cf. Rowe et al., 2001)) affect the production and parsing of blends?
- *Studies in experimental semantics:*
 - Eye-tracking studies could offer insight into processes of decompression in visual blends such as caricature.
 - Drawing experiments that provide subjects with a decontextualised neologism and which utilises conceptual blending (such as *land yacht*) could be used to extract details as to the information transfer occurring between domains and the resulting construal.
 - The role played by priming effects in the construal of such novel constructions could be investigated in order to shed more light on how mappings are guided by the quest for relevance.
 - Studies tracing eye movement could be employed to discover whether separate conceptual domains are accessed in the production and parsing of discursively dissonant blends.

MSCI in particular and Cognitive Semantics in general have enabled us to perceive meaning as a dynamic and essentially perceptual process. By taking into account the embodied nature of cognising, first-generation cognitive linguists overcame the rigid

constraints posited by analytic language philosophy and linguistic studies conducted in the tradition. In order to fully reclaim language and cognition as inherently human phenomena (and thus rejecting technical metaphors, such as the computing mind popular in the 20th century and evident in formal approaches), an explicit consideration of the social grounding of cognition, Theory of Mind and cooperation is required.

REFERENCES

- Allwood, J. (2008). Multimodal corpora. In A. Lüdeling & M. Kytö (Eds.), *Corpus Linguistics. An International Handbook* (Vol. 1, pp. 207-224). Berlin, New York: Mouton de Gruyter.
- Alter, J. (1981). From Text to Performance. Semiotics of Theatricality. *Poetics Today*, 2(3), 113-139.
- Aristotle (1996). *Poetics* (M. Heath, Trans.). London: Penguin.
- Aristotle (1989). *Metaphysics* (H. Tredennick, Trans.). Cambridge: Harvard University Press.
- Aristotle (1992). *The Art of Rhetoric* (H. Lawson-Tancred, Trans. 3rd ed.). London: Penguin.
- Astington, J. W. (2001). The Future of Theory-of-Mind Research: Understanding Motivational States, the Role of Language, and Real-World Consequences. *Child Development*, 72(3), 685-687.
- Atkinson, J. M. (1984). Public speaking and audience responses: some techniques for inviting applause. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Interaction* (pp. 370-410). Cambridge: Cambridge University Press.
- Attardo, S. (1994). *Linguistic Theories of Humor*. Berlin; New York: Mouton de Gruyter.
- Attardo, S. (2001). *Humorous Texts: A Semantic and Pragmatic Analysis*. Berlin; New York: Mouton de Gruyter.
- Austin, J. L. (1962). *How to do things with words*. Oxford: Oxford University Press.
- Bache, C. (2005). Constraining conceptual integration theory: Levels of blending and disintegration. *Journal of Pragmatics*, 37, 1615-1653.
- Bakhtin, M. ((1963)1984). *Problems of Dostoevsky's Poetics* (C. Emerson, Trans.). Minneapolis: University of Minnesota Press.
- Bateson, G. ((1955)1972). A Theory of Play and Fantasy, *Steps to an Ecology of Mind. Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology* (pp. 177-193). Northvale, NJ: Jason Aronson.
- Bergmann, J. R. (2004a). Conversation Analysis. In U. Flick & E. v. Kardorff & I. Steinke (Eds.), *A Companion to Qualitative Research* (pp. 296-302). London; Thousand Oaks; New Delhi: Sage.

- Bergmann, J. R. (2004b). Ethnomethodology. In U. Flick & E. v. Kardorff & I. Steinke (Eds.), *A Companion to Qualitative Research* (pp. 72-80). London; Thousand Oaks; New Delhi: Sage.
- Bergmann, J. R. (2004c). Harold Garfinkel and Harvey Sacks. In U. Flick & E. v. Kardorff & I. Steinke (Eds.), *A Companion to Qualitative Research* (pp. 29-34). London; Thousand Oaks; New Delhi: Sage.
- Bergson, H. (1900). *Le rire. Essai sur la signification du comique*. Paris: Alcan.
- Black, M. (1962). *Models and Metaphors*. Ithaca, NY: Cornell University Press.
- Bolter, J. D., & Grusin, R. (1999). *Remediation. Understanding New Media*. Cambridge: The MIT Press.
- Bourdieu, P. ((1972)1977). *Outline of a Theory of Practice* (R. Nice, Trans.). Cambridge: Cambridge University Press.
- Bousmalis, K., Morency, L., & Pantic, M. (2011). Modeling Hidden Dynamics of Multimodal Cues for Spontaneous Agreement and Disagreement Recognition, *IEEE International Conference on Automatic Face and Gesture Recognition* (pp. 746-752); IEEE Xplore Digital Library.
- Boxer, D., & Cortés-Conde, F. (1997). From bonding to biting: Conversational joking and identity display. *Journal of Pragmatics*, 27, 275-294.
- Brandt, L. (2008). A semiotic approach to fictive interaction as a representational strategy in communicative meaning construction. In T. Oakley & A. Hougaard (Eds.), *Mental Spaces in Discourse and Interaction* (pp. 109-148). Amsterdam: John Benjamins.
- Brandt, L. (2010). *Language and enunciation - A cognitive inquiry with special focus on conceptual integration in semiotic meaning construction*. Unpublished doctoral dissertation, Aarhus Universitet.
- Brandt, L., & Brandt, P. A. (2005). Making sense of a blend. A cognitive-semiotic approach to metaphor. *Annual Review of Cognitive Linguistics*, 3, 216-249.
- Brandt, P. A. (2005). Mental spaces and cognitive semantics: a critical comment. *Journal of Pragmatics*, 37, 1578-1594.
- Bratman, M. E. (1992). Shared Cooperative Activity. *The Philosophical Review*, 101(2), 327-341.
- Brône, G. (2008). Hyper- and misunderstanding in interactional humor. *Journal of Pragmatics* 40(12). 2027-2061.
- Brône, G. (2010). *Bedeutungskonstitution in verbalem Humor. Ein kognitiv-linguistischer und diskurssemantischer Ansatz*. Frankfurt: Peter Lang.

- Brône, G., & Feyaerts, K. (2003). The cognitive linguistics of incongruity resolution: Marked reference-point structures in humor, *8th International Cognitive Linguistics Conference*. Logroño: Universidad de La Rioja. Electronic copy available at <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.111.4193&rep=rep1&type=pdf> [last accessed 19 Feb 2012].
- Brown, P., & Levinson, S. C. (1987). *Politeness. Some universals in language usage*. (4 ed.). Cambridge: Cambridge University Press.
- Bublitz, W. (1988). *Supportive Fellow-Speakers and Cooperative Conversations. Discourse Topics and Topical Actions, Participant Roles and 'Recipient Action' in a Particular Type of Everyday Conversation*. Amsterdam; Philadelphia: John Benjamins.
- Buchstaller, I. (2006). Diagnostics of age-graded linguistic behaviour: The case of the quotative system. *Journal of Sociolinguistics*, 10, 3-30.
- Buchstaller, I. (2008). The localization of global linguistic variants. *English World-Wide*, 29(1), 15-44.
- Bücker, J. (fthc). Doing doing talking and bilateral positioning: Towards a linguistic narratology of direct speech in everyday story telling. Manuscript.
- Bühler, K. ((1934)1999). *Sprachtheorie: die Darstellungsfunktion der Sprache*. (3rd ed.). Stuttgart: Fischer.
- Carston, R. (2002). Linguistic Meaning, Communicated Meaning and Cognitive Pragmatics. *Mind & Language*, 17(1-2), 127-148.
- Chilton, P. (2008). Reflections on blends and discourse. In T. Oakley & A. Hougaard (Eds.), *Mental Spaces in Discourse and Interaction* (pp. 251-256). Amsterdam: John Benjamins.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Clark, H. H. (1982). Hearers and Speech Acts. *Language*, 58(2), 332-373.
- Clark, H. H. (1996). *Using Language*. Cambridge: Cambridge University Press.
- Clark, H. H., & Gerrig, R. J. (1990). Quotations as Demonstrations. *Language*, 66(4), 764-805.
- Coulson, S. (2000). *Semantic Leaps: Frame-Shifting and Conceptual Blending in Meaning Construction*. Cambridge: Cambridge University Press.
- Coulson, S. (2005). Extemporaneous blending: Conceptual integration in humorous discourse from talk radio. *Style* 39(2), 107-122.
- Coulson, S. (2007). Electrifying results. ERP data and cognitive linguistics. In M. Gonzalez-Marquez & I. Mittelberg & S. Coulson & M. J. Spivey (Eds.), *Methods in Cognitive Linguistics* (pp. 400-423). Amsterdam, Philadelphia: John Benjamins.

- Coulson, S., & Matlock, T. (2005). Cognitive Science. In J. Verschueren & J.-O. Östman (Eds.), *Handbook of Pragmatics. 2003-2005 Installment* (Vol. 9, pp. 86-109). Amsterdam: John Benjamins.
- Coulson, S., & Oakley, T. (2000). Blending basics. *Cognitive Linguistics*, 11(3/4), 175-196.
- Couper-Kuhlen, E. (1986). *An Introduction to English Prosody*. Tübingen: Max Niemeyer.
- Couper-Kuhlen, E. (1999). Coherent voicing. On prosody in conversational reported speech. In W. Bublitz & U. Lenk & E. Ventola (Eds.), *Coherence in Spoken and written Discourse*. Amsterdam: John Benjamins.
- Couper-Kuhlen, E. (2001). Interactional prosody: High onsets in reasons-for-the-call turns. *Language in Society*, 30(1), 29-53.
- Couper-Kuhlen, E. (2009). Prosody. In S. D'hondt & J.-O. Östman & J. Verschueren (Eds.), *The Pragmatics of Interaction* (pp. 174-189). Amsterdam; Philadelphia: John Benjamins.
- Croft, W. (2009). Toward a social cognitive linguistics. In V. Evans & S. Pourcel (Eds.), *New Directions in Cognitive Linguistics* (pp. 395-420). Amsterdam: John Benjamins.
- Cruttenden, A. (2007). Intonational diglossia: a case study of Glasgow. *Journal of the International Phonetic Association*, 37(2), 257-274.
- D'Andrade, R. G. (1984). Cultural Meaning Systems. In R. A. Shweder & R. A. Levine (Eds.), *Culture theory. Essays on Mind, Self, and Emotion* (pp. 88-119). Cambridge: Cambridge University Press.
- Dancygier, B. (2008). The text and the story: Levels of blending in fictional narratives. In T. Oakley & A. Hougaard (Eds.), *Mental Spaces in Discourse and Interaction* (pp. 51-78). Amsterdam: John Benjamins.
- Davidson, D. (1973). Radical Interpretation. *Dialectica*, 27(3-4), 313-328.
- Davies, C. (1982). Ethnic jokes, moral values and social boundaries. *British Journal of Sociology*, 33(3), 383-403.
- Davies, C. (1998). *Jokes and their Relation to Society*. Berlin: Mouton de Gruyter.
- Davies, C. (2005). European ethnic scripts and the translation and switching of jokes. *Humor. International Journal of Humor Research*, 18(2), 147-160.
- de Kok, B. C. (2008). The role of context in conversation analysis: Reviving an interest in ethno-methods. *Journal of Pragmatics*, 40, 886-903.
- Dobrovolskij, D. (2011). The structure of metaphor and idiom semantics (a cognitive approach). In S. Handl & H.-J. Schmid (Eds.), *Windows to the Mind. Metaphor, Metonymy and Conceptual Blending* (pp. 41-62). Berlin; New York: De Gruyter Mouton.

- Drew, P. (1992). Contested evidence in a courtroom cross examination: the case of a trial for rape. In P. Drew & J. Heritage (Eds.), *Talk at Work: Social Interaction in Institutional Settings* (pp. 470-520). Cambridge: Cambridge University Press.
- Drew, P. (2006). Mis-alignments between caller and doctor in 'out-of-hours' telephone calls to a British GP's practice. In J. Heritage & D. Maynard (Eds.), *Communication in Medial Care: Interaction Between Primary Care Physicians and Patients* (pp. 416-444). Cambridge: Cambridge University Press.
- Drew, P., & Heritage, J. (2006). Editors' Introduction. In P. Drew & J. Heritage (Eds.), *Conversation Analysis* (Vol. 1, pp. xxi-xxxvii). London: Sage.
- Drew, P., & Heritage, J. (Eds.). (1992). *Talk at work: interaction in institutional settings*. Cambridge: Cambridge University Press.
- Ehrlich, S. (2002). Legal Institutions, Nonspeaking Reciprocity and Participants? Orientations. *Discourse & Society*, 13, 731-747.
- Ehrlich, S. (2006). Trial Discourse and Judicial Decision-Making: Constraining the Boundaries of Gendered Identities. In J. Baxter (Ed.), *Speaking Out: The Female Voice in Public Contexts* (pp. 139-158). London: Palgrave Macmillan.
- Erickson, F. (1992). They know all the lines: rhythmic organization and contextualization in a conversational listing routine. In P. Auer & A. di Luzio (Eds.), *The Contextualisation of Language* (pp. 365-397). Amsterdam: John Benjamins.
- Evans, V., & Green, M. (2006). *Cognitive Linguistics. An Introduction*. Edinburgh: Edinburgh University Press.
- Fauconnier, G. (1984). *Espaces mentaux. Aspects de la construction du sens dans les langues naturelles*. Paris: Ed. de Minuit.
- Fauconnier, G. (1994). *Mental Space: Aspects of Meaning Construction in Natural Language*. Cambridge: The MIT Press.
- Fauconnier, G. (1997). *Mappings in Thought and Language*. Cambridge: Cambridge University Press.
- Fauconnier, G. (1999). Methods and generalizations. In T. Janssen & G. Redeker (Eds.), *Cognitive Linguistics: Foundations, Scope, and Methodology* (pp. 98-128). Berlin; New York: Mouton de Gruyter.
- Fauconnier, G. (2006). Pragmatics and Cognitive Linguistics. In L. R. Horn & G. Ward (Eds.), *The Handbook of Pragmatics* (pp. 657-674). Malden; Oxford; Victoria: Blackwell.
- Fauconnier, G., & Lakoff, G. (fthc). On Metaphor and Blending. *Cognitive Semiotics* (Special Issue on Conceptual Metaphor Theory: Thirty years after).

- Fauconnier, G., & Turner, M. (1994). Conceptual Projection and Middle Spaces, *UCSD Department of Cognitive Science Technical Report 9401*. Electronic copy available at <http://www.cogsci.ucsd.edu/research/documents/technical/9401.pdf> [last accessed 12 Feb 2012]
- Fauconnier, G., & Turner, M. (1996). Blending as a Central Process of Grammar. In A. E. Goldberg (Ed.), *Conceptual Structure, Discourse, and Language* (pp. 113-130). Stanford: CSLI Publications.
- Fauconnier, G., & Turner, M. (1998). Principles of Conceptual Integration. In J.-P. König (Ed.), *Discourse and Cognition* (pp. 269-283). Stanford: CSLI Publications.
- Fauconnier, G., & Turner, M. (2002). *The Way We Think. Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.
- Feldman, J. A. (2006). *From Molecule to Metaphor. A Neural Theory of Language*. Cambridge: The MIT Press.
- Ferrara, K. (1997). Form and function of the discourse marker anyway: implications for discourse analysis. *Linguistics*, 35(2), 343-378.
- Ferrara, K., & Bell, B. (1995). Sociolinguistic Variation and Discourse Function of Constructed Dialogue Introducers: The Case of Be + Like. *American Speech*, 70(3), 265-290.
- Fillmore, C. (1982). Frame Semantics. In The Linguistic Society of Korea (Ed.), *Linguistics in the Morning Calm* (pp. 111-137). Seoul: Hanshin.
- Fine, G. A., & DeSoucey, M. (2005). Joking cultures and social regulation. *Humor. International Journal of Humor Research*, 18(1), 1-22.
- Fischer-Lichte, E. (2002). Grenzgänge und Tauschhandel. Auf dem Weg zu einer performativen Kultur. In U. Wirth (Ed.), *Performanz. Zwischen Sprachphilosophie und Kulturwissenschaften*. Frankfurt: suhrkamp taschenbuch wissenschaft.
- Fiske, J. (1987). *Television Culture*. London: Routledge.
- Flick, U., Kardorff, E. v., & Steinke, I. (2004). What is Qualitative Research? An Introduction to the Field. In U. Flick & E. v. Kardorff & I. Steinke (Eds.), *A Companion to Qualitative Research* (pp. 3-11). London; Thousand Oaks; New Delhi: Sage.
- Fludernik, M., Freeman, D. C., & Freeman, M. H. (1999). Metaphor and Beyond: An Introduction. *Poetics Today*, 20(3: Metaphor and Beyond: New Cognitive Developments), 383-396.
- Fox, B. A., & Robles, J. (2010). It's like mmm: Enactments with it's like. *Discourse Studies*, 12(6), 715-738.
- Freud, S. (1905). *Der Witz und seine Beziehung zum Unterbewussten*. Frankfurt: Fischer.

- Friedman, N. (1975). *Form and Meaning in Fiction*. Athens: University of Georgia Press.
- Friedrich, A. (2012). *Im Netz der Metapher. Zur Theorie kultureller Leitmetaphern*. Unpublished doctoral dissertation, Justus-Liebig-Universität, Gießen.
- Garfinkel, H. (1963). A Conception Of, and Experiments With, 'Trust' as a Condition of Stable Concerted Actions. In O. J. Harvey (Ed.), *Motivation and Social Interaction* (pp. 187-238). New York: Free Press.
- Garfinkel, H. (1967). *Studies in Ethnomethodology*. Malden; Oxford: Blackwell.
- Geeraerts, D. (2005). Lectal variation and empirical data in Cognitive Linguistics. In F. J. Ruiz de Mendoza Ibáñez & M. S. Peña Cervel (Eds.), *Cognitive linguistics: internal dynamics and interdisciplinary interaction* (pp. 163-190). Berlin: Mouton de Gruyter.
- Geertz, C. (1973). *The Interpretation of Cultures*. New York: Basic Books.
- Gibbs, R. W. (2000). Making good psychology out of blending theory. *Cognitive Linguistics*, 11(3/4), 347-358.
- Giere, R. N., & Moffatt, B. (2003). Distributed Cognition: Where the Cognitive and Social Merge. *Social Studies of Science*, 33(2), 301-310.
- Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*, 8(3), 183-206.
- Giora, R. (1999). On the priority of salient meanings: Studies of literal and figurative language. *Journal of Pragmatics*, 31, 919-929.
- Giora, R. (2003). *On our Mind: Salience, Context and Figurative Language*. New York: Oxford University Press.
- Givón, T. (2005). *Context as Other Minds. The Pragmatics of Sociality, Cognition and Communication*. Amsterdam; Philadelphia: John Benjamins.
- Glaznieks, A. (2011). Why focus on target domains? The importance of domain knowledge in children's understanding of metaphors. In S. Handl & H.-J. Schmid (Eds.), *Windows to the Mind. Metaphor, Metonymy and Conceptual Blending* (pp. 63-84). Berlin; New York: De Gruyter Mouton.
- Glenn, P. (2003). *Laughter in Interaction*. Cambridge: Cambridge University Press.
- Glick, D. J. (2007). Some performative techniques of stand-up comedy: An exercise in the textuality of temporalization. *Language & Communication*, 27, 291-306.
- Goffman, E. (1974). *Frame Analysis. An Essay on the Organization of Experience*. Boston: Northeastern University Press.
- Goffman, E. (1981). *Forms of Talk*. Oxford: Blackwell.
- Goffman, E. (1983). The Interaction Order. *American Sociological Review* 48, 1-17.

- Golato, A. (2000). An innovative German quotative for reporting on embodied actions: Und ich so/und er so 'and I'm like/and he's like'. *Journal of Pragmatics*, 32, 29-54.
- Golato, A. (2002). Grammar and Interaction: Reported Discourse and Subjunctive in German. *Zeitschrift für Sprachwissenschaft*, 21(1), 25-56.
- Goldberg, A. E. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
- Goldberg, A. E. (1999). The Emergence of Argument Structure Semantics. In B. MacWhinney (Ed.), *The Emergence of Language* (pp. 197-212). Hillsdale: Lawrence Erlbaum Publications.
- Gonzalez-Marquez, M., Mittelberg, I., Coulson, S., & Spivey, M. J. (Eds.). (2006). *Methods in Cognitive Linguistics*. Amsterdam, Philadelphia: John Benjamins.
- Goodwin, C. (1980). Restarts, Pauses, and the Achievement of a State of Mutual Gaze at Turn-Beginning. *Sociological Inquiry*, 50(3-4), 272-302.
- Goodwin, C. (1981). *Conversational Organization: Interaction Between Speakers and Hearers*. New York: Academic Press.
- Goodwin, C. (1984). Notes on story structure and the organization of participation. In J. M. Atkinson & J. Heritage (Eds.), *Structure of social action* (pp. 225-246). Cambridge: Cambridge University Press.
- Goodwin, C. (1986). Gesture as resource for the organization of mutual orientation. *Semiotica*, 62(1-2), 29-49.
- Goodwin, C. (1994). Professional Vision. *American Anthropologist*, 96(3), 606-633.
- Goodwin, C. (2000a). Action and embodiment within situated human interaction. *Journal of Pragmatics*, 32, 1489-1522.
- Goodwin, C. (2000b). Practices of Seeing: Visual Analysis: An Ethnomethodological Approach. In T. Van Leeuwen & C. Jewitt (Eds.), *Handbook of Visual Analysis* (pp. 157-182). London: Sage Publications.
- Goodwin, C. (2003a). Pointing as Situated Practice. In S. Kita (Ed.), *Pointing: Where Language, Culture and Cognition Meet* (pp. 217-241). Mahwah: Lawrence Erlbaum.
- Goodwin, C. (2003b). The Body in Action. In J. Coupland & R. Gwyn (Eds.), *Discourse, the Body and Identity* (pp. 19-42). New York: Palgrave/Macmillan.
- Goodwin, C., & Goodwin, M. H. (1992). Assessments and the construction of context. In A. Duranti & C. Goodwin (Eds.), *Rethinking Context: Language as an Interactive Phenomenon* (pp. 151-189). Cambridge: Cambridge University Press.
- Grady, J. (2000). Cognitive mechanisms of conceptual integration. *Cognitive Linguistics*, 11(3/4), 335-345.

- Greatbatch, D. (1988). A Turn-Taking System for British News Interviews. *Language in Society*, 17(3), 401-430.
- Grice, P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics* (Vol. 3: Speech Acts, pp. 43-58). New York: Academic Press.
- Gruner, C. R. (1997). *The Game of Humor: A comprehensive Theory of Why We Laugh*. Brunswick: Transaction Publishers.
- Gülich, E., & Quasthoff, U. (1986). Story-Telling in Conversation. Cognitive and Interactive Aspects. *Poetics* 15, 217-241.
- Günthner, S. (1999). Polyphony and the 'layering of voices' in reported dialogues: An analysis of the use of prosodic devices in everyday reported speech. *Journal of Pragmatics*, 31, 685-708.
- Habib, R. (2008). Humor and disagreement: Identity construction and cross-cultural enrichment. *Journal of Pragmatics*, 40, 1117-1145.
- Harder, P. (2003). Mental Spaces: Exactly when do we need them? *Cognitive Linguistics*, 14(1), 91-96.
- Harder, P. (2007). Cognitive Linguistics and Philosophy. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford Handbook of Cognitive Linguistics* (pp. 1241-1265). Oxford: Oxford University Press.
- Harder, P. (2010). *Meaning in Mind and Society. A Functional Contribution to the Social Turn in Cognitive Linguistics*. Berlin; New York: De Gruyter Mouton.
- Haugh, M. (2010). Jocular mockery, (dis)affiliation, and face. *Journal of Pragmatics*, 42(8), 2106-2119.
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science*, 298, 1569-1579.
- Heath, C., & Hindmarsh, J. (2002). Analysing interaction: Video ethnography and situated conduct. In T. May (Ed.), *Qualitative Research in Action* (pp. 99-121). London: Sage.
- Heath, C., & Luff, P. (2006). Video Analysis and Organisational Practice. In H. Knoblauch & B. Schnettler & J. Raab & H.-G. Soeffner (Eds.), *Video-Analysis. Methodology and Methods*. London: Peter Lang.
- Heath, C., & Luff, P. (2007a). Gesture and institutional interaction: figuring bids in auctions of fine art and antiques. *Gesture*, 7(2), 215-241.
- Heath, C., & Luff, P. (2007b). Ordering competition: the interactional accomplishment of the sale of fine art and antiques at auction. *British Journal of Sociology*, 58(1), 63-85.
- Heath, C., & Luff, P. (2010). The Strike of a Hammer: Institutional Talk and 'Multi-Modal' Interaction. Paper presented at ICCA 10, Mannheim.

- Heritage, J. (1984). A change-of-state token and aspects of its sequential placement. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action: Studies in Conversation Analysis* (pp. 299-345). Cambridge: Cambridge University Press.
- Heritage, J. (1985). Analyzing News Interviews: Aspects of the Production of Talk for an "Overhearing" Audience. In T. A. Van Dijk (Ed.), *Handbook of Discourse Analysis* (Vol. 3: Discourse and Dialogue, pp. 95-117). London: Academic Press.
- Heritage, J. (1995). Conversation Analysis: Methodological Aspects. In U. Quasthoff (Ed.), *Aspects of Oral Communication* (pp. 391-418). Berlin: De Gruyter.
- Heritage, J. (2004). Conversation Analysis and Institutional Talk. In R. Sanders & K. L. Fitch (Eds.), *Handbook of Language and Social Interaction* (pp. 184-202). Mahwah: Erlbaum.
- Heritage, J. (2010). Questioning in Medicine. In A. Freed & S. Ehrlich (Eds.), *"Why Do You Ask?": The Function of Questions in Institutional Discourse* (pp. 42-68). New York: Oxford University Press.
- Heritage, J., & Atkinson, J. M. (1984). Introduction. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action. Studies in Conversation Analysis* (pp. 1-16). Cambridge: Cambridge University Press.
- Heritage, J., & Goodwin, C. (1990). Conversation Analysis. *Annual Review of Anthropology*, 19, 283-307.
- Heritage, J., & Greatbatch, D. (1986). Generating Applause: A Study of Rhetoric and Response at Party Political Conferences. *American Journal of Sociology*, 92, 110-157.
- Heritage, J., & Maynard, D. (Eds.). (2006). *Communication in Medical Care: Interaction Between Primary Care Physicians and Patients*. Cambridge: Cambridge University Press.
- Heritage, J., & Roth, A. (1995). Grammar and Institution: Questions and Questioning in the Broadcast News Interview. *Research on Language and Social Interaction*, 28(1), 1-60.
- Herrmann, E., Call, J., Hernández-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans Have Evolved Specialized Skills of Social Cognition: The Cultural Intelligence Hypothesis. *Science*, 317, 1360-1366.
- Holland, P. (2000). *The Television Handbook* (2nd ed.). London; New York: Routledge.
- Holmes, S. (2006). "The Question Is - Is It All Worth Knowing?" The cultural circulation of the early British quiz show. *Media, Culture & Society*, 29(1), 53-74.
- Hopper, R. (1990). Ethnography and conversation analysis after talking culture. *Research on Language and Social Interaction*, 24(1-4), 161-171.

- Hopper, R. (2005). A Cognitive Agnostic in Conversation Analysis: When Do Strategies Affect Spoken Interaction? In J. Potter & H. Te Molder (Eds.), *Conversation and Cognition* (pp. 134-158). Cambridge: Cambridge University Press.
- Hougaard, A. (2004). *"How're we doing?": An Interactional Approach to Cognitive Processes of Online Meaning Construction*. Unpublished doctoral dissertation, University of Southern Denmark, Odense.
- Hougaard, A. (2005). Conceptual disintegration and blending in interactional sequences: A discussion of new phenomena, processes vs. products, and methodology. *Journal of Pragmatics*, 37, 1653-1685.
- Hougaard, A. (2008). Compression in interaction. In T. Oakley & A. Hougaard (Eds.), *Mental Spaces in Discourse and Interaction* (pp. 179-208). Amsterdam: John Benjamins.
- Hougaard, A., & Hougaard, G. R. (in preparation). Fused bodies: on the interrelatedness of cognition and interaction, *Body - Language - Communication / Körper - Sprache- und Kommunikation. Vol. 1. Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science*. Berlin, New York: Mouton de Gruyter.
- Hougaard, A., & Oakley, T. (2008). Mental spaces and discourse analysis. In T. Oakley & A. Hougaard (Eds.), *Mental Spaces in Discourse and Interaction* (pp. 1-26). Amsterdam: John Benjamins.
- Hougaard, G. R. (2008). "Mental spaces" and "blending" in discourse and interaction: A response. In T. Oakley & A. Hougaard (Eds.), *Mental Spaces in Discourse and Interaction* (pp. 247-250). Amsterdam: John Benjamins.
- Houtkoop, H., & Mazeland, H. (1985). Turns and discourse units in everyday conversation. *Journal of Pragmatics*, 9, 595-619.
- Huizinga, J. (1949). *Homo Ludens: A Study of the Play-Element in Culture*. London: Routledge.
- Hutchby, I., & Woffitt, R. (1998). *Conversation Analysis. Principles, Practices and Applications*. Cambridge: Polity Press.
- Hutchins, E. (1995). *Cognition in the Wild*. Cambridge: The MIT Press.
- Hutchins, E. (2005). Material anchors for conceptual blends. *Journal of Pragmatics*, 37, 1555-1577.
- Hutchins, E., & Palen, L. (1997). Constructing meaning from space, gesture, and speech. In L. B. Resnick & R. Säljö & C. Pontecorvo & B. Burge (Eds.), *Discourse, Tools, and Reasoning: Essays on Situated Cognition*. Berlin, Heidelberg, New York: Springer.

- Jefferson, G. (1978). Sequential aspects of storytelling in conversation. In J. Schenkein (Ed.), *Studies in the organization of conversational interaction* (pp. 219-248). New York: Academic Press.
- Jefferson, G. (1979). A technique for inviting laughter and its subsequent acceptance/declination. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 79-96). New York: Irvington Publishers.
- Jefferson, G. (1990). List construction as a task and resource. In G. Psathas (Ed.), *Interactional competence* (pp. 63-92). New York: Irvington Publishers.
- Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. H. Lerner (Ed.), *Conversation Analysis: Studies from the First Generation* (pp. 13-31). Amsterdam: John Benjamins.
- Jefferson, G., Sacks, H., & Schegloff, E. A. (1977). Preliminary notes on the sequential organization of laughter. *Pragmatics Microfiche*. Cambridge: Cambridge University Department of Linguistics. Electronic copy available at www.liso.ucsb.edu [last accessed 12 Feb 2012].
- Jefferson, G., Sacks, H., & Schegloff, E. A. (1987). Notes on Laughter in the Pursuit of Intimacy. In G. Button & J. R. E. Lee (Eds.), *Talk and social Organisation* (pp. 152-205). Clevedon: Multilingual Matters.
- Johnstone, B. (1987). 'He says ... so I said': Verb tense alternation and narrative depictions of authority in American English. *Linguistics*, 25, 33-52.
- Kangasharju, H. (1996). Aligning as a team in multiparty conversation. *Journal of Pragmatics*, 26, 291-319.
- Kendon, A. (1990a). *Conducting Interaction: Patterns of Behavior in Focused Encounters*. Cambridge: Cambridge University Press.
- Kendon, A. (1990b). Gesticulation, quotable gestures and signs. In M. Moerman & M. Nomura (Eds.), *Culture embodied* (pp. 53-77). Osaka: National Museum of Ethnography.
- Kendon, A. (1997). Gesture. *Annual Review of Anthropology*, 26, 109-128.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Kendon, A. (2008). Some reflections on the relationship between 'gesture' and 'sign'. *Gesture*, 8(3), 348-366.
- Knoblauch, H., Schnettler, B., & Raab, J. (2006). Video-Analysis. Methodological Aspects of Interpretive Audiovisual Analysis in Social Research. In H. Knoblauch & B. Schnettler & J. Raab & H.-G. Soeffner (Eds.), *Video Analysis: Methodology and Methods. Qualitative Audiovisual Data Analysis in Sociology*. Frankfurt a.M.: Peter Lang.
- Koestler, A. (1964). *The Act of Creation*. London: Hutchinson.

- Kohn, M. E., & Franz, H. A. (2009). Localized Patterns for Global Variants: The Case of Quotative Systems of African American and Latino Speakers. *American Speech*, 84(3), 259-297.
- Kok, S., & Bublitz, W. (2011). Conceptual blending, evaluation and common ground: George W. Bush and Saddam as friends or foe? In H.-J. Schmid & S. Handl (Eds.), *Windows to the Mind: Metaphor, Metonymy and Conceptual Blending* (pp. 291-309). Berlin: De Gruyter Mouton.
- Kotthoff, H. (1998). *Spaß Verstehen. Zur Pragmatik von konversationellem Humor*. Tübingen: Niemeyer.
- Kotthoff, H. (2002). Irony, citation, and other forms of staged intertextuality. In C. F. Graumann & W. Kallmeyer (Eds.), *Perspectives in Discourse* (pp. 201-233). Amsterdam: John Benjamins.
- Kotthoff, H. (2004). Overdoing culture? Sketch-Komik, Typenstilisierung und Identitätskonstruktion bei Kaya Yanar. In K. Hörning & J. Reuter (Eds.), *Doing culture. Neue Positionen zum Verhältnis von Kultur und sozialer Praxis* (pp. 184-200). Bielefeld: Transcript-Verlag.
- Kotthoff, H. (2006). Pragmatics of performance and the analysis of conversational humor. *Humor. International Journal of Humor Research*, 19(3), 271-304.
- Kotthoff, H. (2007). Oral Genres of Humor: On the Dialectic of Genre Knowledge and Creative Authoring. *Pragmatics* 17(2), 263-296.
- Kotthoff, H. (2009). An interactional approach to irony development. In N. R. Norrick & D. Chiaro (Eds.), *Humor in Interaction* (pp. 49-77). Amsterdam: John Benjamins.
- Kövecses, Z. (2010). *Metaphor. A Practical Introduction* (2nd ed.). Oxford: Oxford University Press.
- Koven, M. (2002). An analysis of speaker role inhabitation in narratives of personal experience. *Journal of Pragmatics*, 34, 167-217.
- Kowal, S., & O'Connell, D. (2004). The Transcription of Conversations. In U. Flick & E. v. Kardorff & I. Steinke (Eds.), *A Companion to Qualitative Research*. London; Thousand Oaks; New Delhi: Sage.
- Kurzman, C., Anderson, C., Key, C., Lee, Y. O., Moloney, M., Silver, A., & Van Ryn, M. W. (2007). Celebrity Status. *Sociological Theory*, 25(4), 347-367.
- Labov, W. (1972a). *Language in the Inner City. Studies in Black English Vernacular*. Philadelphia: University of Philadelphia Press.
- Labov, W. (1972b). Rules for Ritual Insults. In D. Sudnow (Ed.), *Studies in Social Interaction* (pp. 120-169). New York: The Free Press.
- Labov, W., & Waletzky, J. (1967). Narrative analysis. In J. Helm (Ed.), *Essays on the Verbal and Visual Arts* (pp. 12-44). Seattle: University of Washington Press.

- Lakoff, G. (1986). A Figure of Thought. *Metaphor and Symbol*, 1(3), 215-225.
- Lakoff, G. (1987). *Women, Fire, and Dangerous Things. What Categories Reveal about the Mind*. Chicago: The University of Chicago Press.
- Lakoff, G. (2008). The neural theory of metaphor. In R. W. Gibbs (Ed.), *The Cambridge Handbook of Metaphor and Thought* (pp. 17-38). Cambridge: Cambridge University Press.
- Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. Chicago: The University of Chicago Press.
- Lakoff, G., & Núñez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York: Basic Books.
- Lakoff, G., & Turner, M. (1989). *More than Cool Reason. A Field Guide to Poetic Metaphor*. Chicago; London: The University of Chicago Press.
- Langacker, R. W. (1999). *Grammar and Conceptualization*. Berlin: Mouton de Gruyter.
- Langacker, R. W. (2001). Discourse in Cognitive Grammar. *Cognitive Linguistics*, 12(2), 143-188.
- Lee, R. G., Neidle, C., MacLaughlin, D., Bahan, B., & Kegl, J. (1997). Role Shift in ASL: A Syntactic Look at Direct Speech. *American Sign Language Linguistic Research Project Report No. 4* (pp. 24-45). Boston: Boston University. Electronic copy available at <ftp://louis-xiv.bu.edu/pub/asl/rpt4/ASLLRPr4.pdf> [last accessed 29 Feb 2012].
- Lenk, U. (1998). *Marking discourse coherence: functions of discourse markers in spoken English*. Tübingen: Narr.
- Lerner, G. H. (1994). Responsive list construction: a conversational resource for accomplishing multifaceted social action. *Journal of Language and Social Psychology*, 13, 20-33.
- Lewes, G. H. (1875). *Problems of Life and Mind* (Vol. 2). London: Trübner & Co.
- Liebal, K., Behne, T., Carpenter, M., & Tomasello, M. (2009). Infants use shared experience to interpret pointing gestures. *Developmental Science*, 12(2), 264-271.
- Luckmann, T. (2006). Some Remarks on Scores in Multimodal Sequential Analysis. In H. Knoblauch & B. Schnettler & J. Raab & H.-G. Soeffner (Eds.), *Video Analysis: Methodology and Methods. Qualitative Audiovisual Data Analysis in Sociology* (pp. 29-34). Frankfurt a.M.: Peter Lang.
- Macaulay, R. (1987). Polyphonic monologue: Quoted direct speech in oral narrative. *IPrA Papers in Pragmatics*, 2, 1-34.
- Macaulay, R. (2001). You're like 'why not?' The quotative expressions of Glasgow adolescents. *Journal of Sociolinguistics*, 5(1), 3-21.

- Mandelblit, N. (1997). *Grammatical Blending: Creative and Schematic Aspects in Sentence Processing and Translation*. Unpublished doctoral dissertation, University of California, San Diego.
Electronic copy available at www.cogsci.ucsd.edu/~faucon/NILI/content.pdf [last accessed 12 Feb 2012].
- Mandelblit, N., & Fauconnier, G. (2000). How I got myself arrested: Underspecificity in Grammatical Blends as a Source for Constructional Ambiguity. In A. Foolen & F. van der Leek (Eds.), *Constructions in Cognitive Linguistics. Selected papers from the Fifth International Cognitive Linguistics Conference, Amsterdam, 1997* (pp. 167-190). Amsterdam: John Benjamins.
- Marin-Arrese, J. I. (2008). Cognition and culture in political cartoons. *Intercultural Pragmatics*, 5(1), 1-18.
- Mathis, T., & Yule, G. (1994). Zero Quotatives. *Discourse Processes*, 18(1), 63-76.
- McClave, E. Z. (2000). Linguistic functions of head movements in the context of speech. *Journal of Pragmatics*, 32, 855-878.
- McLuhan, M. (1964). *Understanding Media: The Extensions of Man*. New York: Mentor.
- Mead, G. H. (1934). *Mind, Self and Society*. Chicago: University of Chicago Press.
- Miller, J., & Weinert, R. (1995). The function of LIKE in spoken language. *Journal of Pragmatics*, 23, 365-393.
- Moerman, M. (1988). *Talking Culture. Ethnography and Conversation Analysis*. Philadelphia: University of Pennsylvania Press.
- Mondada, L. (2006). Video Recording as the Reflexive Preservation and Configuration of Phenomenal Features for Analysis. In H. Knoblauch & B. Schnettler & J. Raab & H.-G. Soeffner (Eds.), *Video Analysis: Methodology and Methods. Qualitative Audiovisual Data Analysis in Sociology* (pp. 51-67). Frankfurt a.M.: Peter Lang.
- Mondada, L. (2009). Video Recording Practices and the Reflexive Constitution of the Interactional Order: Some Systematic Uses of the Split-Screen Technique. *Human Studies*, 32, 67-99.
- Mukherjee, J. (2001). *Form and function of parasyntactic presentation structures: a corpus-based study of talk units in spoken English*. Amsterdam; Atlanta: Rodopi.
- Mukherjee, J. (2004). Corpus data in a usage-based cognitive grammar. In K. Aijmer & B. Altenberg (Eds.), *The Theory and Use of Corpora: Papers from the 23rd International Conference on English Language Research on Computerized Corpora (ICAME 23)* (pp. 85-100). Amsterdam: Rodopi.
- Niemelä, M. (2010). The reporting space in conversational storytelling: Orchestrating all semiotic channels for taking a stance. *Journal of Pragmatics*, 42, 3258-3270.

- Norrick, N. R. (2000). *Conversational Narrative. Storytelling in everyday talk*. Amsterdam; Philadelphia: John Benjamins.
- Norrick, N. R. (2010a). Laughter before the punch line during the performance of narrative jokes in conversation. *Text & Talk*, 30(1), 75-95.
- Norrick, N. R. (2010b). Listening practices in television celebrity interviews. *Journal of Pragmatics*, 42, 525-543.
- Oakley, T. (1998). Conceptual blending, narrative discourse, and rhetoric. *Cognitive Linguistics*, 9, 321-360.
- Oakley, T., & Coulson, S. (2008). Connecting the dots: Mental spaces and metaphoric language in discourse. In T. Oakley & A. Hougaard (Eds.), *Mental Spaces in Discourse and Interaction* (pp. 27-50). Amsterdam: John Benjamins.
- Oakley, T., & Hougaard, A. (Eds.). (2008). *Mental Spaces in Discourse and Interaction*. Amsterdam, Philadelphia: John Benjamins.
- Oakley, T., & Kaufer, D. (2008). Designing clinical experiences with words: Three layers of analysis. In T. Oakley & A. Hougaard (Eds.), *Mental Spaces in Discourse and Interaction* (pp. 149-178). Amsterdam: John Benjamins.
- Oring, E. (1989). Between jokes and tales: on the nature of punch lines. *Humor. International Journal of Humor Research*, 2(4), 349-364.
- Pascual, E. (2002). *Imaginary Dialogues: Conceptual Blending and Fictive Interaction in Criminal Courts*. Utrecht: LOT.
- Pascual, E. (2006). Fictive interaction within the sentence: A communicative type of fictivity in grammar. *Cognitive Linguistics*, 17(2), 245-267.
- Pascual, E. (2008). Fictive interaction blends in everyday life and courtroom settings. In T. Oakley & A. Hougaard (Eds.), *Mental Spaces in Discourse and Interaction* (pp. 79-108). Amsterdam: John Benjamins.
- Pascual, E. (2009). "I was in that room!": Conceptual integration of content and context in a writer's vs. a prosecutor's description of a murder. In V. Evans & S. Pourcel (Eds.), *New Directions in Cognitive Linguistics* (pp. 499-516). Amsterdam: John Benjamins.
- Paterniti, D., Fancher, T., Cipri, C., Timmersmans, S., Heritage, J., & Kravitz, R. (2010). Getting to "No": Strategies Primary Care Physicians Use to Deny Patient Requests. *Archives of Internal Medicine*, 170, 381-388.
- Peirce, C. S. (1868). On a New List of Categories. *Proceedings of the American Academy of Arts and Sciences*, 7, 287-298. Electronic copy available at <http://www.peirce.org/writings/p32.html> [last accessed 12 Feb 2012].
- Peter de Ruiter, J. (2007). Postcards from the mind. The relationship between speech, imagistic gesture, and thought. *Gesture*, 7(1), 21-38.

- Plato (2007). *The Republic* (D. Lee, Trans. 2nd ed.). London: Penguin Classics.
- Platow, M. J., Haslam, S. A., Both, A., Chew, I., Cuddon, M., Goharpey, N., Maurer, J., Rosini, S., Tsekouras, A., & Grace, D. M. (2005). "It's not funny if they're laughing": Self-categorization, social influence, and responses to canned laughter. *Journal of Experimental Social Psychology*, 41, 542-550.
- Plester, B. A., & Sayers, J. (2007). 'Taking the piss': Functions of banter in the IT industry. *Humor. International Journal of Humor Research*, 20(2), 157-187.
- Plum, A. (1998). *Die Karikatur im Spannungsfeld von Kunstgeschichte und Politikwissenschaft. Eine ikonologische Untersuchung zu Feindbildern in Karikaturen*. Doctoral dissertation, Technische Hochschule, Aachen. Electronic copy available at <http://darwin.bth.rwth-aachen.de/opus3/volltexte/1999/9/> [last accessed 12 Feb 2012].
- Priego-Valverde, B. (2009). Failed humor in conversation: a double voicing analysis. In N. R. Norrick & D. Chiaro (Eds.), *Humor in Interaction* (pp. 165-183). Amsterdam: John Benjamins.
- Psathas, G. (1995). *Conversation Analysis. The Study of Talk-in-Interaction*. Thousand Oaks; London; New Delhi: Sage Publications.
- Rampton, B. (1999). Styling the Other: Introduction. *Journal of Sociolinguistics*, 3(4), 421-427.
- Raskin, V. (1985). *Semantic Mechanisms of Humor*. Dordrecht: D. Reidel.
- Rhodes, G., & Brennan, S. (1987). Identification and Ratings of Caricatures: Implications for Mental Representations of Faces. *Cognitive Psychology*, 19, 473-497.
- Richards, I. A. (1936). *The Philosophy of Rhetoric*. New York; London: Oxford University Press.
- Rickheit, G., Weiss, S., Eikmeyer, H.-J. (2010). *Kognitive Linguistik. Theorien, Modelle, Methoden*. Tübingen; Basel: A. Francke Verlag UTB.
- Ricoeur, P. (1975). *La métaphore vive*. Paris: Éditions du Seuil.
- Rohrer, T. (2005). Mimesis, artistic inspiration and the blends we live by. *Journal of Pragmatics*, 37, 1686-1716.
- Romaine, S., & Lange, D. (1991). The Use of Like as a Marker of Reported Speech and Thought: A Case of Grammaticalization in Progress. *American Speech*, 66, 227-279.
- Rowe, A. D., Bullock, P. R., Polkey, C. E., & Morris, R. G. (2001). 'Theory of mind' impairments and their relationship to executive functioning following frontal lobe excisions. *Brain*, 124(3), 600-616.

- Rühlemann, C. (2008). Conversational grammar - bad grammar? A situation-based description of quotative I goes in the BNC. *ICAME Journal*, 32, 157-177.
- Sacks, H. (1974). An analysis of the course of a joke's telling in conversation. In R. Bauman & J. Sherzer (Eds.), *Explorations in the Ethnography of Speaking* (pp. 227-353). Cambridge: Cambridge University Press.
- Sacks, H. (1984). Notes on methodology. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action: Studies in Conversation Analysis* (pp. 21-27). Cambridge: Cambridge University Press.
- Sacks, H. (1992). *Lectures on Conversation* (Vol. 1&2). Oxford: Blackwell.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language* 50(1/4), 696-735.
- Saussure, F. de (1916). *Cours de linguistique générale*. Lausanne; Paris: Payot.
- Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Hillsdale: Erlbaum.
- Schegloff, E. A. (1987). Some sources of misunderstanding in talk-in-interaction. *Linguistics*, 25(1), 201-218.
- Schegloff, E. A. (1992). Introduction. In H. Sacks, *Lectures on Conversation* (Vol. 1, pp. ix-lxiv). Malden; Oxford: Blackwell.
- Schegloff, E. A. (1996a). Issues of Relevance for Discourse Analysis: Contingency in Action, Interaction and Co-Participant Context. In E. H. Hovy & D. Scott (Eds.), *Interdisciplinary Perspectives on Discourse* (pp. 3-35). Heidelberg: Springer
- Schegloff, E. A. (1996b). Turn organization: one intersection of grammar and interaction. In E. Ochs & E. A. Schegloff & S. A. Thompson (Eds.), *Interaction and grammar* (pp. 52-133). Cambridge: Cambridge University Press.
- Schegloff, E. A. (2003). *Discussing Conversation Analysis*. Amsterdam: John Benjamins.
- Schegloff, E. A. (2009). A Practice for (Re-)Exiting a Sequence: And/But/So + Uh(m) + Silence. In B. Fraser & K. Turner (Eds.), *Language in Life, and a Life in Language: Jacob Mey - A Festschrift* (pp. 365-374). Bingley: Emerald Group.
- Schegloff, E. A. (2011). Word repeats as unit ends. *Discourse Studies*, 13(3), 367-380.
- Schickel, R. (1985). *Intimate Strangers: The Culture of Celebrity*. Garden City, NY: Doubleday.
- Schmid, H.-J. (2011). Conceptual blending, relevance and novel N+N compounds. In S. Handl & H.-J. Schmid (Eds.), *Windows to the Mind. Metaphor, Metonymy and Conceptual Blending* (pp. 219-246). Berlin: De Gruyter Mouton.
- Schmid, H.-J. (Ed.). (2012). *Cognitive Pragmatics*. Berlin: De Gruyter Mouton.

- Schütz, A., & Luckmann, T. (1979). *Strukturen der Lebenswelt* (Vol. 1). Frankfurt: Suhrkamp.
- Schwarz, M. (2008). *Einführung in die kognitive Linguistik*. 3rd ed. Tübingen; Basel: A. Francke UTB.
- Searle, J. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.
- Selting, M. (2000). The construction of units in conversational talk. *Language in Society*, 29, 477-517.
- Selting, M. (2007). Lists as embedded structures and the prosody of list construction as an interactional resource. *Journal of Pragmatics*, 39, 483-526.
- Selting, M. (2010). Prosody in interaction: State of the art. In D. Barth-Weingarten & E. Reber & M. Selting (Eds.), *Prosody in Interaction* (pp. 3-40). Amsterdam; Philadelphia: John Benjamins.
- Selting, M., Auer, P., Barden, B., Bergmann, J., Couper-Kuhlen, E., Günthner, S., Meier, C., Quasthoff, U., Schlobinski, P., & Uhmman, S. (1998). Gesprächsanalytische Transkriptionssysteme. *Linguistische Berichte*, 173, 91-122.
- Selting, M., Auer, P., Barth-Weingarten, D., Bergmann, J., Bergmann, P., Birkner, K., Couper-Kuhlen, E., Deppermann, A., Gilles, P., Günthner, S., Hartung, M., Kern, F., Mertzluff, C., Meyer, C., Morek, M., Oberzaucher, F., Peters, J., Quasthoff, U., Schütte, W., Sturkenbrock, A., & Uhmman, S. (2009). Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion*, 10, 353-402.
- Sidnell, J. (2006). Coordinating Gesture, Talk, and Gaze in Reenactments. *Research on Language and Social Interaction*, 39(4), 377-409.
- Singler, J. (2001). Why you can't do a VARBRUL study of quotatives and what such a study can show us. *Selected Papers from NWAV 29. Penn Working Papers in Linguistics*, 7(3), 257-278.
- Sinha, C. (1999). Grounding, mapping, and acts of meaning. In T. Janssen & G. Redeker (Eds.), *Cognitive Linguistics: Foundations, Scope and Methodology* (pp. 223-255). Berlin: Mouton de Gruyter.
- Solomon, O. (2004). Narrative introductions: discourse competence of children with autistic spectrum disorders. *Discourse Studies*, 6(2), 253-276.
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and Cognition*. Oxford: Blackwell.
- Stern, D. G. (2004). *Wittgenstein's Philosophical Investigations: an introduction*. Cambridge: Cambridge University Press.
- Stivers, T., & Sidnell, J. (2005). Introduction: Multimodal Interaction. *Semiotica*, 156(1), 1-20.

- Streeck, J. (2002). Grammars, Words, and Embodied Meanings: On the Uses and Evolution of So and Like. *Journal of Communication*, 52(3), 581-596.
- Suls, J. M. (1977). Cognitive and disparagement theories of humour: A theoretical and empirical synthesis. In A. J. Chapman & H. C. Foot (Eds.), *It's a funny thing, humour* (pp. 41-46). Oxford: Pergamon Press.
- Sweetser, E. (1999). Compositionality and blending: semantic composition in a cognitively realistic framework. In G. Redeker & T. Janssen (Eds.), *Cognitive Linguistics: Foundations, Scope and Methodology* (pp. 129-162). Berlin: Mouton de Gruyter.
- Sweetser, E. (2000). Blended spaces and performativity. *Cognitive Linguistics*, 11(3/4), 305-333.
- Sweetser, E. (2007). Looking at space to study mental spaces: Co-speech gesture as a crucial data source in cognitive linguistics. In M. Gonzalez-Marquez & I. Mittelberg & S. Coulson & M. J. Spivey (Eds.), *Methods in Cognitive Linguistics* (pp. 201-224). Amsterdam: John Benjamins.
- Sweetser, E., & Sizemore, M. (2008). Personal and interpersonal gesture spaces: Functional contrasts in language and gesture. In A. Tyler & Y. Kim & M. Takada (Eds.), *Language in the Context of Use. Discourse and Cognitive Approaches to Language* (pp. 25-52). Berlin; New York: Mouton de Gruyter.
- Szczepek Reed, B. (2006). *Prosodic Orientation in English Conversation*. Basingstoke; New York: Palgrave Macmillan.
- Tagliamonte, S., & Hudson, R. (1999). Be like et al. beyond America: The quotative system in British and Canadian youth. *Journal of Sociolinguistics*, 3(2), 147-172.
- Talmy, L. (2000). *Toward a Cognitive Semantics* (Vol. I&II). Cambridge: The MIT Press.
- Talmy, L. (2007). Foreword. In M. Gonzalez-Marquez & I. Mittelberg & S. Coulson & M. J. Spivey (Eds.), *Methods in Cognitive Linguistics* (pp. xi-xxi). Amsterdam, Philadelphia: John Benjamins.
- Tannen, D. (1986). Introducing Constructed Dialogue in Greek and American Conversational and Literary Narratives. In F. Coulmas (Ed.), *Direct and Indirect Speech* (pp. 311-322). Berlin: Mouton.
- Tannen, D. (1989). *Talking voices. Repetition, dialogue, and imagery in conversational discourse*. Cambridge: Cambridge University Press.
- Ten Have, P. (1999). *Doing Conversation Analysis. A Practical Guide*. London; Thousand Oaks; New Delhi: Sage.
- Tomasello, M. (2008). *Origins of Human Communication*. Cambridge: The MIT Press.
- Tragesser, S. L., & Lippman, L. G. (2005). Teasing: For Superiority or Solidarity? *The Journal of General Psychology*, 132(3), 255-266.

- Tsakona, V. (2009). Language and image interaction in cartoons: Towards a multimodal theory of humor. *Journal of Pragmatics*, 41, 1171-1188.
- Turner, M. (2000). Backstage Cognition in Reason and Choice. In A. Lupia & M. D. McCubbins & S. L. Popkin (Eds.), *Elements of Reason: Cognition, Choice, and the Boundaries of Rationality* (pp. 264-286). Cambridge: Cambridge University Press. Electronic copy available at <http://ssrn.com/abstract=1249963> [last accessed 12 Feb 2012].
- Turner, M. (2010). Blending Box Experiments, Build 1.0. Working Paper. Electronic copy available at <http://ssrn.com/abstract=1541062>.
- Vandelanotte, L., & Davidse, K. (2009). The emergence and structure of be like and related quotatives: A constructional account. *Cognitive Linguistics*, 20(4), 777-807.
- Veale, T., & O'Donoghue, D. (2000). Computation and Blending. *Cognitive Linguistics*, 11(3/4), 253-281.
- Veale, T., Feyaerts, K., & Brône, G. (2006). The cognitive mechanisms of adversarial humor. *Humor. International Journal of Humor Research*, 19(3), 305-339.
- Wells, J. C. (2006). *English intonation: An introduction*. Cambridge: Cambridge University Press.
- Wetherell, M. (1998). Positioning and interpretative repertoires: conversation analysis and post-structuralism in dialogue. *Discourse & Society*, 9(3), 387-412.
- Wignall, J. (2008, 7 November). Never Mind the Buzzcocks has lost its edge. *The Guardian*. Electronic copy available at <http://www.guardian.co.uk/culture/tvandradioblog/2008/nov/07/never-mind-the-buzzcocks> [last accessed 12 Feb 2012].
- Willems, H. (2004). Erving Goffman. In U. Flick & E. v. Kardorff & I. Steinke (Eds.), *A Companion to Qualitative Research* (pp. 24-28). London; Thousand Oaks; New Delhi: Sage.
- Williams, R. F. (2004). *Making meaning from a clock: Material artifacts and conceptual blending in time-telling instruction*. Unpublished Doctoral Dissertation, University of California at San Diego, San Diego. Electronic copy available at <http://www.cogsci.ucsd.edu/research/publications-dissertations/publications-dissertations-rfwilliams.php?travel=> [last accessed 12 Feb 2012].
- Williams, R. F. (2008). Guided conceptualization: Mental spaces in instructional discourse. In T. Oakley & A. Hougaard (Eds.), *Mental Spaces in Discourse and Interaction* (pp. 209-234). Amsterdam: John Benjamins.
- Wilson, D., & Sperber, D. (2006). Relevance Theory. In L. R. Horn & G. Ward (Eds.), *Handbook of Pragmatics* (pp. 607-632). Malden; Oxford: Blackwell.

- Wirth, U. (2002). Der Performanzbegriff im Spannungsfeld von Illokution, Iteration und Indexikalität. In U. Wirth (Ed.), *Performanz. Zwischen Sprachphilosophie und Kulturwissenschaften*. Frankfurt: suhrkamp taschenbuch wissenschaft.
- Wittgenstein, L. (1953). *Philosophical Investigations*. New York: Macmillan.
- Young, R. F. (2008). What Is Discursive Practice? *Language Learning*, 58(2), 1-8.
- Yule, G. (1998). *Explaining English Grammar*. Oxford: Oxford University Press.
- Zajdman, A. (1995). Humorous face-threatening acts: Humor as strategy. *Journal of Pragmatics*, 23, 325-339.
- Zillmann, D. (1983). Disparagement Humor. In P. E. McGhee & J. H. Goldstein (Eds.), *Handbook of Humor Research* (Vol. 1. Basic Issues, pp. 85-107). New York: Springer.
- Zlatev, J. (2010). Phenomenology and Cognitive Linguistics. In S. Gallagher & D. Schmicking (Eds.), *Handbook on Phenomenology and Cognitive Science* (pp. 415-446). Dordrecht: Springer.

APPENDIX

SPRACHE, KOGNITION, INTERAKTION. KONZEPTUELLE INTEGRATION ALS DISKURSIVE PRAXIS

Kapitel 1: Konzeptuelle Integration in einer interaktionalen Linguistik.

Einführung

Die kognitive Linguistik entstand in den 1970er Jahren in Reaktion auf die damals vorherrschenden formalen Zugänge zur Spracherforschung. Im Gegensatz zur generativen Linguistik wird dabei angenommen, dass Sprache kein „autonomes Modul“ (Rickheit et al., 2010:14) darstellt, sondern auf allgemeinen kognitiven Möglichkeiten des Menschen, wie Wahrnehmung und Kategorisierung, beruht. Darüber hinaus wird die „Körpergebundenheit“ (Schwarz, 2008) des menschlichen Denkens und Sprechens betont, die sich beispielsweise in prä-linguistischen Bild-Schemata (Rickheit et al, 2010) zeigt und so Sprache und Denken strukturiert. Innerhalb der kognitiven Linguistik beschäftigt sich die kognitive Semantik mit Bedeutungskonstitution. Die kognitive Linguistik geht hier von einem situierten, dynamischen Prozess aus (die Konzeptualisierung), der auf bildlicher mentaler Simulation, Weltwissen und Körpergebundenheit basiert.

Obwohl die kognitive Linguistik durch die Einbindung des Körpers in die Analyse von Prozessen der Bedeutungskonstitution zu einer produktiven Erweiterung des Forschungsfeldes geführt hat, wurde ihr in den letzten Jahren zunehmend vorgeworfen, den sozialen Aspekt der Bedeutungskonstitution zu vernachlässigen. Croft (2009) fordert so beispielsweise eine ‚soziale Wende‘, die zur Überwindung der tradierten Teilung in soziale und psychologische Aspekte von menschlichem Verhalten führen soll. Trotz einiger erster Versuche ist die Einbindung der vorrangig sozialen Funktion von Sprache innerhalb der anglophonen kognitiven Linguistik noch nicht vollzogen worden.

Die vorliegende Arbeit setzt hier an und bietet einen Beitrag zur Entwicklung einer dezidiert sozial-interaktional fundierten kognitiven Linguistik und Semantik. Sie konzentriert sich dabei auf die von Fauconnier und Turner (2002) formulierte Theorie zu mentalen Räumen und konzeptueller Integration (*Mental Spaces and*

Conceptual Integration Theory). Dieser (hier unter MSCI gefasste) Theoriekomplex versucht, „das Herz menschlicher Vorstellungskraft“ (ibid: 89, Übersetzung VS) zu erschließen, und geht davon aus, dass aus der Integration zweier separater semantischer Szenarien (so genannter Input-Domänen) über einen generischen Raum, der Strukturen und Elemente umfasst, die beide Input-Domänen gemeinsam haben, eine neue Bedeutungsstruktur emergiert.

Trotz der Attraktivität und intuitiven Plausibilität des Theoriekomplexes ergeben sich einige Probleme. Die oben bereits skizzierte Vernachlässigung sozialer, interaktionaler Faktoren in der Bedeutungskonstitution stellt einen wichtigen Punkt dar. Außerdem werden in der Theorie Phylogenese und Ontogenese von Bedeutungen nicht immer klar getrennt, sowie die Etymologie und Historizität von figurativen Ausdrücken nicht berücksichtigt. Die vorliegende Arbeit fragt daher nach zwei miteinander verwobenen Aspekten:

- Forschungsfrage 1) betrifft die interaktionale Aushandlung von konzeptioneller Integration. Wie werden mentale Räume in der konkreten Interaktion kenntlich gemacht? Wie werden sie in Beziehung zueinander gesetzt (*mapping*)?
- Um Forschungsfrage 1) weiter zu konkretisieren, bezieht sich Forschungsfrage 2) auf einen theorieimmanenten Aspekt und untersucht die von Fauconnier und Turner (2002) formulierte Typologie von konzeptueller Integration. So soll gefragt werden, inwieweit die hochkomplexen doppelt gerahmten Integrationsnetzwerke (*double-scope blending*), die in der Forschungsliteratur als Bedingung menschlicher Einzigartigkeit beschrieben werden, eine eigene diskursive Praxis¹ darstellen und sich von weniger komplexer konzeptioneller Integration interaktional unterscheiden lassen.

Zur Beantwortung dieser beiden Forschungsfragen untersuche ich Fallbeispiele von polyphonem Stimmwechsel in der Scherzkommunikation mithilfe der Methoden der Konversationsanalyse. Die Stimmwechsel beruhen in prototypischer Weise auf konzeptioneller Integration unterschiedlichen Komplexitätsgrades und erlauben so

¹ Meine Verwendung des Ausdrucks 'diskursive Praxis' unterscheidet sich hier explizit von der kulturwissenschaftlichen Tradition im Sinne von Foucault oder Butler. Ich beziehe mich auf das interaktionale, diskursive ‚Tun‘, das gemeinsam zwischen Sprecher_innen ausgehandelt wird und dabei auf Abstraktionen bisheriger Erfahrungen basiert. In meiner Arbeit ist von Interesse, ob die verschiedenen Typen von konzeptioneller Integration unter Rückgriff auf unterschiedliche Methoden und Praxen verhandelt werden.

die Beschreibung von interaktionalen Aushandlungsprozessen der von Fauconnier und Turner (2002) beschriebenen kognitiv-semantischen Prozesse. So wird Forschungsfrage 2) beantwortet. Im Anschluss an die Analyse wird ein Modell vorgeschlagen, das auf den Analyseergebnissen basiert, der sozialen Dimension von Bedeutungsaushandlung Rechnung trägt und somit auf Forschungsfrage 1) eingeht.

Kapitel 2: Der Theoriekomplex. Mentale Räume und konzeptuelle Integration

Dieser Abschnitt bietet eine Einführung in das von Fauconnier (1994, 1997) entwickelte Konzept zu mentalen Räumen und die darauf aufbauende Theorie der konzeptuellen Integration (Fauconnier & Turner, 2002). In der kognitiven Semantik wird davon ausgegangen, dass Sprache nicht inhärent bedeutungstragend ist, sondern Bedeutungskonstitutionen anregt. Diese Bedeutungskonstitutionen selbst sind konzeptueller Natur, prozedural und dynamisch. Aus dieser Perspektive stellen Sätze, die Bedeutung stets unterspezifizieren, eine Art Anleitung zur Bedeutungskonstitution dar, die auf der Konstruktion von komplexen temporalen konzeptuellen Domänen, sogenannter mentaler Räume, beruht. Sprache selbst regt – unter Bezug auf enzyklopädisches Welt- und Kontextwissen – zum Aufbau solcher Domänen an und führt auf diese Weise zu situativ bedingter Bedeutungskonstitution. Im Gegensatz zu analytischen und formalen Sprachphilosophien ist Bedeutung somit weder in der Sprache selbst noch in der externen Welt verortet, sondern entsteht dynamisch-konzeptuell und ist stets situativ eingebunden.

Die durch Sprache konstruierten konzeptuellen Domänen werden bei Fauconnier (1994, 1997) unter dem Begriff ‚mentale Räume‘ gefasst. Die im Zuge der Bedeutungskonstitution konstruierten mentalen Räume werden miteinander auf verschiedene Weise verbunden und in Beziehung gesetzt (*mappings*). Dieser Aspekt wird in der Weiterentwicklung des Modells durch Fauconnier und Turner (2002) genauer diskutiert. Dabei wird darauf verwiesen, dass sich aus der konzeptionellen Verbindung und Interaktion zweier Domänen emergente Strukturen entwickeln können, die *blends* genannt werden. Das folgende Beispiel (adaptiert von Fauconnier,

1997:155ff) illustriert den Prozess. Im Jahr 1993 folgt der Katamaran ‚Great America II‘ dem Kurs, den der Klipper ‚Northern Light‘ bereits im Jahre 1853 einschlug. Segelzeitschriften verglichen die Törns der beiden Yachten in Äußerungen wie „Die Great America II liegt zur Zeit 4,5 Seemeilen vor der Northern Light.“ Obwohl 150 Jahre zwischen den beiden Törns liegen, entsteht der Eindruck einer Regatta. In der ‚Verschmelzung‘, dem *blend*, treten die Yachten gegeneinander an. Aus den beiden Input-Domänen (dem Szenario des Katamaran und dem des Klippers) werden selektiv Strukturelemente in den *blend* projiziert: das Jahr, die Wetterkonditionen und das eine Schiff werden aus der Input-Domäne des Katamarans genommen. Aus der Domäne des Klippers wird das zweite Schiff entnommen. Weitere Wissensstrukturen werden hingegen vernachlässigt. Der Vergleich wird möglich durch einen generischen Raum (*generic space*), der die Struktur, die beide Input-Domänen gemeinsam haben, subsumiert. Abbildung 1 zeigt eine schematische Visualisierung dieses Prozesses.

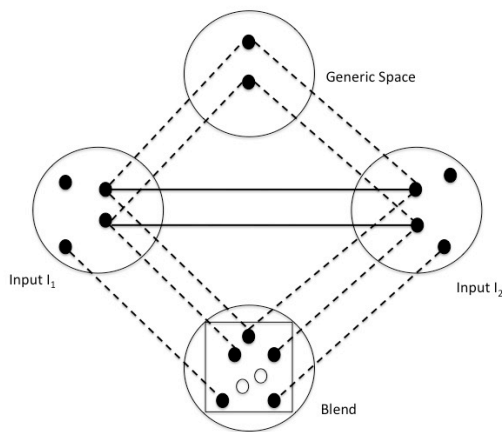


Abbildung 1: Konzeptionelle Integration: Schematisches Diagramm

Durch die Verschmelzung ist ein direkter Vergleich der beiden Schiffe möglich. Die dynamische Simulation des entstandenen Regattaszenarios bringt emergierende Wissensstrukturen, wie beispielsweise das jeweilige Tempo der Yachten, hervor, die wiederum in die Input-Domänen zurückgespiegelt werden können und die jeweiligen Wissensstrukturen weiter anreichern.

Die Autoren unterscheiden zwischen vier Typen von konzeptueller Integration, die unterschiedliche Komplexitätsstufen repräsentieren. Die sogenannten Simplex-Netzwerke beschreiben *blends*, in denen ein Input eine bekannte Rahmungsstruktur (wie zum Beispiel ‚Familie‘) und ein anderer die Werte (‚Peter‘, ‚Maria‘) bereitstellt. So werden Werte mit rahmender Struktur verbunden (‚Peter ist Marias Vater‘). An zweiter Stelle des Integrations-Kontinuums stehen sogenannte

Spiegel-Netzwerke. In diesem Typ haben alle Input-Räume die gleiche Organisationsstruktur, die „relevante Aktivitäten, Ereignisse und Teilnehmer_innen“ (Fauconnier und Turner, 2002:123, Übersetzung VS) umfasst. Das oben beschriebene Regattabeispiel stellt ein Spiegel-Netzwerk dar. Auf der nächsthöheren Komplexitätsstufe verorten die Autoren sog. ‚einfach gerahmte Netzwerke‘. Im Gegensatz zu Simplex- und Spiegel-Netzwerken integrieren einfach gerahmte Netzwerke semantisch nicht kompatible Input-Domänen. Im *blend* selbst ist lediglich die Organisationsstruktur eines Inputs zu finden. Die Organisationsstruktur des zweiten Inputs wird fallengelassen. Das Beispiel der boxenden Firmen illustriert diesen Typus: Wenn ein Unternehmen einem anderen einen herben Schlag versetzt oder es gar k.o. schlägt, so tragen zwei widersprüchliche Wissensdomänen zum Integrationsnetzwerk bei: das Szenario der Geschäftswelt und das der Boxwelt. Doch nur die Ereignisstruktur der Boxwelt strukturiert den *blend*, da aus dem Szenario der Geschäftswelt lediglich die Figuren stammen. Viele konzeptionelle Metaphern folgen diesem Schema einer asymmetrischen, einfach gerahmten konzeptuellen Integration.

Die komplexeste Form der konzeptuellen Integration stellen schließlich ‚doppelt gerahmte Integrationsnetzwerke‘ dar. Wie bei einfach gerahmten Integrationsnetzwerken finden sich semantisch nicht kompatible Input-Räume. Bei der Integration dieser Domäne finden jedoch Rahmungsaspekte beider Inputs Eingang in den *blend* und resultieren in einer emergenten Struktur, die in keiner der beiden Inputs vorhanden ist. Als Beispiel kann die Transsubstantiation genannt werden, in der Brot und Wein zu Leib und Blut Jesu werden. Die Fähigkeit zur Konstruktion doppelt gerahmter konzeptueller Integrationsnetzwerke ist, so die Autoren, die Wurzel aller menschlicher Kulturleistungen.

Kritische Würdigung

Trotz der Attraktivität der Theorie finden sich einige Problemlagen philosophischer wie methodologischer Natur, die im Folgenden zusammengefasst werden.

1. Post-hoc Analysen: In der kanonischen MSCI-Literatur wird vom ‚Produkt‘ ausgegangen und versucht, die Schritte, die zur einer gewissen Konzeptualisierung geführt haben, nachzuzeichnen. Dies kann nicht nur zu psychologisch zweifelhaften Ergebnissen führen, sondern auch zu einer

- Vernachlässigung von interaktionalen Aushandlungsprozessen und der Kontextgebundenheit von Bedeutung.
2. Der generische Raum: Dieser tertium comparationis-Raum fügt dem konzeptuellen Integrationsnetzwerk keine weitere relevante Wissensstruktur zu und wird lediglich in dekontextualisierten Beispielen zur Erklärung der Motivation von bestimmten Ausdrücken benötigt.
 3. Generalisierungen, Terminologieprobleme und Falsifizierbarkeit: Aufgrund der sehr generellen und breiten Annahmen fällt es schwer, das Theoriegebäude empirisch zu testen. Außerdem werden die jeweiligen Spezifiken der untersuchten Phänomene vernachlässigt, was die Beschreibung dichter Bedeutung erschwert.
 4. Demarkation mentaler Räume: Die einzelnen Räume sind nur schwer voneinander abzugrenzen und eine einheitliche Definition des Terminus ist noch nicht gefunden worden.
 5. Wessen mentale Räume werden beschrieben? Wessen *blends*? In der MSCILiteratur scheint oft von einem generischen Rezipienten ausgegangen zu werden, der zu einer Spiegelung der kognitiven Prozesse, die den Sender zu einer bestimmten Äußerung bewogen haben, fähig ist.
 6. Loci und Ebenen konzeptueller Integration: Eine Unterscheidung in ‚tote‘ und ‚lebendige‘ Metaphern (vgl. Ricoeur, 1975) wird nicht getroffen und, so scheint es, gleichartige kognitive Prozesse für die Bedeutungskonstitution in beiden Phänomenen angenommen.
 7. Homunculus-Problematik: MSCI unterscheidet zwischen vordergründiger und hintergründiger Kognition und hofft, letztere zu beschreiben. Doch wie sich die vordergründige von der hintergründigen Kognition unterscheidet und wie Bedeutung „in das Bewusstsein geliefert“ wird (ibid:57, Übersetzung VS), wird nicht weiter thematisiert.
 8. Konzeptuelle Integration, Kontext und Interaktion: Die situative Eingebundenheit und interaktionale ‚Erdung‘ von Konzeptualisierungsprozessen wird in der MSCILiteratur nicht näher behandelt. Die Prozesse, die zu kontextualisierter (d.h. nicht ausschließlich sedimentierter) Bedeutung führen, werden somit nicht thematisiert.
 9. Typologie: Von den oben beschriebenen vier Typen von *blends* ist vor allem der erste, die Simplex-Netzwerke auffallend. Was ist ein mentaler Raum ohne

Rahmenstrukturierung? Doch auch doppelt gerahmte Integrationsnetzwerke erscheinen bei genauerer Betrachtung problematisch. Besonders die Abgrenzung von einfach gerahmten Integrationsnetzwerken zeigt sich in der Praxis schwierig, da die Trennlinie willkürlich gezogen werden kann.

So wird deutlich, dass MSCI zwar eine durchaus einleuchtende Theorie menschlicher Kognition darstellt, doch auch eine Reihe von Schwächen methodischer sowie theoretischer Natur aufweist. Um diesen zu begegnen und somit zu einer produktiven Weiterentwicklung des Ansatzes beizutragen, werden in Kapitel drei erste Hypothesen zu einer methodologischen wie konzeptuellen Neupositionierung von MSCI vorgeschlagen, die eine empirische Analyse kontextualisierter *blends* ermöglichen soll.

Kapitel 3: Vorüberlegungen zur theoretischen Positionierung und empirischen Validierung einer sozial-interaktionalen kognitiven Semantik

Die Methodik der kognitiven Linguistik ist traditionell introspektiv, hat sich in den letzten Jahren aber zunehmend empirischen und experimentellen Zugängen geöffnet. Die vorliegende Arbeit folgt dabei weitgehend dem von Hougaard (2004, 2005, 2008) vorgeschlagenem diskursiven Ansatz, der auf Methoden der Konversationsanalyse zurückgreift. Dieser Ansatz wird gewählt, da mit seiner Hilfe voll kontextualisierte, nicht simulierte Daten berücksichtigt werden können und eine ad-hoc-Beschreibung von interaktionalen Bedeutungsaushandlungsprozessen möglich wird. Dieser interaktionale Zugang wurde nicht nur von Hougaard verfolgt, sondern führte bereits zu ersten Ansätzen einer Erweiterung des MSCI-Modells um semio-pragmatische Aspekte durch Brandt und Brandt (2005), die jedoch stark auf semiotischen Theorien basieren und die interaktionale Aushandlung von Bedeutung nicht gebührend berücksichtigen. Zudem beschäftigen sich ihre Analysen nicht mit der Validität der Binnenstruktur des Theoriekomplexes. Dieser Aspekt, die MSCI-Typologie, wird jedoch von Bache (2005) und von Brandt (2010) thematisiert, die beide eine Neustrukturierung vorschlagen. Während Baches (2005) Modell die ursprüngliche

Typologie generell rafft, aber besonders die weniger komplexen Integrationstypen stärker differenziert, bietet Brandt (2010) eine ausdifferenzierte Typologie der komplexen Integrationsarten, die methodisch jedoch eher auf Introspektion als auf empirischer Analyse basiert. Die empirische Validität der jeweiligen Typologien wird somit weder bei Brandt noch bei Bache thematisiert. Aus diesem Grund schlägt die vorliegende Arbeit vor, die Binnenstruktur von *blends* mithilfe der Methoden der Konversationsanalyse genauer zu untersuchen.

Die Konversationsanalyse untersucht Alltagsgespräche im Hinblick auf diejenigen Regeln und Verfahren, mit denen Bedeutung interaktional ausgehandelt wird. Dabei geht sie davon aus, dass Bedeutung in der sozialen Aktion entsteht. Äußerungen sind somit Aktionen, die stets in spezifischen interaktionalen und sequentiellen Kontexten situiert sind und methodisch produziert werden. Die Konversationsanalyse fragt nun nach den Mechanismen, mit deren Hilfe die Teilnehmer_innen gemeinsam Bedeutung konstituieren. Dies wird durch die mikrosoziologische Betrachtung von aufgenommenen und transkribierten Alltagsgesprächen erreicht.

Einen dezidiert pragmatischen methodischen Zugang mit einer weitgehend mentalistischen Theorie in Verbindung zu setzen erfordert jedoch Vorüberlegungen zur theoretischen Konzeption der zentralen Konzepte. Diese werden im Folgenden zusammengefasst.

1. Kognition ist ‚vordergründig‘ und nicht ‚hintergründig‘. In der interaktionalen Bedeutungsaushandlung werden kognitive Prozesse koordiniert und somit offen signalisiert, um sich der Intersubjektivität anzunähern. Beschrieben wird in der Analyse also Kognition, wie sie sich interaktional vordergründig zeigt, und nicht, wie sie hintergründig ablaufen könnte.
2. Geteiltes Wissen und mentale Räume: Mentale Räume basieren auf wechselseitigen Annahmen zu geteiltem Wissen und dem jeweiligen Kommunikationspartner. Sie werden daher interaktional ausgehandelt und koordiniert, sind interaktional relevant und phänomenologisch evident.
3. Diskursbasisraum als Kommunikationsvoraussetzung: Zur Repräsentation jener Elemente des geteilten Wissens, die für die lokale Bedeutungsaushandlung relevant sind, wird auf das von Langacker (2001) entworfene Model des Diskursbasisraums (*Current Discourse Space*)

zurückgegriffen. Dieser mentale Raum umfasst Wissen, das „als von Sprecher_in und Hörer_in geteilt angenommen wird und zu einem bestimmten Zeitpunkt im Diskursverlauf als Kommunikationsbasis dient“ (ibid:144, Übersetzung VS). Entgegen dem Modell von Brandt und Brandt (2005) erlaubt Langackers Diskursbasisraum die explizite Würdigung von geteiltem Wissen und sozialer Aktion in der Bedeutungskonstitution und wird ersterem daher vorgezogen.

4. Der generische Raum wird in der Analyse nur dann berücksichtigt, wenn die Kommunikationspartner in ihrer Interaktion explizit darauf Bezug nehmen.

Bedeutung wird in der sozial-interaktionalen kognitiven Semantik, wie sie die vorliegende Arbeit vorschlägt, also nicht als individuelles Phänomen gefasst, sondern als sozial und lokal konstruiert. Zur Analyse der Prozesse, die beim Aushandeln von konzeptuellen Integrationsnetzwerken unterschiedlicher Komplexität zum Tragen kommen, konzentriert sich die Analyse auf ein interaktionales Phänomen, bei dem zwei klar umrissene und interaktional verhandelte semantische Szenarien als Input-Domänen fungieren: Stimmwechsel im polyphonen Humor. Dieses Phänomen wird im Folgenden kurz beschrieben. In einem weiteren Schritt werden Datenlage und Sammlung skizziert.

Forschungsgegenstand: Polyphone Stimmwechsel im interaktionalen Humor

Basierend auf dem von Bakhtin (1963) skizzierten Konzept der polyphonen Erzählung werden Stimmwechsel im interaktionalen Humor (wie beispielsweise Nachäffen oder Parodie) als kondensierte und fragmentierte polyphone Narrative gefasst, die welterzeugende Funktion haben. Die Stimmwechsel sind körperlicher Natur, bühnenhaft (cf. Fischer-Lichte, 2002) und ähneln Karikaturen, da sie typifizieren (Schütz & Luckmann, 1979), Normabweichungen betonen und der Lächerlichkeit preisgeben. Die Stimmwechsel stellen jedoch auch eine Art soziales Spiel dar (Goffman (1974)), das basierend auf unterschiedlichen Humormechanismen, wie der Freude an Mimesis und der überraschenden Einblicke, die das Zusammenführen inkongruenter semantischer Szenarien mit sich bringen kann, zu einem Zelebrieren von Gruppenzugehörigkeit und der damit einhergehenden Differenzkonstitution führt.

Solche polyphonen Stimmwechsel stellen prototypische Integrationsnetzwerke dar, da eine *persona* durch den Körper einer anderen präsentiert wird. Durch Dekomprimierung wird aus der integrierten Performanz des Stimmwechsel eine polyphone Erzählung, die emergente Rückschlüsse auf die Input-Domänen ermöglicht. Abbildung 2 visualisiert diesen Gedanken.

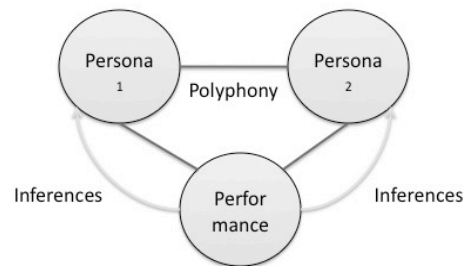


Abbildung 2: Polyphone Integrationsnetzwerke: Schematische Darstellung

Polyphone Stimmwechsel basieren auf konzeptuellen Integrationsnetzwerken unterschiedlicher Komplexität. Es finden sich Spiegel-, einfach gerahmte und doppelt gerahmte Integrationsnetzwerke, die anhand der Faktoren ‚Diskursszenario‘ und ‚semiotische Codes‘ unterschieden werden können. Während Stimmwechsel, in denen Menschen andere Menschen desselben Diskursszenarios darstellen (wie zum Beispiel im Falle von Nachäffen), auf Spiegelnetzwerken basieren, greifen solche Stimmwechsel, in denen Teilnehmer_innen Menschen in temporal, epistemisch und räumlich abgesetzten Diskursszenarien verkörpern, auf einfach gerahmte Integrationsnetzwerke zurück. Diejenigen Stimmwechsel, in denen Teilnehmer_innen nicht-menschliche Entitäten wie Maschinen, Tiere oder Pflanzen verkörpern, stellen hingegen doppelt gerahmte Integrationsnetzwerke dar. Die folgende Tabelle fasst die Kriterien zur Typologie der Stimmwechsel zusammen.

Integrationskomplexität	Diskursszenario	Semiotischer Code
Spiegel-Netzwerk (z.B. Nachäffen)	Diskursszenario ist räumlich identisch und temporal nah in beiden Input-Domänen → geringe Inkompatibilität	‚Menschliche‘ semiotische Codes → geringe Inkompatibilität
Einfach gerahmte Netzwerke (z.B. Verkörperung einer historischen Figur)	Deutlich voneinander abgesetzte Diskursszenarien. Diskursszenario der dargestellten Figur stellt Rahmenstruktur der Darstellung.	‚Menschliche‘ semiotische Codes → geringe Inkompatibilität
Doppelt gerahmte Netzwerke (z.B. Verkörperung eines Tieres)	Deutlich voneinander abgesetzte Diskursszenarien. Die Domäne der verkörperten Figur stellt nur Teile der emergenten Rahmenstruktur der Darstellung.	Nicht-menschliche semiotische Codes werden mithilfe menschlicher Codes dargestellt → Inkompatibilität

Tabelle 1: Typologie polyphoner Stimmwechselnetzwerke

Die interaktionale Analyse polyphoner Stimmwechsel erlaubt es, der oben zusammengefassten Kritik an MSCi zu begegnen, da sie a) klare Kriterien zur Differenzierung von mentalen Räumen anbietet, b) lediglich ein Phänomen in seiner Situiertheit beleuchtet und somit den Ubiquitätsvorwurf berücksichtigt, c) die Prozesshaftigkeit von Bedeutsaushandlung betont, und d) Bedeutung und Kognition als öffentliche Phänomene begreift.

Datensammlung und Analyseschritte

Die in der vorliegenden Arbeit untersuchten insgesamt 109 Stimmwechselsequenzen, stammen aus Aufnahmen der britischen Fernsehsendung *Never Mind The Buzzcocks* (NMTB) aus den Jahren 2006 bis 2010. NMTB ist ein wöchentlich ausgestrahltes Popmusik-Quiz, das vornehmlich ein junges, männliches Publikum anspricht. Zwei Teams treten gegeneinander an. Die Teams bestehen aus festen Teamkapitänen und jeweils zwei Gästen, die meist Popstars, Schauspieler_innen oder Komiker_innen sind. Moderiert wird das Quiz von einem Moderator. Da es keine Preise zu gewinnen gibt und die Punkte auch recht arbiträr vergeben werden, geht es der Sendung vorrangig nicht um Spannung oder die Zelebrierung von Schulwissen, sondern um die Präsentation (und oft Verspottung) der berühmten Gäste. NMTB ist berüchtigt für seinen bissigen Humor und die oft rüde Art, mit den Gästen umzuspringen. Da sich Moderatoren, Teamkapitäne, aber auch Gäste gerne und viel über andere lustig machen, ergibt sich eine hohe Dichte an polyphonen Stimmwechseln.

Fernsehdaten sind in Studien, die der Tradition der Konversationsanalyse folgen, nicht unumstritten. Die vorliegende Arbeit folgt aus diesem Grund dem Zugang von Mondada (2006, 2009) und schlägt vor, die vorliegenden Sequenzen als „Rundfunkinteraktion“ zu untersuchen. Videodaten, wie sie auch Rundfunkinteraktion darstellen, ermöglichen nie ungetrübte Einblicke in soziale Interaktion, sondern sind stets „situiertere Produkte von Videopraktiken“ (Mondada, 2009:67), denen Editierungsentscheidungen, wie das Schneiden, vorangehen. Rundfunkinteraktion ist daher Interaktion, die bereits interpretiert wurde und aktiv von weiteren Teilnehmer_innen, wie Kameralenten, Regisseur_innen und Cutter_innen, mitgestaltet wurde. Rundfunkinteraktion ist somit Interaktion, in der gewisse Aspekte von anderen Teilnehmer_innen durch Schnitt und hinzugefügten Tonspuren (wie Lachen) betont werden, um sie einem größeren Publikum verständlich zu präsentieren.

In der Analyse wurden in einem ersten Schritt Stimmwechselsequenzen im Korpus identifiziert und im Anschluss nach GAT 2 (Selting et al., 2009) transkribiert. In einem nächsten Schritt wurden zentrale Beispiele beschrieben und alle Sequenzen nach der oben beschriebenen Typologie klassifiziert. Von den 109 Sequenzen sind 23 Spiegelnetzwerke, 62 einfach gerahmte und 24 doppelt gerahmte Netzwerke. Der folgende Analyseschritt beinhaltet die interaktionale Analyse zentraler Fallbeispiele jedes Typs und die Formulierung erster Regeln. Diese wurden im Anschluss an weiteren Fällen getestet, um möglichst generelle Regeln zu finden. Diese wurden in einem letzten Schritt mit anderen Typen verglichen.

Im Folgenden wird exemplarisch zunächst eine Fallstudie zusammengefasst, die den Analyseprozess verdeutlichen soll. Im Anschluss werden die Ergebnisse des Typologievergleichs in Kapitel fünf präsentiert.

Kapitel 4: Polyphone Stimmwechsel als konzeptuelle Integrationsnetzwerke: Fallstudie

Die Sequenz „Der Metalldetektor“ stellt ein prototypisches doppelt gerahmtes Integrationsnetzwerk dar. Teilnehmer NF verkörpert einen überaus höflichen Metalldetektor aus Metall, der sich dafür entschuldigt, stets immer nur sich selbst aufzuspüren. Sowohl auf der Diskurs- als auch auf der Ebene der semiotischen Codes findet sich semantische Inkompatibilität. Vor dem eigentlichen Stimmwechsel wird eine Art ‚Vorwort‘ präsentiert, das die interaktionale Relevanz der folgenden Sequenz verdeutlicht und somit aufzeigt, dass Stimmwechsel stets lokal veranlasst sind und ihre Relation zum vorhergegangenen thematischen Diskurs transparent gemacht werden muss. Zweitens führt dieses ‚Vorwort‘ in die Problemstellung (d.h. der sich aufgrund seines Materials selbst aufspürende Metalldetektor) des Szenarios der folgenden Sequenz ein und liefert so wichtige Informationen zur narrativen Struktur der fragmentierten Erzählung. Doch nicht nur die narrative Struktur wird im Vorwort deutlich – auch auf der Ebene der semiotischen Codes werden Erläuterungen gegeben. Es wird diegetisch darauf hingewiesen, dass Metalldetektoren piepen. Mit dieser klanglichen Dimension, mit einem Wechsel der semiotischen Codes, wird dann

auch der Stimmwechsel eingeleitet: NF piept zur Einleitung der Imitationssequenz. Gleichzeitig richtet er seine Blickrichtung geradeaus, wendet seinen Körper von den anderen Teilnehmer_innen ab und orientiert sich somit nicht mehr in Richtung seiner vorherigen Gesprächspartnerin. NF zeigt dadurch, dass seine kommunikativen Signale einem anderen Diskursszenario zugeordnet sind. Diese von den anderen Teilnehmer_innen abgewandte Position behält er während der Imitationssequenz bei. Die Imitation umfasst insgesamt drei kleinere Sequenzen (bestehend aus jeweils ein bis zwei Turnkonstruktionseinheiten (TCUs)), die jeweils mit einem Piepen eingeleitet werden. Zwischen den einzelnen Blocks finden sich Pausen, in denen das Publikum seine Wertschätzung und sein Verständnis in Form von Lachen signalisiert. Der dritte und letzte Block ist prosodisch als *Pointe* markiert. Unter Rückgriff auf die im ‚Vorwort‘ eingeführte Problematik der eigenen Materialität fasst die *Pointe* das Problem nochmals zusammen und fungiert als prosodische wie auch pragmatische Endmarkierung. Nach einem abschließenden Lachen durch das Publikum und die anderen Teilnehmer_innen wird wieder zum thematischen Gespräch zurückgekehrt. Dies macht deutlich, dass solche Sequenzen tatsächlich ‚Gesprächsspiel‘ sind und als nicht-implikativ für den weiteren thematischen Gesprächsverlauf behandelt werden.

Auf kognitiv-semantischer Ebene kann das aufgebaute Integrationsnetzwerk wie in Abbildung 3 dargestellt werden. Die emergente Struktur des *blends* (d.h. der Imitationssequenz) ist doppelt gerahmt, da einer Maschine menschliche Gefühle und Sprache anverwandt werden, jedoch auch Aspekte der ‚Lebens‘- und Zeichenwelt des Metalldetektors einbezogen werden.

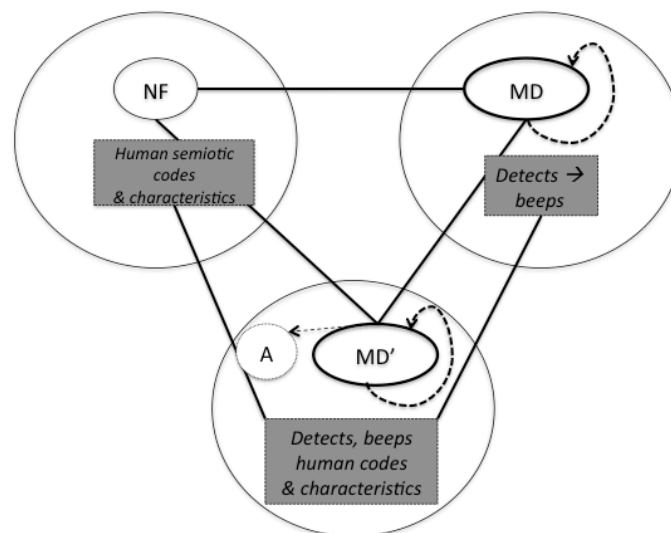


Abbildung 3: Konzeptuelles Integrationsnetzwerk in *Der Metalldetektor*

Aus dieser sowie den anderen in Kapitel vier verhandelten Fallstudien, lassen sich erste Hypothesen für die interaktionale Aushandlung von Imitationssequenzen ableiten. Die Stimmwechsel müssen aus der vorhergegangenen Interaktion resultieren und die Beziehung zum vorhergegangenen Gespräch systematisch transparent machen. Des Weiteren kann eine prototypische Struktur von Stimmwechselsequenzen beobachtet werden. ‚Vorworte‘ beziehen die präsentierten Diskursszenarien auf das vorhergegangene thematische Gespräch (z.B. „das erinnert mich daran, als...“, „Genau wie damals als ...“) und präsentieren die narrative Makrostruktur des Diskursszenarios. Demarkationsmarker, wie ein Wechsel der Blickrichtung und der Körperhaltung, signalisieren den Stimmwechsel und verdeutlichen den Referenzwechsel von der eigenen zur fremden Stimme. Die Stimmwechselsequenzen selbst sind unterteilt in kleinere Subeinheiten, die aus ein bis zwei TCUs bestehen. Nach jeder Subeinheit wird den Rezeptionssignalen des Publikums Raum gegeben und das Rederecht neu verhandelt. Spezielle Methoden zur Beendigung des Stimmwechsels beinhalten prosodisches Design im Pointen-Format sowie die markierte Wiederholung von Elementen aus dem ‚Vorwort‘. Solche Wiederholungen verdeutlichen nochmals die Relevanz der Imitationssequenz und führen gleichzeitig wieder zurück in das vorangegangene thematische Gespräch.

Im Folgenden werden diese Aspekte genauer untersucht und die Methoden, die zur Aushandlung von Stimmwechseln in den unterschiedlichen Netzwerktypen herangezogen werden, verglichen, um festzustellen, ob komplexere Netzwerke auf andere Methoden zurückgreifen als weniger komplexe.

Kapitel 5: Interaktionale Organisation von polyphonen Stimmwechseln

In diesem Abschnitt fasse ich die Ergebnisse der in Kapitel 5 der Arbeit präsentierten Analyse zusammen. Dabei orientiert sich die Struktur an den Ergebnissen der oben skizzierten Fallstudie.

Lokale Anlässe und diegetische Einführung

Spiegelnetzwerke zeigen meist sehr kurze Einführungen (‚Vorworte‘), die bereits eine Bewertung von vorangegangenen Aussagen anderer Teilnehmer_innen beinhalten.

Zudem spielt die markierte Wiederholung von Elementen dieser Aussagen eine zentrale Rolle.

In einfach gerahmten Netzwerken muss die temporale, räumliche, epistemische und modale Differenz der Input-Domänen in den Einführungen verdeutlicht werden. So finden sich Simile-Konstruktionen („x ist wie y“), aber auch die von Schegloff (1987) beschriebene ‚Zuerst ein Witz‘-Strategien, in denen ein Aspekt des vorangegangenen Gesprächs die Teilnehmer dazu veranlasst, einen Witz einzuschieben, bevor das thematische Gespräch weitergeführt wird. Die ‚Vorworte‘ zu einfach gerahmten Imitationssequenzen präsentieren somit kurze Zusammenfassungen des einzuführenden Diskursszenarios und beschreiben die Relevanz, die eine Imitation zum vorhergegangenen Gespräch hat.

Doppelt gerahmte Integrationsnetzwerke verwenden ähnliche Methoden wie die einfach gerahmten Netzwerke. Die ‚Vorworte‘ tendieren jedoch dazu, länger und detaillierter zu sein. Rhetorische Fragen sind ein gern genutztes Werkzeug, um Imitationen einzuführen.

	Distanzmarker	Verankerung im Gespräch	Pragmatische Funktion
Spiegel-Netzwerke	<ul style="list-style-type: none"> • Abgrenzungsmarkierung (z.B. „Oh“) • Nicht vorhanden 	<ul style="list-style-type: none"> • Auslöser wird durch (rhetorische) Frage profiliert • Markierte Wiederholung eines Elements des vorangegangenen Gesprächs 	<ul style="list-style-type: none"> • Zuschreibung von Identitäten, Affekt oder hintergründigen Motiven
	<ul style="list-style-type: none"> • Wiederbe- und Umschreibung der vorangegangenen Aktivität 	<ul style="list-style-type: none"> • Simile („ist wie“) • Beschreibung der vorhergegangene Aktivität 	<ul style="list-style-type: none"> • Bewertung vorangegangener Äußerungen anderer Teilnehmer_innen • Zuschreibung von Identitäten oder Affekt
Einfach gerahmte Stimmwechsel	<ul style="list-style-type: none"> • Signalisierung neuer mentaler Räume (epistemisch, modal, räumlich und temporal separat) 	<ul style="list-style-type: none"> • Gesprächsobjekt wird in eine andere Diskurswelt transferiert (Auslöser kann durch Frage oder markierte Wiederholung gekennzeichnet werden) 	<ul style="list-style-type: none"> • Illustration einer ‘zuerst ein Witz’-Sequenz
	<ul style="list-style-type: none"> • Wiederbeschreibung einer vorangegangenen Aktivität im Gespräch und Signalisierung neuer Räume 	<ul style="list-style-type: none"> • Simile („ist wie“) 	<ul style="list-style-type: none"> • Beurteilung; Objekt des Gesprächs ins Lächerliche ziehen
Doppelt gerahmte Stimmwechsel	Wie oben	Wie oben	Wie oben

Tabelle 2: Diegetische 'Vorworte' in Verkörperungssequenzen unterschiedlicher Komplexitätsstufen

Im Vergleich der drei Netzwerktypen (s. Tabelle 2) werden somit unterschiedliche Strategien deutlich, die Evidenz für die von Fauconnier & Turner (2002) vorgeschlagene Typologie beanspruchen könnten. Es muss jedoch berücksichtigt werden, dass diese Unterschiede vor allem Spiegelnetzwerke von asymmetrischen Netzwerken (einfach und doppelt gerahmten) unterscheiden. In Imitationssequenzen ermöglichen diese unterschiedlichen Netzwerktypen auch unterschiedliche soziale Aktionen: die Illustration eines Witzes in einem doppelt gerahmten Netzwerk muss sich interaktional vom Nachäffen eines Teammitgliedes unterscheiden. Der Frage, ob es die soziale Aktion oder die Netzwerkkomplexität ist, nach der sich die Methoden richten, wird aus diesem Grund in den nachfolgenden Analysen im Detail beleuchtet.

Demarkationsmarkierung: Diegese vs. Mimese

Dieser Abschnitt beschäftigt sich zum Einen mit Verben zur Einleitung von Redewiedergabe und zum Anderen mit dem in der Fallstudie bereits erwähnten Blick- und Positionswechsel zur Demarkation von diegetischen Vorwörtern und mimetischen Vorführungen.

Bei den Redewiedergabeverben fällt zunächst auf, dass die meisten Stimmwechsel ohne ein *verbum dicendi* eingeleitet werden. Wenn zur Einführung von Stimmwechseln Redewiedergabeverben verwendet werden, so sind es meist die Verben *go* und *be like*, die eine hohe Toleranz für stark körperliche Stimmwechsel aufweisen. Werden die Funktionen, die *verba dicendi* in den unterschiedlichen Integrationstypen haben, verglichen, zeigt sich, dass der Gebrauch der Verben stark vom Verkörperungsgrad abhängt und zweitens von Notwendigkeiten des lokalen Stimmen-Managements: Sobald die Charaktere einer Verkörperungssequenz etabliert sind, können die Stimmen ohne die Rahmung durch Verben des Tuns und des Sprechens verhandelt werden. Doch auch funktional unterscheiden sich die Verben. Wenn angebliche hintergründige Motive mithilfe einer Verkörperungssequenz sichtbar gemacht werden sollen, wird generell auf ein *verbum dicendi* verzichtet. Die Verwendung der *verba dicendi* erlaubt daher keine Rückschlüsse auf die Validität der MSCI-Binnenstruktur.

Bei der Betrachtung von Körperhaltung und Blickrichtung zeigt sich, dass in Spiegel-Netzwerken zwar der Blick von den bisherigen Gesprächspartner_innen abgewandt, die Körperhaltung jedoch beibehalten wird. In aggressiven Verkörperungssequenzen, in denen von der Norm abweichende Verhaltensmuster

anwesenden Personen eingeschrieben werden, wird der Blick wieder aufgenommen. In einfach gerahmten Verkörperungssequenzen wird der (Ober-)körper weiter von den bisherigen Gesprächspartner_innen abgewandt. Dabei zeigt sich auch, dass die Länge der Verkörperungssequenz Auswirkungen auf den Grad der Abwendung hat: Je länger eine Verkörperung andauert, desto mehr Abwendung lässt sich beobachten. Nochmals deutlicher wird der Grad der Abwendung in doppelt gerahmten Integrationsnetzwerken. Hier wird schon von Beginn an deutliche Diskontinuität signalisiert. Dies könnte die von Fauconnier und Turner (2002) vorgeschlagene Binnenstruktur unterstützen. Jedoch finden sich auch Sequenzen, in denen einfach und doppelt gerahmte Stimmwechsel ohne eine körperliche Abwendung stattfinden. Diese Stimmwechsel dienen dazu, anderen Teilnehmer_innen abweichende Identitäten zuzuschreiben. Zur Signalisierung der Adressat_innen werden Blick und Körper eingesetzt. Diejenige Person, auf die der Blick gerichtet wird, wird dargestellt. Kognitiv-semantische Komplexität spielt dabei keine Rolle; lediglich der Handlungsaspekt bestimmt, welche Blickrichtung und Körperhaltung eingenommen wird. Die nachfolgende Tabelle bietet einen Überblick über die sozialen Aktionen, zu denen die oben besprochenen Demarkationsmarker beitragen.

Aktion	Verbum dicendi	Blick	Körperpositionierung
Identitätszuschreibung	Ø	Zunächst vom Adressaten abgewandt; kann während der Sequenz wieder aufgenommen werden	Dem Adressaten zugewandt
Illustrationen von modal, raumzeitlich und epistemisch abgegrenzten Diskursszenarien	<i>Go</i> : Verkörperung treibt Narration voran <i>Be like</i> : visuelle / onomatopoetische Illustration diegetisch eingeführter Narrative Ø: Illustration diegetisch eingeführter Narrative	Von anderen Teilnehmer_innen abgewandt	Von anderen Teilnehmer_innen abgewandt

Tabelle 3: Demarkationsmarkierung und soziale Aktion

Gestaltung der Darbietung: Rederecht, Segmentalisierung und Listenphänomene

Da sich Verkörperungssequenzen über mehrere Turns erstrecken können, wird das ansonsten in Alltagsgesprächen greifende Sprecherwechselsystem ausgesetzt. Um solche längeren ‚Projekte‘ durchzuführen, müssen Teilnehmer_innen das Rederecht

für eine ausgeweitete Turnsequenz erhalten. Dieser Abschnitt beschäftigt sich mit den Instrumenten (wie beispielsweise prosodische Markierung), mithilfe derer das Rederecht gesichert wird. Hierzu gehört auch eine Untersuchung der Segmentalisierung solcher längeren Projekte in kleine Subeinheiten, da sich daran besonders gut die fein abgestimmte Verhandlung des Rederechts aufzeigen lässt. In einem letzten Schritt wird in diesem Abschnitt ein besonders häufiges Verkörperungsformat, das Listenphänomen, besprochen. Hier zeigt sich, dass viele Stimmwechselformen einem 3+1-Format folgen. Drei Subeinheiten liefern fragmentierte Redebeiträge der verkörperten Person. Nach einer längeren Pause wird im prosodisch markierten Pointen-Modus ein letzter Beitrag präsentiert, der die ‚Vorstellung‘ beendet und zurück zum vorhergegangenen thematischen Gespräch führt (vgl. „Der Metalldetektor“).

Im Vergleich der drei unterschiedlichen Netzwerktypen zeigt sich, dass einfach gerahmte Verkörperungen den prototypischen Typ darstellen: Sie umfassen generell mehr TCUs als die anderen Komplexitätstypen und verfügen über längere und häufigere Pausen zwischen den einzelnen Subeinheiten, die verdeutlichen, dass es sich hier um eine etablierte kommunikative Praxis handelt. Spiegel- und doppelt gerahmte Netzwerke sind hingegen kürzer und verfügen über weniger Pausen zwischen den wenigen Subeinheiten. Es scheint daher, als sei das Verkörpern von Personen außerhalb der eigenen Gruppe eine etablierte kommunikative Praxis, die eigene Werte betont und somit verstärkt, und von der Norm abweichende Aspekte auf Außenseiter überträgt. Weitere Evidenz für die von Fauconnier und Turner (2002) vorgeschlagene Binnentypologie lässt sich in der methodischen Gestaltung von Stimmwechseln jedoch nicht finden.

Beendigung von Verkörperungssequenzen

Dieser Abschnitt beschäftigt sich mit den Methoden und Praktiken, mithilfe derer eine Verkörperungssequenz beendet und ein Übergang zum vorherigen thematischen Gespräch geschaffen wird. Es zeigt sich hier, dass alle Integrationstypen auf ähnliche Mechanismen zurückgreifen. Wiederholungen, der prosodisch markierte Pointen-Modus, kurze Coda oder auch Lachpartikel in der letzten Turnkonstruktionseinheit tragen dazu bei, das Ende einer Verkörperungssequenz zu markieren. Dabei scheinen Spiegelverkörperungen tendenziell Wiederholungen zu bevorzugen, während sich der Pointen-Modus vor allem in einfach und doppelt gerahmten Verkörperungen findet.

Coda und Lachpartikel finden sich hingegen gleichmäßig über alle Typen verteilt. Um zurück zum vorigen thematischen Gespräch zu führen, wird ebenfalls unabhängig von der Komplexitätsstufe auf ähnliche Mechanismen zurückgegriffen. Beispielsweise leiten abschließende Diskursmarker wie „anyway“ über zu einer neuen Diskurseinheit.

Aktionsbasierte Typologie von Verkörperungssequenzen

In der Analyse zeigte sich, dass Teilnehmer_innen an Bedeutungsaushandlungsprozessen in polyphonen Stimmwechseln sich in der Gestaltung ihrer Äußerungen nicht an kognitiver Komplexität orientieren, sondern hauptsächlich an der sozialen Aktion. Bedeutungsaushandlung resultiert daher aus und wird geleitet von sich entwickelnden, rekursiven Annahmen (*beliefs*) über die gerade ausgehandelte soziale Aktion. Die Integrationskomplexität (im Sinne von Fauconniers und Turners Binnendifferenzierung) spielt dabei eine untergeordnete Rolle. Die Fähigkeit zur konzeptuellen Integration ist ein Werkzeug, das in der Bedeutungsaushandlung herangezogen und dann explizit signalisiert wird. Das von den Autoren vorgeschlagene Analysemodell kann aber keinerlei Aussagen über tatsächliche Bedeutungskonstitution treffen (vgl. Forschungsfrage 2). Daher wird im Folgenden zuerst eine aktionsbasierte Typologie von Verkörperungssequenzen vorgestellt und in einem nächsten Schritt ein Modell präsentiert, das die oben beschriebenen Ergebnisse berücksichtigt und so voll kontextualisierte Bedeutungsaushandlungsprozesse, die mit konzeptueller Integration einhergehen, schematisch darstellen kann.

Polyphone Stimmwechsel im interaktionalen Humor sind soziale, gemeinsam verhandelte Aktivitäten, die vor allem zwei Ziele verfolgen. Zuerst tragen Verkörperungssequenzen zur Beurteilung von vorhergegangenen Redebeiträgen bei. Diese Art von Verkörperungssequenzen folgt unterschiedlichen Formaten, wie Abbildung 4 verdeutlicht.

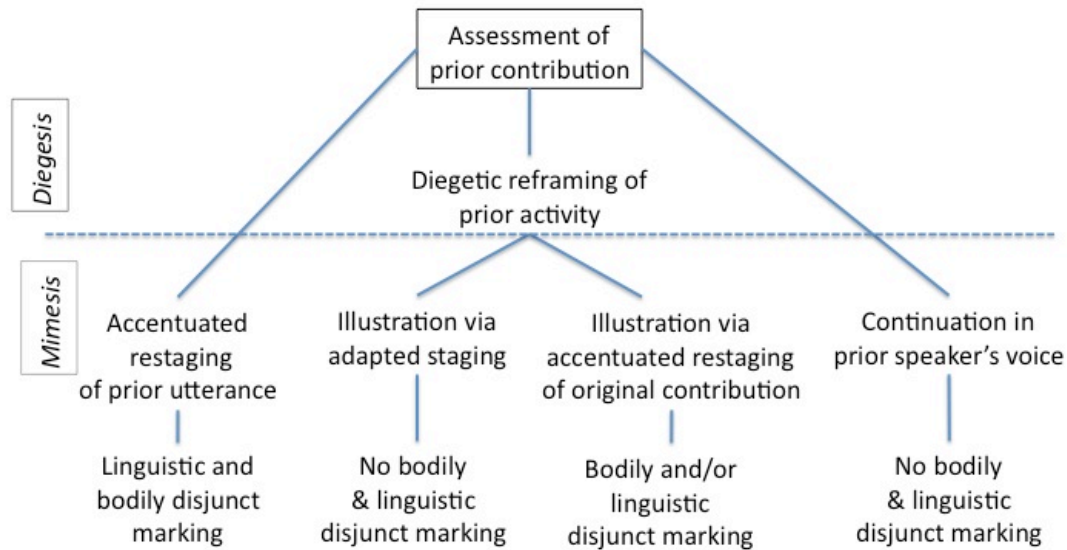


Abbildung 4: Verkörperungssequenzen zur Wertung von Redebeiträgen

Die zweite Gruppe von Verkörperungssequenzen beinhaltet narrative ‚Ein Witz zuerst‘-Sequenzen, die in separaten Diskursszenarien verortet sind. Abbildung 5 fasst die Untertypen und Strategien zusammen.

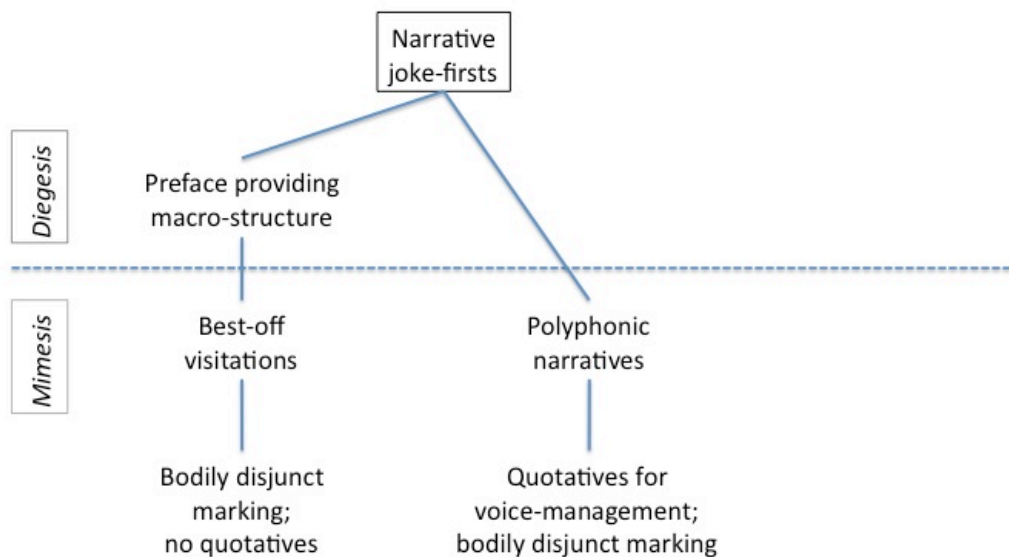


Abbildung 5: Verkörperungen zur Illustration narrativer 'Zuerst ein Witz'-Sequenzen

Die Methoden, mithilfe derer polyphone Verkörperungssequenzen interaktional ausgehandelt werden, hängen von der sozialen Aktion, die gemeinsam gestaltet wird, ab und nicht von den von Fauconnier und Turner vorgeschlagenen Typologie. Bedeutung ist eingebettet in soziale Aktion (vgl. pragmatische, gebrauchorientierte

Sprachphilosophien). Damit MSCi eine allgemeine Theorie des Denkens und der Bedeutungsaushandlung werden kann, muss sie das kooperative, aktionsorientierte und zielgeleitete Wesen von kontextuell situierten Bedeutungsaushandlungsprozessen berücksichtigen. Im Folgenden wird ein solches Modell vorgeschlagen.

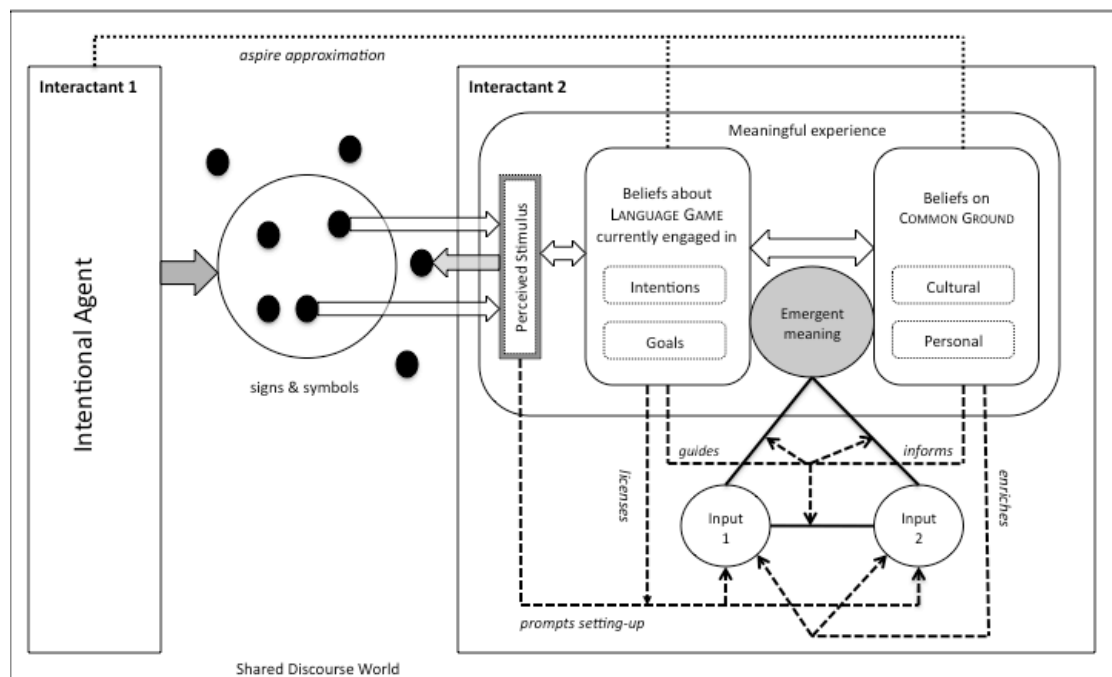
Kapitel 6: Das Primat der Pragmatik

Die oben zusammengefassten Analysen haben verdeutlicht, dass sowohl die emergente Struktur als auch die Art und Weise, wie zwei Input-Domänen in Verbindung gesetzt werden (*mappings*) von einer Suche nach interaktional relevanter Bedeutung geleitet wird. Fauconnier und Turner (2002) erwähnen Relevanz als eines der leitenden Prinzipien, die hinter einem erfolgreichen *blend* stehen können, widmen dem Phänomen jedoch nur eine Seite ihres knapp 400 Seiten umfassenden *magnum opus* und konzentrieren sich zweitens auf netzwerkimmanente Relevanz. D.h. pragmatische Relevanz wird nicht berücksichtigt. Die Analyse hat jedoch gezeigt, dass es genau diese pragmatische Relevanz ist, die auch die interne Struktur eines Integrationsnetzwerkes bedingt. Aus diesem Grund ist das Relevanzkonzept (vgl. dazu Sperber & Wilson, 1986) von zentraler Wichtigkeit im vorgeschlagenen Modell.

Das in Abbildung 6 schematisch dargestellte Modell geht von der Rezipientenseite aus und betont neben Relevanz auch die grundlegende Annahme von Kooperativität (vgl. u.a. Davidson, 1973; Grice, 1975; Tomasello, 2008) und, dem untergeordnet, Intentionalität. Es wird davon ausgegangen, dass Teilnehmer_innen in gemeinsamen Bedeutungsaushandlungsprozessen prinzipiell in der Lage sind, ähnlich kognitive Operationen durchzuführen. Diese Bedeutungsaushandlungsprozesse werden mithilfe von Zeichen² gestaltet und koordiniert. Den von den Teilnehmer_innen ausgesendeten Zeichen wird Relevanz unterstellt, d.h. ein Beitrag zu einer Form von Bedeutung. Aufgrund der Unmenge an potentiell zeichenhaften Signalen und ihrer multiplexen Relevanzen werden solche Signale durch einen Filter betrachtet, der Vorannahmen zu in der aktuell

² Der Zeichenbegriff sowie weitere Terminologien können in der Zusammenfassung nicht ausführlich dargestellt werden. Eine detaillierte Verhandlung der sich ergebenden Problemlagen findet sich in Kapitel 6 der vorliegenden Dissertation.

verhandelten Interaktion erwartbaren Zeichen bereitstellt. Zeichen, die diesen Filter passieren, werden zu Stimuli (vgl. Sperber & Wilson, 1986) und ermöglichen den interaktional koordinierten Aufbau mentaler Räume. Der ‚Inhalt‘ der mentalen Räume wird dabei von Annahmen über potentiell geteiltes Wissen, Wissen über die Gesprächspartner_innen sowie die gemeinsame Diskursgeschichte geleitet. Beschränkt wird der mentale Raum durch Annahmen zum aktuell verhandelten Sprachspiel. Das Wittgenstein’sche (1953) Konzept des Sprachspiels wird hier herangezogen, da es die gleichzeitige Berücksichtigung der Konventionalität linguistischer Zeichen, der Regelgeleitetheit von Kommunikation und der intrinsischen Vernetzung von Aktion und Bedeutung ermöglicht. Diese Annahmen über das gerade verhandelte Sprachspiel begrenzen nicht nur den ‚Inhalt‘ der mentalen Räume; sie leiten auch die zwischen den Räumen hergestellten Verbindungen (*mappings*). Die Bedeutung, die aus der Integration der beiden mentalen Räume entsteht, ist geleitet von der Suche nach Relevanz sowohl in Hinblick auf das gerade verhandelte Sprachspiel wie auch auf die Annahmen über geteiltes Wissen, Werte und Diskursgeschichte (*common ground*).



Das hier kurz zusammengefasste Modell erlaubt es, voll kontextualisierte Bedeutungsaushandlung auf Basis von konzeptueller Integration zu beschreiben und

so die von der kognitiven Semantik propagierte Aufhebung der Unterscheidung in Semantik und Pragmatik zu erreichen. Die Arbeit leistet somit einen Beitrag zur sozialen Wende in der kognitiven Linguistik.