Narrative in English Conversation

Storytelling is a fundamental mode of everyday interaction. This book is based upon the Narrative Corpus (NC), a specialized corpus of naturally occurring narratives, and provides new paths for its study. Christoph Rühlemann uses the NC’s narrative-specific annotation and XPath and XQuery, query languages that allow the retrieval of complex data structures, to facilitate large-scale quantitative investigations into how narrators and recipients collaborate in storytelling. Empirical analyses are validated using R, a programming language and environment for statistical computing and graphics. Using this unique data and methodological base, Rühlemann reveals new insights, including the discovery of turntaking patterns specific to narrative; the first investigation of textual colligation in spoken data; the unearthing of how speech reports, as discourse units, form striking patterns at utterance level; and the identification of the story climax as the sequential context in which recipient dialogue is preferentially positioned.

CHRISTOPH RÜHLEMANN is a researcher at Philipps University Marburg, Germany. He is the author of Conversation in Context (2007) and co-editor, with Karin Aijmer, of The Cambridge Handbook of Corpus Pragmatics (forthcoming).
STUDIES IN ENGLISH LANGUAGE

General editor
Merja Kytö (Uppsala University)

Editorial Board
Bas Aarts (University College London), John Algeo (University of Georgia), Susan Fitzmaurice (University of Sheffield), Christian Mair (University of Freiburg), Charles F. Meyer (University of Massachusetts)

The aim of this series is to provide a framework for original studies of English, both present-day and past. All books are based securely on empirical research, and represent theoretical and descriptive contributions to our knowledge of national and international varieties of English, both written and spoken. The series covers a broad range of topics and approaches, including syntax, phonology, grammar, vocabulary, discourse, pragmatics, and sociolinguistics, and is aimed at an international readership.

Already published in this series:

- Geoffrey Leech, Marianne Hundt, Christian Mair and Nicholas Smith: *Change in Contemporary English: A Grammatical Study*
- Jonathan Culpeper and Merja Kytö: *Early Modern English Dialogues: Spoken Interaction as Writing*
- Daniel Schreier, Peter Trudgill, Edgar Schneider and Jeffrey Williams: *The Lesser-known Varieties of English: An Introduction*
- Hilde Hasselgård: *Adjunct Adverbials in English*
- Raymond Hickey: *Eighteenth-Century English: Ideology and Change*
- Claudia Claridge: *Hyperbole in English: A Corpus-based Study of Exaggeration*
- Päivi Pahta and Andreas H. Jucker (eds.): *Communicating Early English Manuscripts*
- Irma Taavitsainen and Päivi Pahta (eds.): *Medical Writing in Early Modern English*
- Colette Moore: *Quoting Speech in Early English*
- Steven Jones, M. Lynne Murphy, Carita Paradis and Caroline Willners: *Antonyms in English: Construals, Constructions and Canonicity*
- Haruko Momma: *From Philology to English Studies: Language and Culture in the Nineteenth Century*
- Raymond Hickey (ed.): *Standards of English: Codified Varieties Around the World*
- Benedikt Szmarotsanyi: *Grammatical Variation in British English Dialects: A Study in Corpus-based Dialectometry*
- Daniel Schreier and Marianne Hundt (eds.): *English as a Contact Language*
- Bas Aarts, Joanne Close, Geoffrey Leech and Sean Wallis (eds.): *The Verb Phrase in English: Investigating Recent Language Change with Corpora*
- Martin Hilpert: *Construcational Change in English: Developments in Allomorphy, Word Formation, and Syntax*
- Christoph Rühlemann: *Narrative in English Conversation*

Earlier titles not listed are also available.
Narrative in English Conversation

A Corpus Analysis of Storytelling

CHRISTOPH RÜHLEMANN

Philipps University Marburg
To my mother and my father
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of figures</td>
<td>x</td>
</tr>
<tr>
<td>List of tables</td>
<td>xii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>xiv</td>
</tr>
<tr>
<td>List of tags</td>
<td>xvi</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1. Towards a working definition of conversational narrative</td>
<td></td>
</tr>
<tr>
<td>1.1 Introduction</td>
<td>4</td>
</tr>
<tr>
<td>1.2 Genre</td>
<td>4</td>
</tr>
<tr>
<td>1.3 Participation</td>
<td>6</td>
</tr>
<tr>
<td>1.4 Temporal sequence</td>
<td>14</td>
</tr>
<tr>
<td>1.5 Agent orientation</td>
<td>22</td>
</tr>
<tr>
<td>1.6 Narrative structure</td>
<td>24</td>
</tr>
<tr>
<td>1.7 Recipient design</td>
<td>30</td>
</tr>
<tr>
<td>1.8 Sense-making</td>
<td>34</td>
</tr>
<tr>
<td>1.9 Summary</td>
<td>37</td>
</tr>
<tr>
<td>2. Data, methods, and tools</td>
<td></td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>40</td>
</tr>
<tr>
<td>2.2 The Narrative Corpus</td>
<td></td>
</tr>
<tr>
<td>2.2.1 Corpus construction</td>
<td>40</td>
</tr>
<tr>
<td>2.2.2 Corpus annotation</td>
<td>44</td>
</tr>
<tr>
<td>2.3 Methods and tools</td>
<td></td>
</tr>
<tr>
<td>2.3.1 Methods</td>
<td>63</td>
</tr>
<tr>
<td>2.3.2 Tools</td>
<td>65</td>
</tr>
<tr>
<td>3. How do narrators and recipients co-construct turntaking?</td>
<td></td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>76</td>
</tr>
<tr>
<td>3.2 Co-construction of turn order</td>
<td></td>
</tr>
<tr>
<td>3.2.1 Introduction</td>
<td>78</td>
</tr>
<tr>
<td>3.2.2 Data and methods</td>
<td>82</td>
</tr>
</tbody>
</table>
Contents

3.2.3 Results 85
3.2.4 Discussion 88
3.2.5 Summary 90

3.3 Co-construction of turn size 92
3.3.1 Introduction 92
3.3.2 Data and methods 93
3.3.3 Results 98
3.3.4 Discussion 104
3.3.5 Summary 108

4 Recipient design I: How do narrators mark quotation? 110
4.1 Introduction 110
4.2 Interjections as quotation markers 123
4.2.1 Introduction 123
4.2.2 Data and methods 126
4.2.3 Results 128
4.2.4 Discussion 133
4.2.5 Summary 141

4.3 Pauses as quotation markers 142
4.3.1 Introduction 142
4.3.2 Data and methods 144
4.3.3 Results 146
4.3.4 Discussion 151
4.3.5 Summary 153

5 Recipient design II: How do narrators use discourse presentation for dramatization? 155
5.1 Introduction 155
5.2 Climactic structure at micro level: Sequential ordering of report units in utterances 156
5.2.1 Introduction 156
5.2.2 Data and methods 157
5.2.3 Results 160
5.2.4 Discussion 167
5.2.5 Summary 171

6 How do recipients co-author stories? 173
6.1 Introduction 173
6.2 How do recipients influence narrator verbosity? 176
6.2.1 Introduction 176
6.2.2 Data and methods 178
6.2.3 Results 180
6.2.4 Discussion 184
6.2.5 Summary 187
Contents ix

6.3 How do Co-constructive Recipients co-tell stories? The case of recipient dialog 188
  6.3.1 Introduction 188
  6.3.2 Data and methods 197
  6.3.3 Results 204
  6.3.4 Discussion 207
  6.3.5 Summary 216

7 Conclusions and implications 218
  7.1 Introduction 218
  7.2 Summaries of the analytical chapters and conclusion 218
  7.3 Implications and outlook: a plea for annotation-driven corpus research 228

Appendix 1 232
Appendix 2 237
Appendix 3 239
Appendix 4 242
Notes to the text 246
References 274
Index 287
# Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Conversational and narrative subgenres</td>
<td>7</td>
</tr>
<tr>
<td>1.2</td>
<td>Participation framework for conversational narrative</td>
<td>9</td>
</tr>
<tr>
<td>1.3</td>
<td>Cyclical structure of text KB7-N2</td>
<td>30</td>
</tr>
<tr>
<td>2.1</td>
<td>Componential structure of texts in the NC</td>
<td>52</td>
</tr>
<tr>
<td>2.2</td>
<td>Mean utterance length in words across participant roles</td>
<td>55</td>
</tr>
<tr>
<td>2.3</td>
<td>Proportions of quotative lemmas</td>
<td>57</td>
</tr>
<tr>
<td>2.4</td>
<td>Frequencies and proportions of reporting mode types in narrative</td>
<td>61</td>
</tr>
<tr>
<td>2.5</td>
<td>Proportions of six speech presentation categories in the LSWTPSC and the NC</td>
<td>62</td>
</tr>
<tr>
<td>2.6</td>
<td>Boxplots of lengths of PRR and PRC responses</td>
<td>67</td>
</tr>
<tr>
<td>2.7</td>
<td>Tree structure of text KB1-N1</td>
<td>72</td>
</tr>
<tr>
<td>3.1</td>
<td>Linear regression model of the proportions of N-not N-N trigrams as a function of all trigrams per story</td>
<td>88</td>
</tr>
<tr>
<td>3.2</td>
<td>Turn size fluctuation formula adapted from Leech and Fallon (1992)</td>
<td>95</td>
</tr>
<tr>
<td>3.3</td>
<td>TSF coefficients for text KB9-N2</td>
<td>96</td>
</tr>
<tr>
<td>3.4</td>
<td>Boxplots of turn size fluctuation coefficients per narrative across pre-narrative, narrative, and post-narrative components</td>
<td>99</td>
</tr>
<tr>
<td>3.5</td>
<td>Lengths of turns by Primary Narrator immediately preceding turns by Co-narrator, Co-constructive Recipient, and Responsive Recipient as well as lengths of PNC, PRC, and PRR turns</td>
<td>102</td>
</tr>
<tr>
<td>3.6</td>
<td>Left panel: boxplots of TSF coefficients for Responsive Recipient, Co-constructive Recipient and Co-narrator; right panel: barplots of respective IQRs.</td>
<td>103</td>
</tr>
<tr>
<td>3.7</td>
<td>Cline for TSF</td>
<td>104</td>
</tr>
<tr>
<td>3.8</td>
<td>Discoursal intertwining as an effect of discourse presentation</td>
<td>111</td>
</tr>
<tr>
<td>3.9</td>
<td>Cline of immediacy as a function of pretended authorial interference (see Leech &amp; Short 1981)</td>
<td>113</td>
</tr>
</tbody>
</table>
List of figures

Figure 4.3 Histograms with density curves of positions of interjections, including well, in conversational utterances and quotes longer than one word

Figure 4.4 Boxplots of positions of the ten most frequent interjections in conversational utterances

Figure 4.5 Boxplots of positions of the ten most frequent interjections in quotes

Figure 4.6 Boxplots of positions of oh and well occurring (i) within narrative utterances in quote-first position and (ii) within conversational utterances

Figure 4.7 Histograms of positions of silent pauses (<pause/>), and the filled pauses er and erm taken together within the positional scope of quotes

Figure 4.8 Histogram of positions of silent pauses in quotes with confidence bands calculated on the basis of a bootstrap (B = 2000, α = 0.05)

Figure 5.1 Boxplot (left panel) and barplot (right panel) showing the distribution of number of report units across 1371 utterances

Figure 5.2 Immediacy cline with associated ordinal values

Figure 5.3 ECDF plots of immediacy values for first position (pos1) and last position for 3 ≤ n ≤ 8 report unit sequences

Figure 5.4 ECDF plots of immediacy values for all positions for 3 ≤ n ≤ 8 report unit sequences

Figure 5.5 ECDF plots of mimesis values for sequences of 3 ≤ n ≤ 8 report units

Figure 5.6 ECDF plots of immediacy and mimesis values for n=2-sequences of report units

Figure 5.7 Functional framework for immediatization trend

Figure 6.1 Proportions of top 10 most frequent word classes in PRR and PRC utterances

Figure 6.2 Boxplots of total number of words in narrator utterances per story in stories with vs. stories without recipient feedback

Figure 6.3 Meansplots for number of words in narrator utterances per story as a function of number of recipient responses per story in subsets 1–3

Figure 6.4 3D plot of Model 4

Figure 6.5 Barplots of frequencies of discourse presentation types in Co-constructive Recipient utterances

Figure 6.6 Quote-climax-quote triangle

Figure 6.7 Left panel: Density lines and medians for positions of quotes by Co-constructive Recipients and narrators; right panel: Bootstrap confidence band for PRC quotes
Tables

Table 1.1 Characteristics of the recipient roles Responsive Recipient and Co-constructive Recipient
Table 2.1 Distribution of male and female participants in the narrative components of the NC
Table 2.2 Distribution of narrative participants in the narrative component according to social class
Table 2.3 Distribution of participants in the narrative component according to age group
Table 2.4 Narrative Corpus – basic statistics
Table 2.5 Distribution of the 531 narratives in the NC across narrative types
Table 2.6 Participant values in narrative components
Table 2.7 Quotatives and their tags in the NC
Table 3.1 Possible turn trigrams in three-party conversational narrative
Table 3.2 Turn trigrams in KE1-N1
Table 3.3 Number of trigrams and N-notN-N trigrams in data (21 three-party narratives)
Table 3.4 Number of words per turn and TSF coefficients in the pre-narrative, narrative, and post-narrative components of text KB9-N2
Table 3.5 Medians and IQRs for components of text KB9-N2
Table 3.6 Data analyzed for turn size fluctuation
Table 3.7 Average medians and IQRs of TSF coefficients
Table 3.8 Median lengths of narrator and recipient turn pairs
Table 4.1 Number of interjections, including well, in quotes and conversational utterances longer than one word (n > 1)
Table 4.2 10 most frequent ITJ in conversational utterances
Table 4.3 10 most frequent ITJ in quotes
Table 4.4 Chi-squared statistics for oh and well in quote- and utterance-first position (df = 1)
Table 4.5 Frequencies of fuck, wow, urgh/ugh, and aargh/ow in quote- and utterance-first position
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.6</td>
<td>Frequencies of <em>er, erm</em> and <code>&lt;pause/&gt;</code> occurring within the positional scope of quotes</td>
<td>146</td>
</tr>
<tr>
<td>5.1</td>
<td>Basic statistics of utterances with sequences of $3 \leq n \leq 8$ report units</td>
<td>159</td>
</tr>
<tr>
<td>5.2</td>
<td>Results of tests for trends in proportions for sequences of $3 \leq n \leq 8$ report units within utterances (df = 1)</td>
<td>163</td>
</tr>
<tr>
<td>5.3</td>
<td>Report types with associated mimetic values</td>
<td>164</td>
</tr>
<tr>
<td>5.4</td>
<td>Results of test for trends in proportions for sequences of $3 \leq n \leq 8$ report units within utterances (binary values)</td>
<td>166</td>
</tr>
<tr>
<td>6.1</td>
<td>10 random examples of PRR and PRC utterances each</td>
<td>174</td>
</tr>
<tr>
<td>6.2</td>
<td>Test statistics for Kendall’s rank correlation tests for subsets 1–3</td>
<td>181</td>
</tr>
<tr>
<td>6.3</td>
<td>4 Negative Binomial models for the association of narrator verbosity and recipient activity</td>
<td>182</td>
</tr>
<tr>
<td>6.4</td>
<td>Forms of Co-constructive Recipient contributions</td>
<td>189</td>
</tr>
</tbody>
</table>
This book is about the co-construction of narrative in conversation. Fittingly, it is itself the result of intense co-construction between the author and numerous people who have contributed substantially and whom I owe gratitude.

The first thank you goes to Hans-Jörg Schmid at Ludwig-Maximilians-University, Munich, who provided generous assistance in the early stages when the corpus had to be built and annotated from scratch. As the book progressed he offered invaluable criticism of earlier drafts of individual chapters. His comments and suggestions were most insightful and constructive, motivating me to go to extreme lengths to make this study a worthwhile read. I am also deeply indebted to Rolf Kreyer at Marburg University, who not only read and commented on the manuscript in very great detail, providing most invaluable feedback, but also supervised the final stages of the post-doctoral thesis from which this book arose.

Another great thank you is due to Neal Norrick at Saarbrücken University. His outstandingly diverse and innovative research on conversational narrative has inspired a great many questions I have pursued in the book. He also supported this work immensely by inviting me to his workshop at the Boston ISLE2 conference and by inviting me to give a guest lecture at Saarbrücken University in fall 2011. On both occasions I presented case studies reported on in this book.

Matthew Brook O’Donnell at the University of Michigan provided invaluable help not only with the compilation and annotation of the Narrative Corpus but also with the query languages XPath and XQuery, without which hardly any case study in this book would have been feasible. His willingness to share his outstanding expertise in handling these technologies as well as his friendship are greatly appreciated.

Another friend to whom I am greatly indebted is Andreas Bagoutdinov, who did most of the ground work for the corpus and who often helped me with computer and query issues despite difficult personal circumstances. Franca Kirchberg’s assistance too was essential. She not only helped me get started with R and statistical thinking but also carefully advised me on how to use it with the case studies.
Further, I’ve learnt a lot from Stefan Th. Gries’s books on statistics in corpus linguistics and his R workshop held in Denton/Texas in 2011. Stefan was the first to demonstrate to me in fuller detail the breathtaking possibilities R offers. I am also greatly indebted to him for invaluable comments on an earlier version of the case study on turn order in Section 3.2 and for useful tips given in response to queries posted to the Google Group “Corp Ling with R” (http://groups.google.com/group/corpling-with-r) which he created and maintains. I am likewise grateful for inspirational thoughts on earlier versions of individual case studies contributed by Michael Hoey (Section 4.2) and Rebecca Clift (Section 6.4). Special warm thanks are due to Merja Kytö, general editor of the Studies in English Language series, and Helen Barton, commissioning editor at Cambridge, whose endorsement all the way through from proposal to publication was unfailing. The Cambridge University Press production team too deserve a big thank you. Under the guidance of assistant editor Fleur Jones, production editor Beata Mako and copy editor Cheryl Prophett have greatly helped out the finishing touches to the manuscript. Despite all these significant contributions, the usual disclaimer applies: all remaining errors and weaknesses of this book are entirely my fault.

Finally, I thank my family, my wife Andrea and my sons Ricardo and Lionel, not only for bearing with me during the years I have diverted much of my attention to this book (rather than to them) but also for providing the lively and loving environment without which such an endeavor would hardly be possible. Also, I owe much to my father-in-law Fred and, even more, to my parents, Günther and Jolanthe Rühlemann. It was undoubtedly their stories of experiences of love, hunger, despair, struggle, and hope that taught me the centrality of storytelling to who we believe we are and who we wish to be.
## List of tags

### Textual component
- CPR = Pre-narrative conversation  
- CNN = Narrative  
- CNI = Narrative-initial utterance  
- CNF = Narrative-final utterance  
- CNI-CNF= Utterance which is both narrative-initial and narrative-final  
- CPO = Post-narrative conversation  

### Embed level
- ES = Stand-alone narrative  
- EC = Narrative chain (with sublevels EC1, EC2, EC3)  
- EN = Nested narrative  

### Reporting mode
- MDD = Direct  
- MDF = Free Direct  
- MII = Indirect  
- MIF = Free Indirect  
- MSS = Representation of Speech Act  
- MVV = Representation of Voice  
- MVT = Representation of Voice with Topic  
- MUU = Representation of Use  
- MRR = Anaphoric Reference to Discourse Presentation  
- MRQ = Request for Discourse Presentation  
- MXX = Unclear Mode  

### Participation role
- PNU = Unsupported Narrator  
- PNS = Supported Narrator  

xvi
List of tags  xvii

PNP = Primary Narrator
PNC = Ratified Co-narrator
PRR = Responsive Recipient
PRC = Co-constructive Recipient
PXX = Unclear Role
PX0 = Non-narrative Role

Quotatives
QSB = Base form “say”
QSZ = Third-person singular present tense form “says”
QSD = Past tense form “said”
QSG = Progressive form “saying”
QSN = Past participle form “said”
QTB = Base form “think,” QTZ, QTD, QTG, QTN
QGB = Base form “go,” QGZ, QGD, QGG, QG
QBB = Base form “be like,” QBM = “m/am like”, QBZ = “s/is like”,
     QBR = “re/are like,” QBD = “was/were like,” QBG = “being like,”
     QBN = “been like”
QAB = Base form “ask,” QAZ, QAD, QAG, QAN
QLB = Base form “tell,” QLZ, QLD, QLG, QLN
QLK = Quotative “like” without preceding form of “be”
QOO = Other Quotatives

Narrative type
T1o = First-person experience narrative
T1G = First-person generalized recurrent experience story
T1D = First-person dreamtelling
T1F = First-person fantasy
T1M = First-person mediated story
T3o = Third-person experience narrative
T3G = Third-person generalized recurrent experience story
T3D = Third-person dream
T3F = Third-person fantasy
T3M = Third-person mediated story
T3J = Joke