

## Multimodality and Genre

*This page intentionally left blank*

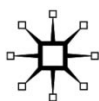
# Multimodality and Genre

## A Foundation for the Systematic Analysis of Multimodal Documents

John A. Bateman

*University of Bremen*

palgrave  
macmillan



© John A. Bateman 2008

All rights reserved. No reproduction, copy or transmission of this publication may be made without written permission.

No paragraph of this publication may be reproduced, copied or transmitted save with written permission or in accordance with the provisions of the Copyright, Designs and Patents Act 1988, or under the terms of any licence permitting limited copying issued by the Copyright Licensing Agency, 90 Tottenham Court Road, London W1T 4LP.

Any person who does any unauthorised act in relation to this publication may be liable to criminal prosecution and civil claims for damages.

The author has asserted his right to be identified as the author of this work in accordance with the Copyright, Designs and Patents Act 1988.

AD  
First published 2008 by  
PALGRAVE MACMILLAN  
Houndmills, Basingstoke, Hampshire RG21 6XS and  
175 Fifth Avenue, New York, N. Y. 10010  
Companies and representatives throughout the world

PALGRAVE MACMILLAN is the global academic imprint of the Palgrave Macmillan division of St. Martin's Press, LLC and of Palgrave Macmillan Ltd. Macmillan® is a registered trademark in the United States, United Kingdom and other countries. Palgrave is a registered trademark in the European Union and other countries.

ISBN 978-1-349-28079-7 ISBN 978-0-230-58232-3 (eBook)  
DOI 10.1057/9780230582323

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources. Logging, pulping and manufacturing processes are expected to conform to the environmental regulations of the country of origin.

A catalogue record for this book is available from the British Library.

A catalog record for this book is available from the Library of Congress.

10	9	8	7	6	5	4	3	2	1
17	16	15	14	13	12	11	10	09	08

# Contents

<i>List of Tables</i>	<b>ix</b>
<i>List of Figures</i>	<b>xi</b>
<i>Acknowledgements</i>	<b>xvii</b>
<i>Preface</i>	<b>xix</b>
<b>1 Introduction: Four Whys and a How</b>	<b>1</b>
1.1 Learning to walk: framing issues and analytic focus . . . . .	3
1.1.1 Why multimodality? . . . . .	3
1.1.2 Why ‘documents’? . . . . .	7
1.1.3 Why genre? . . . . .	9
1.1.4 Why analysis? . . . . .	11
1.2 How can we analyse multimodal documents? . . . . .	13
1.2.1 An orientation for analysis: empirical linguistics . . . . .	14
1.2.2 A framework for empirical analysis: the GeM model . . . . .	15
1.2.3 Structure of the book . . . . .	19
<b>2 Multimodal Documents and their Components</b>	<b>21</b>
2.1 Starting points: how to find document parts? . . . . .	24
2.2 The page as an object of interpretation . . . . .	27
2.2.1 Interpretation within document design . . . . .	28
2.2.2 Multimodal linguistics . . . . .	38
2.3 The page as object of perception . . . . .	57
2.4 Page as signal . . . . .	65
2.5 The Page as object of production . . . . .	74
2.5.1 Describing a page for design . . . . .	75
2.5.2 Describing a page for rendering . . . . .	85
2.5.3 Producing a page from intentions: automatic document generation . . . . .	91
2.6 Combining viewpoints on document parts . . . . .	103
<b>3 The GeM Model: Treating the Multimodal Page as a Multilayered Semiotic Artefact</b>	<b>107</b>
3.1 The GeM Model: the base layer . . . . .	110
3.2 The GeM presentation layers: the layout base . . . . .	115
3.2.1 Layout segmentation: identification of layout units . . . . .	116

3.2.2	Realisation information . . . . .	117
3.2.3	Layout structure . . . . .	121
3.3	A more complicated example of layout analysis . . . . .	129
3.3.1	The parts of the Louvre . . . . .	130
3.3.2	The layout of the Louvre . . . . .	134
3.4	Conclusion . . . . .	142
<b>4</b>	<b>The Rhetorical Organisation of Multimodal Documents</b>	<b>143</b>
4.1	Rhetoric and multimodal documents: our starting points . .	144
4.2	A brief introduction to <i>Rhetorical Structure Theory</i> . . . .	146
4.2.1	The RST rhetorical relations . . . . .	147
4.2.2	The RST rhetorical structure . . . . .	150
4.3	The move to multimodal RST: the GeM rhetorical layer . .	151
4.3.1	André's extension of RST . . . . .	152
4.3.2	Problems with traditional multimodal RST . . . . .	155
4.3.3	Multimodal relationals: subnuclear elaboration . .	160
4.4	Example analyses: rhetorical relations between layout units	163
4.4.1	Mismatches between layout structure and intended rhetorical structure . . . . .	166
4.4.2	Explaining how to use a telephone . . . . .	171
4.5	Conclusion . . . . .	174
<b>5</b>	<b>Multimodal Documents and Genre</b>	<b>177</b>
5.1	Perspectives on genre . . . . .	183
5.1.1	Genre as social semiotic . . . . .	184
5.1.2	Genre as social action . . . . .	188
5.1.3	Genre: the need for fine-grained descriptions . . .	194
5.2	The move to multimodal genre . . . . .	196
5.2.1	Multimodal moves within linguistic and rhetorical approaches to genre . . . . .	197
5.2.2	Moving in on genre from the visual . . . . .	201
5.2.3	Cybergenres: a brief critique . . . . .	209
5.3	Representing genre . . . . .	217
5.3.1	Genre typology . . . . .	219
5.3.2	Genre topology . . . . .	223
5.4	The multimodal genre space . . . . .	225
5.5	Illustrations of genre: tracking change . . . . .	229
5.5.1	Field guides across time . . . . .	229
5.5.2	Wildlife fact files across time . . . . .	240
5.6	Discussion and conclusion . . . . .	246

<b>6</b>	<b>Building Multimodal Document Corpora: the State of the Art</b>	<b>249</b>
6.1	Corpus-based linguistics . . . . .	250
6.2	The origin and representation of annotated corpora . . . .	252
6.2.1	Annotated corpora: early days . . . . .	252
6.2.2	Applying XML to corpus design . . . . .	254
6.2.3	Annotation problems with complex data . . . . .	260
6.3	The move to multimodal corpora . . . . .	264
6.4	The GeM model as a corpus annotation scheme . . . . .	267
6.5	Conclusions and recommendations . . . . .	272
<b>7</b>	<b>Conclusions and Outlook: What Next?</b>	<b>273</b>
	<i>Bibliography</i>	<b>279</b>
	<i>Author Index</i>	<b>301</b>
	<i>Subject Index</i>	<b>307</b>

*This page intentionally left blank*



# List of Tables

1.1	The primary sources of constraints adopted by the Genre and Multimodality framework . . . . .	18
1.2	The primary layers of the Genre and Multimodality framework . . . . .	19
2.1	The primary Gestalt laws of perception . . . . .	61
3.1	Table of page elements to be identified as base units during analysis . . . . .	111
3.2	Base units of the Gannet page . . . . .	113
3.3	Base units of the Louvre two-page spread . . . . .	133
3.4	Extract of the correspondence between layout units and base units . . . . .	136
4.1	List of RST relations with the constraints they impose on their respective nuclei and satellites . . . . .	149
5.1	Extract from Orlikowski and Yates's definition of coding criteria for their investigation of genres deployed within an extended email-mediated organisational effort (Orlikowski and Yates 1994, p552). . . . .	195
5.2	Selection of the results of Shepherd and Watters' analysis of 96 websites according to their content, form and functionality scheme . . . . .	213

*This page intentionally left blank*

# List of Figures

1.1	A sample of varied document pages which raise, to differing degrees and in differing ways, issues of multimodal interpretation . . . . .	3
1.2	An extract from a domestic gas bill that is attempting to communicate how much the consumer should pay (or not)	4
1.3	A continuum of visual-textual deployment . . . . .	10
1.4	The GeM model . . . . .	16
2.1	Spectrum of approaches to a document and its pages . . . .	25
2.2	An example page from an early field guide book on birds .	26
2.3	Possible layout compositions according to the information value portion of Kress and van Leeuwen's 'visual grammar'	45
2.4	Examples of a diagram and a page layout that have received ideological interpretations according to Kress and van Leeuwen's framework . . . . .	47
2.5	A page layout and its functionally-motivated layout clusters	55
2.6	Good continuation and bad continuation . . . . .	59
2.7	Examples of pre-attentive perception at work . . . . .	62
2.8	Eye-tracking results . . . . .	63
2.9	Successive reduction of resolution in a page image . . . .	67
2.10	Using discriminability at different resolutions to reveal layout elements . . . . .	68
2.11	Alternative visual filters applied to an extract of a newspaper page . . . . .	70
2.12	Consecutive construction of an XY-tree . . . . .	72
2.13	Two alternative layouts that receive the same XY-tree description . . . . .	74
2.14	Construction of line segments obeying the proportionalities of the Golden Ratio . . . . .	77
2.15	Examples of page layouts produced by Le Corbusier . . . .	79
2.16	Grid designed by Gerstner . . . . .	82
2.17	A modular grid for a more complex Gannet page . . . . .	84
2.18	HTML source and its rendering in a web-browser . . . . .	87
2.19	Flow of information in an XML-based document preparation scheme . . . . .	89

2.20	A formatted extract from a patient information leaflet adapted from Power, Scott and Bouayad-Agha (2003, p226) together with a more simply formatted variant . . . .	95
2.21	Document structure from Power <i>et al.</i> . . . . .	97
2.22	Example of layout structure and its realisation on the page as developed within the DArt <sub>bio</sub> multimodal page generation system . . . . .	102
2.23	The page as a site of cooperation and integration of distinct semiotic modes . . . . .	106
3.1	The distribution of base elements to layout, rhetorical, generic and navigational elements . . . . .	109
3.2	The Gannet example page repeated . . . . .	112
3.3	Example of differential use of leading in the example Gannet page . . . . .	119
3.4	Layout structure for the 1972 Gannet example page according to the GeM model . . . . .	123
3.5	Graphical representation of the general method of correspondence used to relate layout structure and the area model in the GeM framework . . . . .	126
3.6	Correspondence between the layout structure and the area model of the 1972 Gannet page . . . . .	128
3.7	A page from a Dorling-Kindersley guide to Paris: description of the Louvre . . . . .	131
3.8	Base units of the Louvre page shown for ease of reference using the page as background . . . . .	135
3.9	Example layout positioning of the ‘facade’ information from several pages of the Dorling-Kindersley <i>Eyewitness Travel: Paris</i> guide . . . . .	138
3.10	Layout structure for the Louvre page according to the GeM model . . . . .	139
3.11	A grid for the Louvre page . . . . .	140
3.12	The complex area model for the Louvre page, combining a geometric shape and a vertical grid . . . . .	141
3.13	Correspondence between the layout structure and the area model for the Louvre page . . . . .	142

4.1	The graphical representation of the RST analysis of the 'text': "In the event of fire, open the doors by pressing the red button. Then leave the building."	151
4.2	A multimodal rhetorical structure diagram describing instructions for using a coffee machine from (André 1995)	153
4.3	Example labels taken from the Louvre page	161
4.4	Problematic Louvre inset	165
4.5	Fact sheet extract: features of the Bengal Tiger from the <i>Wildlife Explorer</i> series	167
4.6	Layout structure for the tiger page extract	167
4.7	Rhetorical structure for the tiger page extract	169
4.8	Mismatching layout and rhetorical organisations in the tiger page extract	169
4.9	Extract from the fact file from the <i>Discovering Wildlife</i> series	170
4.10	Instructions for setting up a phone	172
4.11	Layout structure and a corresponding rhetorical structure for the telephone page extract	173
4.12	Intended rhetorical structure for the telephone page extract	174
4.13	Three semiotic modes commonly deployed within document pages	175
5.1	The positions taken by genre in the design process according to Waller (1987a, pp298–301; Figures 9.5–9.9)	179
5.2	Two versions of <i>The Guardian</i> newspaper—one print, the other online	180
5.3	Graphical representation of the genre structures of Hasan (1984) for nursery tales and of Labov and Waletzky (1978) for narrative	186
5.4	Characterisation of the cyclic process of genre change according to Yoshioka and Herman (2000)	191
5.5	Graphical representation of the development of the memo genre as discussed by Yates and Orlikowski (1992, p315)	193
5.6	The verbal and numerical row of Twyman's schema of 'graphic language' together with examples from Twyman (2004)	203
5.7	Constructed analysis of an example newspaper page	209
5.8	Cybergenres and their development according to Shepherd and Watters (1998)	214
5.9	Development of parallel distinct online news genres	215

5.10	Contrasting Shepherd and Watters (1998) view of cybergenres with the approach proposed within the GeM model . . .	216
5.11	Network for classifying members of the factual genre family (Martin 1992) . . . . .	220
5.12	An example of spatial layout constraints and a possible page solution . . . . .	222
5.13	Example of overlapping genres adapted from Lemke (1999)	224
5.14	The genre space and some recognisable regions lying along its boundary . . . . .	226
5.15	Two example documents that share a centre-margin organisation but which vary considerably in other respects . . . .	227
5.16	Documents that all remain within broadly typographical, linear-interrupted genres ('text-flow') . . . . .	228
5.17	Documents ranging from single entry points to multiple entry points (non-linear directed viewing) . . . . .	228
5.18	Four Gannet pages . . . . .	231
5.19	Generalised content structure for the Gannet pages . . . .	232
5.20	The area models for the four Gannet pages . . . . .	232
5.21	The contrasting associations of a common layout structure with the area models of the 1972 and 1994 Gannet pages .	233
5.22	Modal presentations of a fragment of the Gannet content structure across time . . . . .	237
5.23	Mode distribution across the four Gannet pages . . . . .	238
5.24	Mode distribution across the four Gannet pages shown as a continuous development . . . . .	239
5.25	Three fact file entries for the Nile Crocodile . . . . .	241
5.26	Composite layout presenting information about the distribution of the Kingfisher . . . . .	243
5.27	Layout grid for the 2002 animal fact sheet for the Northern Gannet . . . . .	245
5.28	Contrasting layout organisation taken up by the central spreads of the three editions of the animal fact files . . . .	246
6.1	Extract from the Lancaster-Oslo-Bergen (LOB) corpus . . .	252
6.2	Extract from the SUSANNE corpus (slightly abbreviated) .	253
6.3	Example of XCES-conformant annotation adapted from the TEI Guidelines . . . . .	258
6.4	Example of a common mark-up situation and an ill-formed XML attempt at its representation . . . . .	261
6.5	Example of stand-off annotation . . . . .	263

6.6	The inter-relationships between base units and the components of the GeM layout model . . . . .	268
6.7	A GeM-style XML representation of a Gannet page showing base, layout and RST layers . . . . .	270
7.1	Three versions of the tiger fact sheet image . . . . .	276

*This page intentionally left blank*



# Acknowledgements

Many of the results and ideas presented in this book are the results of joint development with my colleagues in the Genre and Multimodality project: Judy Delin and Renate Henschel. The project itself was funded by the British Economic and Social Research Council (ESRC) from 1999 to 2002. Former members of the project were Patrick Allen and Stewart Pleace, who also played a significant role in the project's early directions and approach.

The project also involved direct input from our design collaborators: The Guardian (Roger Browning), HarperCollins (Myles Archibald), The Herald, JET Documentation Services (Jane Teather), and Enterprise IDU (Rob Waller). We thank them for their cooperation and for making materials available to us for analysis and presentation.

In addition, the author thanks the following for assistance during the preparation of this book: Maria Brady (International Masters Publishers Ltd., London) for providing QuarkXpress files of the animal fact sheets discussed in Chapters 4 and 5 and Dr. Klimis Mastoridis (University of Macedonia Press) for pdf-files of Prof. Michael Twyman's article discussed in Chapter 5.

Finally, the author is grateful to the following for permission to quote from copyright works:

- Prof. Elisabeth André for permission to use the material shown in Figure 4.2.
- Dorling-Kindersley for permission to show the Louvre page spread from *Eyewitness Travel: Paris*, 2007, pp122-123, used in Figures 1.1, 3.7, 3.8, 4.3, 5.15.
- Elsevier for permission to reproduce Figure 5.5 from *Information Visualization: Perception for Design*, Colin Ware, Page 166, 2000, ISBN 1-55860-511-8; used in Figure 2.7.
- *The Guardian* for QuarkXPress files and permission to show the pages used in Figures 1.1, 2.11, 5.2, 5.7, 5.17.
- HarperCollins for the bird pages and extracts used in Figures 1.1, 2.17, 5.17, 5.18, 6.7.
- Prof. George Herman for permission to use the diagram shown in Figure 5.4.
- Dr. Jana Holsanova for permission to use the eye-tracking displays shown in Figure 2.8.
- International Masters Publishers Ltd. for permission to use the animal fact sheets (*Wildlife Fact File*, 1990; *Wildlife Explorer*, 1997;

*Discovering Wildlife*, 2002) shown in Figures 1.1, 4.5, 4.9, 5.25, 5.26, 5.27, 5.28, 7.1.

- *The Journalist* (Taiwan) for permission to use the magazine spread that appears in Figures 1.1, 5.17.
- Prof. Michael Twyman for permission to use the diagram shown in Figure 5.6.

The author offers his apologies should the rights of any further copyright holders have been infringed upon unknowingly.

# Preface

This book is intended primarily (but not, of course, exclusively!) for three distinct groups of readers; each of which has its own starting points, requirements and interests. It is hoped that the focus of the book, multimodal documents, will provide a meeting point where these communities may find some useful common ground.

The first group is formed by that growing body of researchers and students who are looking for ways of analysing communication anchored in *combinations* of distinct semiotic modes—in the present case, verbal language and visual presentations such as pictures and diagrams. There is a severe deficit in re-usable methodologies for guiding such investigations; this book seeks to provide one.

The second group consists of those who are working in any particular area concerned with documents—be that design, automatic document analysis, literacy, and so on—and who are interested in seeing how other research communities are approaching issues of multimodality. Young disciplines are often fragmented and different but closely related areas of knowledge can develop in distinct research communities without, at the outset, awareness of each other's efforts. Connections across the distinct communities are then surprisingly infrequent. At a more general level, therefore, this book is also concerned with bridging some of these gaps and in further consolidation of the field. For example: what can the practical process of automatic document recognition gain from abstract semiotic analysis of document rhetoric? Or what can abstract analyses learn from image analysis? There are many such questions and also much to learn.

The last group consists of those who have already been working on multimodality and its description for some time—for those who are familiar with the ground-breaking work of Kress and van Leeuwen on the *Grammar of visual design*, of Kress on multimodal genre, of O'Toole on displayed art, of Lemke on multimodal genre topology and the like, and who ask themselves: what now? When we have already moved to a position where meaning-making is accepted to be essentially multimodal and to rely on an extended notion of genre, one which necessarily includes artefactual components of production, dissemination and consumption, what then? What consequences does this have for how we can structure future investigations into multimodal meaning-making?

This book suggests one possible direction, one which is particularly aimed at research in a 'post-Kress & van Leeuwen' world of pervasive multimodality.