Multimodal genre analysis

Hiippala, Tuomo

Genre analysis

Tuomo Hiippala
Department of Modern Languages
University of Helsinki

1   Introduction

The concept of genre is frequently invoked in multimodal analysis for a variety of reasons. In most cases, genre is used to describe multimodal phenomena and their properties. From a methodological perspective, it may be argued that genre is often used to circumscribe the analysed phenomenon. For instance, Held (2005, p. 193) describes magazine covers as a 'contact-and-advertisement' genre, whose function is to showcase the magazine's contents and to attract the reader's attention — a task which the proposed genre fulfills by using multiple semiotic resources. Tan (2010, p. 93), in turn, uses genre to characterise the multimodal features and the semiotic potential of websites as a form of digital media:

> Newly emergent media such as internet web-pages — an innately hybridic genre that generates a multitude of intertextual possibilities by assembling texts from various modes and discourses (e.g. verbiage, image, sound, activity) that are then represented in multiple relations to one another ...

As the work of Held (2005) and Tan (2010) shows, genre is used to cover a wide range of multimodal phenomena, extending from specific page types within an artefact — a magazine cover — to particular forms of digital media, that is, websites. The question that immediately arises is whether genre can be theoretically useful, if the concept can be freely applied to such diverse examples without constraints and criteria.

Indeed, there are growing concerns that genre has not received the theoretical consideration it warrants. Although the previous work includes both theoretical explorations (Lemke 2001, 2005; van Leeuwen 2005) and methodological proposals

(Baldry and Thibault 2005; Bateman 2008), in many cases the concept of multi-modal genre lacks a solid theoretical foundation due to the scarcity of empirical research and analytical scrutiny. The lack of a definition with firm boundaries is of concern, because the concept is often tasked with abstract description of complex multimodal phenomena — that is, the entire artefact or situation — without the required theoretical foundation.

The purpose of this chapter is to outline the current research on multimodal genre, beginning with the initial linguistic models of genre and then moving towards the state-of-the-art in multimodal research. I emphasise the position of genre in multimodal analysis, because a well-formulated theory requires establishing the relations of genre to other central theoretical concepts such as *medium* and *mode*. In this chapter, the discussion of genre is limited to static multimodal artefacts in print media: I will begin by tracing some of the functions that have been traditionally assigned to genre.

2   The underlying theories in genre research

As several influential multimodal theories draw on both social semiotic and systemic functional theories of language (cf. Halliday 1978; Halliday and Matthiessen 2004), it is necessary to consider how genre has been described in these theories. Firstly, the motivation for deploying genre needs to be established. Bateman (forthcoming, p. 2) identifies three broad purposes:

1. comparing texts and events, and describing their properties,
2. describing the expectations that genre creates for the reader,
3. describing the social functions of genre.

The above points have been extensively investigated in linguistic studies of genre (see e.g. Miller 1984; Ventola 1987; Swales 1990; Bhatia 1993; Hasan 1996; Christie and Martin 1997; Lemke 1999; Martin and Rose 2008). The detailed description of each point is not feasible here due to limitations of space; I simply wish to underline that all of the above points are still relevant for multimodal genre. However, novel approaches are required if we wish to extend genre analysis from linguistics to multimodality.

This brings us to the second point of consideration, that is, the structure of linguistic genre. Traditionally, the structure of genre has been considered *linear* and *staged*. In both spoken and written discourse, genre is seen to unfold through multiple stages, which work towards accomplishing the intended social function. However, the principle of linearity becomes problematic when the analysis is extended from a purely linguistic environment to a multimodal one. It appears

that the linear principle applies mainly when linearity is essential for interpreting multimodal discourse, such as in serial graphics (Holsanova et al. 2009, p. 1223).

There have been propositions to reinstate the principle of linearity by moving away from the artefact structure and towards the observer (Bateman forthcoming, p. 4). For instance, van Leeuwen (2005) proposes that multimodal artefacts act as an environment for the genre. In this environment, the notion of a reading path could be used to trace the hypothetical, linear path taken by the observer to accomplish the social function performed by the genre. However, there is considerable evidence that visual perception is largely task-driven and unlikely to follow a strictly predefined pattern, which contests the notions of linear, observer-based approaches to the study of multimodal genre (cf. Hiippala 2012*b*).

To conclude, it appears that the principle of linearity cannot be used to establish reliable criteria for genre analysis, which would support sequential staging in multimodal artefacts. While the attempts to define the social functions of genre remain largely valid, the absence of linearity or any other principle leaves us with few analytical tools to identify genre structures. Consequentially, genre currently appears as a grey area between the social function of a multimodal artefact and its structure. At the same time, the concept of genre can be useful only when it is defined clearly with sufficient analytical constraints and when its relations to other theoretical concepts are clearly stated.

In order to work towards a definition of genre, the first step is to identify where to look for genre structures. Bateman (forthcoming, p. 3) suggests that genre is a 'multi-stratal' phenomenon, which means that the 'semiotic work' required for deploying a genre involves making choices on several strata. Martin (1999, p. 38-39) has proposed a similar principle for language, which he calls *metaredundancy*. According to Martin (1999, p. 38), "genre is a pattern of register patterns just as register patterns represents patterns of language patterns", that is, linguistic genre metaredounds with register, and register metaredounds with language.

In monomodal contexts, describing the contribution of different strata to genre is by no means a simple task due to various approaches to language (and genre), which have their own theoretical frameworks and concepts, and their interrelations (cf. Martin 1997; Biber and Conrad 2009; Dorgeloh and Wanner 2010). With the move to multimodality, the same principles of theory-building should apply to the definition of theoretical concepts, their interrelations and strata. For this reason, I will now move to describe the concepts relevant to multimodal genre.

3   Towards multimodal genre analysis

If multimodal genre is viewed as a stratified phenomenon, then we need analytical tools that could allow us to identify the semiotic choices that contribute to genre

structures on multiple strata. Moreover, we need to be able to do more than state that the artefacts deploy language and image simultaneously. We also need to say what language and image do in different contexts and how they do it. For this purpose, we need the concept of a semiotic mode to describe the structure and functionality of language, image, or combinations thereof in a given context. Only then it does become possible to proceed towards more abstract levels of description in genre analysis. It should be noted, however, that the concepts of a semiotic mode and that of a semiotic resource are still being developed (see e.g. Kress 2009; Elleström 2010). Due to restricted space and methodological focus, the current discussion focuses on the model proposed in Bateman (2011).

3.1 Rhetorical strategies

Before discussing the semiotic modes, the related concept of *rhetorical strategy* needs to be outlined. According to Bateman (forthcoming, p. 8), a broad definition of rhetorical strategy consists of the 'communicative goal' of an artefact and the contribution of the semiotic modes towards the realisation of this goal. In other words, rhetorical strategies are established means of doing particular kinds of communicative work by deploying and making selections in the available semiotic modes.

Previously, Lemke (1999) has suggested that artefacts with similar social functions may also be structurally similar: the aforementioned notion of semiotic mode allows us to capture and compare these multimodal structures. The structural properties of multimodal artefacts can then be described using a *topology* — a genre space — that describes the range of choices in the semiotic modes which realise a particular rhetorical strategy. Thus, the rhetorical strategies establish a space, which is populated by the genres that do similar communicative work (Bateman 2008, p. 223-225).

However, before we may proceed to discuss the abstract semiotic levels of genre or rhetorical strategy, we need a solid analytical tool to account for the deployed semiotic resources. Therefore, I will now turn to the notion of a semiotic mode.

3.2 Semiotic modes and media

Bateman (2011) proposes that a full-blown semiotic mode needs three strata: (1) a *material substrate*, which carries (2) the *semiotic resources*, whose interpretation is guided by (3) their *discourse semantics*. Firstly, the underlying stratum of a material substrate will be discussed shortly in connection with the notion of medium. Secondly, the semiotic resources are modelled as "paradigmatic systems of choice together a with syntagmatic organisation for re-expressing paradigmatic choices in structural configurations" (Bateman 2011, p. 20). This means that the

semiotic resources allow a range of choices, and the results of these choices may be then combined into structures. Thirdly, the stratum of discourse semantics guides the contextual interpretation of the semiotic resources: it directs the reader towards the correct interpretation in a given context (Bateman 2011, p. 21). I will now highlight aspects of both semiotic resources and discourse semantics using Figure 1.
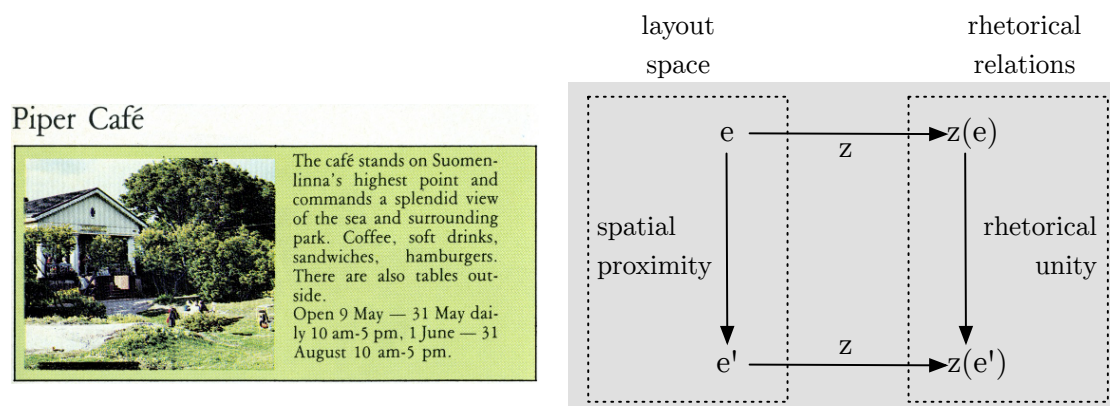

Figure 1: An extract from *Suomenlinna Seafortress in Helsinki* (1988)

The location description in a tourist brochure, shown in Figure 1, is realised using both language and image, and therefore involves choices in both semiotic resources. Previous multimodal research on tourism discourse (see e.g. Hiippala 2007; Kvåle 2010; Francesconi 2011) has described aspects of the choices in lexicogrammar and their relation to the accompanying images. However, what needs to be understood in this context is that the visual-verbal description is a result of choices in both semiotic resources. Consider, for instance, the choice of MOOD in the linguistic structure, where declarative is preferred over alternative choices, such as imperative (Halliday and Matthiessen 2004, p. 23). Similarly, choices in photography, such as camera angle and settings (aperture, exposure, focal length) contribute to the appearance of the photograph (for choices in image and typography, see also Lim 2004; O'Halloran 2008).

The following question is how the instances resulting from these choices are to be interpreted together? In the "back-and-forth" diagram (see Bateman 2011, p. 28-29) in Figure 1, the left-side domain of layout space is mapped with the domain of rhetorical relations on the right-hand side. The mapping relation z indicates a relationship between the spatially proximate entities e and e' in the layout domain and their counterparts $z(e)$ and $z(e')$ in the rhetorical domain. In plain words, the spatial proximity of the photograph (e) and the verbal description (e') indicates that they are to be interpreted together. According to Bateman (2009, p. 61), the

use of the layout space to communicate additional meanings is characteristic of a particular semiotic mode, which he refers to as *page-flow*.

In addition to page-flow, Bateman (2009) identifies two other semiotic modes in print media, which are termed *text-flow* and *static image-flow*. Text-flow is characterised by linear-interrupted text, much like this chapter: diagrams, images, tables and such may occasionally interrupt text-flow, but the two-dimensional space is not used to communicate additional meanings. This distinguishes text-flow from page-flow. Therefore, the discourse semantics of text-flow resemble a much reduced version of those outlined for language in Martin (1992). Static image-flow, in turn, organises graphic elements into meaningful sequences (cf. e.g. Figure 2 below and Hiippala 2012*b*, p. 323). These three semiotic modes are illustrated in Figure 2.
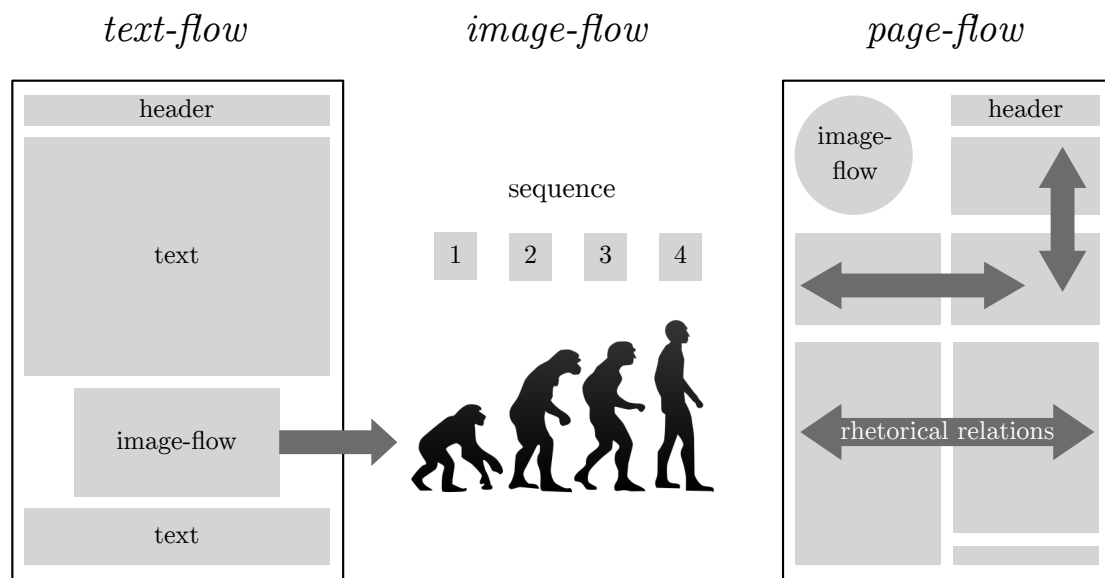

Figure 2: Three semiotic modes: text-flow, image-flow and page-flow

Finally, we arrive at the concept of a *medium* and its relation to the semiotic modes. It was established above that a full-fledged semiotic mode requires a material substrate to carry the semiotic resources. Furthermore, the material substrate needs to be sufficiently controllable so that a group of users may put the material to 'semiotic' work. At the same time, it has to be noted that semiotic modes do not emerge suddenly, but instead they emerge when a suitable substrate comes available and develop over time (cf. O'Halloran 2009). Consider, for instance, the emergence of screen-based displays as a material substrate alongside page-based print media, whose predecessors have included parchment, papyrus and clay tablets. As Bateman (forthcoming, p. 12) writes:

> [A] medium is seen as a historically stabilised site for the deployment of some selection of semiotic modes for the achievement of varied communicative purposes [...] it is possible for an unrestricted range of genres to be carried within [a] medium.

To exemplify, Bateman (forthcoming, p. 12) suggests that books constitute a medium in which several semiotic modes may be deployed. Another proposed medium is the newspaper, which has similar affordances in terms of semiotic modes (cf. Gibson 1979), but significantly different practices of consumption and distribution. Bateman (forthcoming, p. 12) also insists on a clear demarcation between (1) features that arise from the medium, such as page numbering, margins, etc. and (2) selections in the semiotic modes that contribute towards the genre structures in a multimodal artefact, which will be made explicit later in this chapter. This concludes the theoretical discussion of genre and related concepts. In the following section, I will show how the *Genre and Multimodality* model (hereafter GeM; see Bateman 2008) may be used to identify and capture aspects of genre structure in multimodal artefacts.

3.3   Identifying genre structures

The GeM model is a model of multimodal document structure which attends to several aspects of multimodal artefacts: the semiotic resources and their typographic and graphic features, the layout and its hierarchical organisation, rhetorical structure, and navigation. The model also comes with an XML-based annotation scheme for the creation of multimodal corpora. As the name of the GeM model suggests, genre is a foundational notion within the model, whose aim is to provide a consistent analytical method for describing and comparing multimodal artefacts. Because genre is considered a multi-stratal phenomenon, the GeM model provides multiple analytical layers that are cross-referenced in the annotation. A full discussion of each analytical layer is not possible here; the distinct layers and their contribution to the GeM model are listed in Table 1.

The model provides a list of *Recognised Base Units* (RBUs), which can be picked up in subsequent analytical layers (Bateman 2008, p. 111). According to the list of RBUs, sentences, headings, page numbers, list items, photographs, diagrams, etc. are identified as base units. The layout layer describes the hierarchical structure of the identified RBUs, their positioning in the layout, and their typographic or graphic features. The relationships between base units are described using an extension of *Rhetorical Structure Theory* (hereafter RST; see Mann and Thompson 1988; Taboada and Mann 2006). Finally, the navigation structure describes the structures provided to facilitate the use of the artefact.

Table 1: The layers of the GeM model

Layer name	Descriptive function	Analytical unit and examples
Base layer	Provides a list of base units that may be analysed as a part of other layers.	**Base units:** sentences, headings, drawings, figures, photos, captions, list items, etc.
Layout layer	Groups the base units together based on similar properties in the three domains below.	**Layout units**: paragraphs, headings, drawings, figures, photos, captions, list items, etc.
Structure	The hierarchical structure between layout units.	
Area model	The placement of layout units in a layout.	
Realisation	Typographical or visual features of layout units.	
Rhetorical layer	Describes the rhetorical relations holding between the identified rhetorical segments.	**Rhetorical segments**: base units with rhetorical functions.
Navigational layer	Describes the navigational structure by defining pointers, entries and indices.	**Pointers, entries and indices**: base units and layout units with navigation functions.

Source: Hiippala (2012*a*, p. 108)

I will now illustrate how the GeM model may be used to identify multimodal genre structures, using the example shown in Figure 3. The data in Figure 3 were retrieved from a GeM-annotated corpus and visualised automatically using scripts developed for this purpose (Hiippala forthcoming). The upper part of Figure 3 shows the hierarchical layout structure of a double-page in a tourist brochure. The double-page is represented by the parent node in the centre of the diagram. From here, the hierarchical structure extends all the way to the child nodes on the outer circle, which consist of layout units such as headings, paragraphs, photographs, maps and so on (see Table 1). The layout units annotated as graphics are marked using colour. The lower part of Figure 3, which I shall explain shortly below, shows a part of the brochure's rhetorical structure.

As Figure 3 shows, the layout structure of the tourist brochure is fragmented and has a deep hierarchy. In the most extreme case, reaching the outermost child
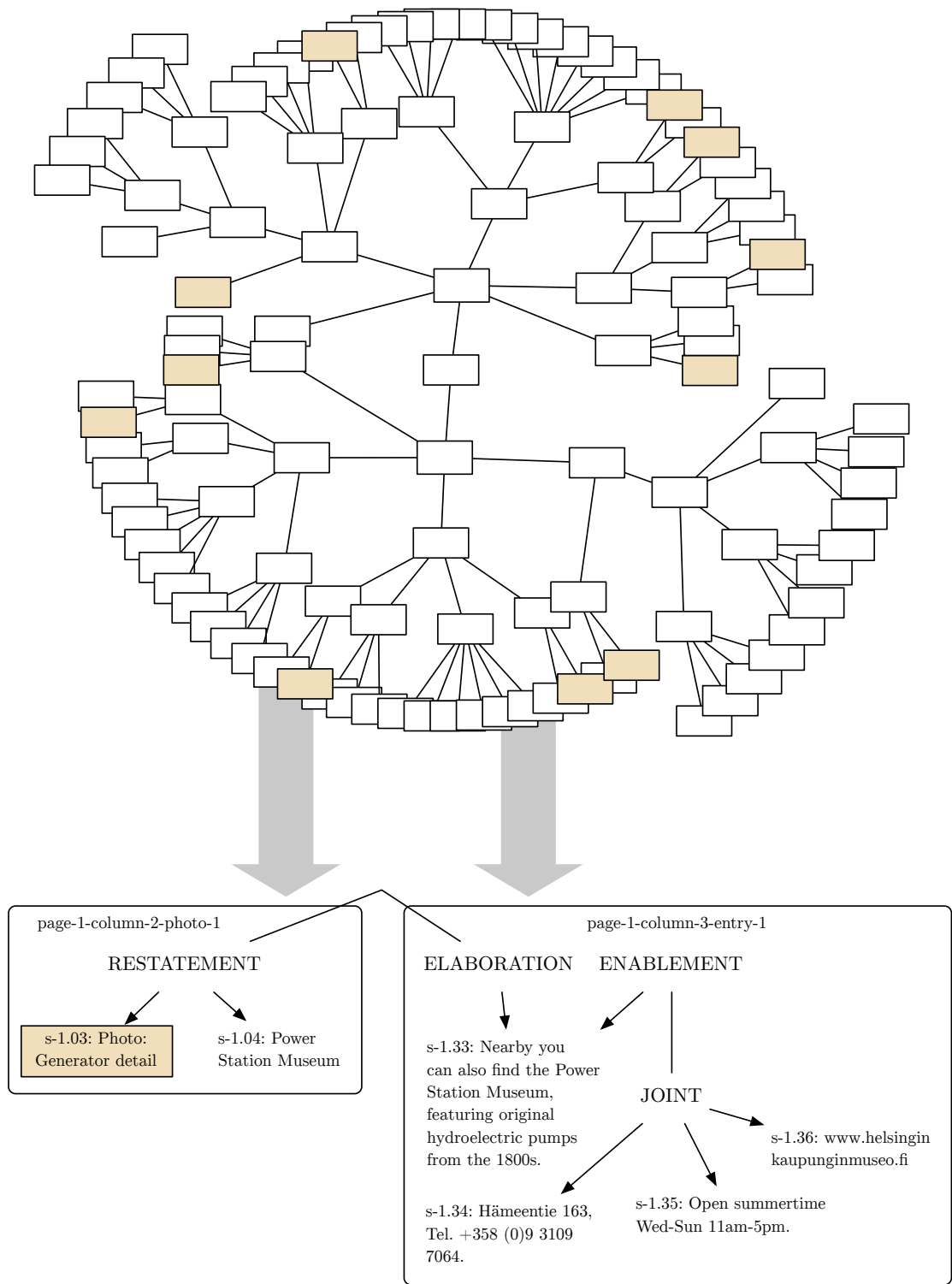
Figure 3: The layout and rhetorical structures in *Helsinki Visitors Guide* (2008) (Hiippala forthcoming, p. 220)

nodes requires a total of five steps from the centre node. Now, considering the layout structure and its functions — organising the content into larger structures and establishing part-whole relationships (Bateman and Schmidt 2012, p. 53) — we may ask what enables the reader to make sense of the complex layout structure in Figure 3? One possible answer may be found in the interaction between the layout structure and the rhetorical structure.

The layout structure is not arbitrary, but functionally motivated. The communicative function of a multimodal artefact defines its structure, which may be captured using the rhetorical layer of the GeM model. This can be seen in the lower part of Figure 3, which shows parts of the rhetorical structure and their relation to the layout structure. In this case, the bounding boxes around the RST structures indicate that the rhetorical segments belong under the same parent node in the layout structure.

On the left-hand side, we have a photograph and an accompanying caption, which participate in a RESTATEMENT relation. Structurally, this configuration of layout and rhetorical structures is close to the example shown in Figure 1, as their structure and spatial proximity in the layout indicate that they are to be interpreted together. Moreover, the photograph and caption elaborate a verbal description located elsewhere in the layout, as indicated by the RST structure on the right-hand side. This RST structure comprises an introductory sentence, which is complemented by travel and contact information using the relation of ENABLEMENT.

What is worth acknowledging here is that these kinds of rhetorical configurations may also be found elsewhere. For instance, the upper part of Figure 3 contains several layout 'chunks' with two child nodes, in which one of the child nodes is a graphic element. In most cases, these image-text combinations in the layout structure possess a similar rhetorical configuration: the same applies to the verbal descriptions which fulfil similar communicative tasks. It may be argued that this kind of patterning allows the reader to negotiate the complex layout structure.

The tourist brochures may prefer certain multimodal structures, while artefacts with different communicative purposes may rely on different types of structure (see e.g. procedural texts in André and Rist 1995; Martinec 2003). These patterned structures have evolved over time to perform certain communicative tasks, and they also respond to the reader's expectations towards the multimodal structure of artefacts (see also Waller et al. 2012). Together, the configuration of these patterns may also signal how the entire artefact is to be interpreted. In the case of Figure 3, the fragmented configuration may encourage an interpretation using the discourse semantics of page-flow, which prefers a selective reading strategy, as opposed to a linear interpretation based on the semiotic mode of text-flow.

Finally, it is also important to note that not all tourist brochures are structured this way in terms of the layout and rhetorical structure. What we have accomplished here is simply a brief description of one possible configuration among many alternatives. And to account for these alternative structures, we need more data, preferably in the form of annotated corpora. For this reason, future work on multimodality and genre is likely to benefit from optical character recognition to speed up the creation of corpora and the development of concordancers for their exploration (see e.g. Thomas 2007, 2009; Parodi 2010).

Further reading

Bateman, John A. (2008) *Multimodality and Genre: A Foundation for the Systematic Analysis of Multimodal Documents.* London: Palgrave Macmillan.
Bateman, John A. (2011) The decomposability of semiotic modes. In K. L. O'Halloran & B. A. Smith (eds) '*Multimodal Studies: Multiple Approaches and Domains*'. London: Routledge.
Bateman, John A. (forthcoming) Genre in the age of multimodality: some conceptual refinements for practical analysis, in P. Evangelisti-Allori, V. K. Bhatia and J. A. Bateman (eds) '*Evolution in Genres: Emergence, Variation, Multimodality*', Peter Lang.

Acknowledgements

References

André, E. and Rist, T. (1995), 'Generating coherent presentations employing textual and visual material', *Artificial Intelligence Review* **9**, 147–165.
Baldry, A. P. and Thibault, P. J. (2005), *Multimodal transcription and text analysis*, Equinox, London.
Bateman, J. A. (2008), *Multimodality and Genre: A Foundation for the Systematic Analysis of Multimodal Documents*, Palgrave Macmillan, London.
Bateman, J. A. (2009), Discourse across semiotic modes, *in* J. Renkema, ed., 'Discourse, of course: An overview of research in discourse studies', Benjamins, Amsterdam/Philadelphia, pp. 55–66.
Bateman, J. A. (2011), The decomposability of semiotic modes, *in* K. L. O'Halloran and B. A. Smith, eds, 'Multimodal Studies: Multiple Approaches and Domains', Routledge, London, pp. 17–38.
Bateman, J. A. (forthcoming), Genre in the age of multimodality: some conceptual refinements for practical analysis, *in* P. Evangelisti-Allori, V. K. Bhatia and J. A. Bateman, eds, 'Evolution

in Genres: Emergence, Variation, Multimodality', Linguistic Insights, Peter Lang, Frankfurt am Main.

Bateman, J. A. and Schmidt, K.-H. (2012), *Multimodal Film Analysis: How Films Mean*, Routledge, London.

Bhatia, V. K. (1993), *Analysing Genre: Language Use in Professional Settings*, Longman, London.

Biber, D. and Conrad, S. (2009), *Register, Genre, and Style*, Cambridge University Press, Cambridge.

Christie, F. and Martin, J. R., eds (1997), *Genre and Institutions: Social Processes in the Workplace and School*, Continuum, London.

Dorgeloh, H. and Wanner, A., eds (2010), *Syntactic Variation and Genre*, de Gruyter Mouton, Berlin & New York.

Elleström, L. (2010), The modalities of media: A model for understanding intermedial relations, *in* L. Elleström, ed., 'Media Borders, Multimodality and Intermediality', Palgrave, London, pp. 11–48.

Francesconi, S. (2011), 'Images and writing in tourist brochures', *Journal of Tourism and Cultural Change* **9**(4), 341–356.

Gibson, J. J. (1979), *The Ecological Approach to Visual Perception*, Houghton Mifflin, Boston.

Halliday, M. A. K. (1978), *Language as a Social Semiotic: The Social Interpretation of Language and Meaning*, Arnold, London.

Halliday, M. A. K. and Matthiessen, C. M. I. M. (2004), *An Introduction to Functional Grammar*, third edn, Arnold, London.

Hasan, R. (1996), *Ways of Saying: Ways of Meaning. Selected Papers of Ruqaiya Hasan*, Cassell, London. Edited by Carmel Cloran, David Butt and Geoffrey Williams.

Held, G. (2005), 'Magazine covers – a multimodal pretext-genre', *Folia Linguistica* **XXXIX**(1-2), 173–196.

Hiippala, T. (2007), Helsinki: A multisemiotic analysis of tourist brochures, Master's thesis, University of Helsinki.

Hiippala, T. (2012*a*), The localisation of advertising print media as a multimodal process, *in* W. L. Bowcher, ed., 'Multimodal Texts from Around the World: Linguistic and Cultural Insights', Palgrave, London, pp. 97–122.

Hiippala, T. (2012*b*), 'Reading paths and visual perception in multimodal research, psychology and brain sciences', *Journal of Pragmatics* **44**(3), 315–327.

Hiippala, T. (forthcoming), Modelling the structure of a multimodal artefact, PhD thesis, Department of Modern Languages, University of Helsinki.

Holsanova, J., Holmberg, N. and Holmqvist, K. (2009), 'Reading information graphics: the role of spatial contiguity and dual attentional guidance', *Applied Cognitive Psychology* **23**, 1215–1226.

Kress, G. (2009), What is mode?, *in* C. Jewitt, ed., 'The Routledge Handbook of Multimodal Analysis', Routledge, London, pp. 54–67.

Kvåle, G. (2010), Invitasjon til destinasjon: Multimodal retorikk i turistkommunikasjon, *in* M. Engebretsen, ed., 'Skrift/bilde/lyd: Analyse av sammensatte tekster', Høyskoleforlaget, Kristiansand, pp. 39–55.

Lemke, J. L. (1999), 'Typology, topology, topography: genre semantics', MS University of Michigan.

Lemke, J. L. (2001), 'Discursive technologies and the social organization of meaning', *Folia Linguistica* **XXXV**(1-2), 79–96.

Lemke, J. L. (2005), 'Multimedia genres and traversals', *Folia Linguistica* **XXXIX**(1-2), 45–56.

13

Lim, F. V. (2004), Developing an integrative multi-semiotic model, *in* K. O'Halloran, ed., 'Multimodal Discourse Analysis: Systemic Functional Perspectives', Continuum, London, pp. 220–246.

Mann, W. C. and Thompson, S. A. (1988), 'Rhetorical Structure Theory: Toward a functional theory of text organization', *Text* **8**(3), 243–281.

Martin, J. R. (1992), *English Text: System and Structure*, Benjamins, Amsterdam.

Martin, J. R. (1997), Analysing genre: functional parameters, *in* F. Christie and J. R. Martin, eds, 'Genre and Institutions: Social Processes in the Workplace and School', Continuum, London, pp. 3–39.

Martin, J. R. (1999), Modelling context: A crooked path of progress in contextual linguistics, *in* M. Ghadessy, ed., 'Text and Context in Functional Linguistics', Vol. 169 of *Amsterdam Studies in the Theory and History of Linguistic Science: Series IV — Current Issues In Linguistic Theory*, Benjamins, Amsterdam, pp. 25–61.

Martin, J. R. and Rose, D. (2008), *Genre Relations: Mapping Culture*, Equinox, London.

Martinec, R. (2003), 'The social semiotics of text and image in Japanese and English software manuals and other procedures', *Social Semiotics* **13**(1), 43–69.

Miller, C. R. (1984), 'Genre as social action', *Quarterly Journal of Speech* **70**, 151–167.

O'Halloran, K. L. (2008), 'Systemic functional-multimodal discourse analysis (SF-MDA): Constructing ideational meaning using language and visual imagery', *Visual Communication* **7**(4), 443–475.

O'Halloran, K. L. (2009), Historical changes in the semiotic landscape: from calculation to computation, *in* C. Jewitt, ed., 'The Routledge Handbook of Multimodal Analysis', Routledge, London, pp. 98–113.

Parodi, G. (2010), 'Research challenges for corpus cross-linguistics and multimodal texts', *Information Design Journal* **18**(1), 69–73.

Swales, J. M. (1990), *Genre analysis: English in Academic and Research Settings*, Cambridge University Press, Cambridge.

Taboada, M. and Mann, W. C. (2006), 'Rhetorical structure theory: looking back and moving ahead', *Discourse Studies* **8**(3), 423–459.

Tan, S. (2010), 'Modelling engagement in a web-based advertising campaign', *Visual Communication* **9**(1), 91–115.

Thomas, M. (2007), Querying multimodal annotation: A concordancer for GeM, *in* 'Proceedings of the Linguistic Annotation Workshop', Association for Computational Linguistics, Prague, Czech Republic, pp. 57–60.
    URL: *http://www.aclweb.org/anthology/W/W07/W07-1510*

Thomas, M. (2009), Localizing pack messages: A framework for corpus-based cross-cultural multimodal analysis, PhD thesis, University of Leeds.

van Leeuwen, T. (2005), Multimodality, genre and design, *in* S. Norris and R. H. Jones, eds, 'Discourse in Action: Introducing Mediated Discourse Analysis', Routledge, London, pp. 73–94.

Ventola, E. (1987), *The Structure of Social Interaction: A Systemic Approach to the Semiotics of Service Encounters*, Pinter, London.

Waller, R. H. W., Delin, J. and Thomas, M. (2012), 'Towards a pattern language approach to document description', *Discours* **10**. DOI: 10.4000/discours.8673.