Modelling Multimodal Genre in Print Media: A Case Study of Tourist Brochures

Tuomo Hiippala
University of Helsinki (Finland)

Abstract
This paper aims to show how the Genre and Multimodality (GeM) model (Bateman 2008) can be applied to the study of multimodal meaning-making and genre as higher level organizational concept in print media. The theoretical framework of the GeM model combines the methodologies of corpus and text linguistics and document design in a multi-layered and cross-referenced XML-based annotation schema, which is suited for the systematical analysis required in modelling the complex phenomenon of genre in print media. An analysis of a series of five tourist brochures published between 1967 and 1988 for the purpose of advertising the city of Helsinki serves as the basis for a discussion of the patterns that may be identified in the rhetorical, layout and navigational structures in this particular print media genre. The semiotic resources of language and image convey ideational meaning for the purpose of construing a persuasive image of the locations, and to guide the reader to and around the location. Consistent patterns in layout and realisation information (typographical and visual features) are used to make the document accessible to the reader. The findings imply that while identifying patterns that may be attributed to the genre of tourist brochure is possible, many of the features can also exist in other print media genres, which suggest that the concept of genre itself is fuzzy and may contain overlapping features. A prototypical model of genre provides a context for more detailed studies of multimodal semiotics, but further investigations and comparative studies are required to in order to work towards a functioning definition genre in print media.

Proceedings of the 10th World Congress of the International Association for Semiotic Studies (IASS/AIS)
Universidade da Coruña (España / Spain), 2012. ISBN: 978-84-9749-522-6 Pp. 1501-1510

Various theoretical and methodological frameworks have been developed for studying the phenomenon of multimodality (see e.g. Baldry and Thibault 2005; Kress and van Leeuwen 2006; O'Halloran 2008) and they have been frequently used to analyse print media artefacts (see e.g. Teo 2004; Maiorani 2009). These multimodal analyses of print media artefacts have focused mainly on language-image interaction and finding corresponding structures and cohesive ties in both semiotic resources. Yet the concept of multimodal genre has received little attention. Bateman (2008), however, has sought to fill this gap by developing a model of genre for multimodal documents (the *Genre and Multimodality* model, hereafter GeM). He has emphasised the need for systematic and empirical anlaysis, drawing on established theories of text and corpus linguistics.

This paper aims to show how the model can be used to observe convergent features in the internal structure of otherwise visually dissimilar multimodal artefacts. The implication is that a systematic approach may also be used to rework the notion of genre in print media.

1. THEORETICAL FRAMEWORK

1.1 From linguistic to multimodal genre

The concept of genre has been widely researched within the field of linguistics in both written and spoken discourse. Traditionally, genre has been viewed as a fuzzy concept responsible for higher-level organisation in various social contexts, serving the function of facilitating social interaction by allowing expectations to be made on the discourse structure. Early research (Swales 1990; Martin 1992) defined the structure of genre as linear and phased, with both required and optional sequential stages, adopting largely typological perspectives into the organisation of genres.

Lemke (1999), in contrast, introduced the concept of genre topology, suggesting that genre patterns (such as linguistic features) may be used as parameters to define similarities and differences between genres: they may share certain parameters, causing the genres to overlap. In more recent work, Lemke (2005) has suggested that a multimodal definition of genre has to be multilinear and more concerned with probabilities of occurrence than sequential structure. This observation is significant, as the co-deployment and interaction of multiple semiotic resources cannot be expected to form an integrative artefact with a single linear structure in terms of genre. Instead, genre in print media is better described by observing the probabilities of occurrence in verbal and visual semiotic resources, and their interaction. The following section describes how the GeM model may be used for this task.

1.2 The *Genre and Multimodality* (GeM) model

In the GeM model, social practices are seen as the context for the production of multimodal artefacts. The established uses of an artefact, its form and the tools used in its production form a 'virtual artefact', which defines the constraints of the physical form, and the production and consumption of the artefact. The multimodal genre, consisting of linguistic, layout, generic and design practices is then realised according to the constraints set by the virtual artefact, as seen in Figure 1, which shows an overview of the GeM model.

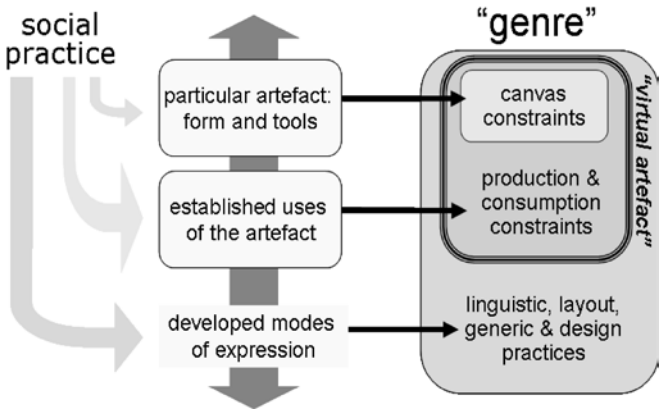

Figure 1: The GeM model (Bateman 2008: 16)

According to Bateman (2008: 13-15), the study of multimodal genre requires a consistent framework for empirical analysis, which allows the observation of multimodal phenomena, upon which further hypotheses may be based. Corpus-based methods are suggested as a solution to provide a framework preferring detailed analyses of data and consequently developed theories over interpretative analysis. In order to describe multimodal artefacts, the GeM model defines four descriptive layers: *base*, *layout*, *rhetorical* and *navigational*, which are now briefly introduced.

The *base layer* is used to distribute verbal and visual elements into base units, which are defined in a list of recognised base units. Examples of recognised base units include, for example, sentences, headings, list items, photos, drawings, and captions. The purpose of the base layer is to provide a comprehensive list of units, which can be subsequently analysed as parts of other layers that describe the actual structure of a multimodal artefact. The base units are grouped together in the *layout layer* to form larger units as they appear in the multimodal artefact: for example, a group of sentences forming a paragraph would be treated as a single layout unit. Layout units are defined according to the principle of visual similarity, based on a shared set of features described in the GeM model annotation. In addition, the layout layer includes a physical representation of the document layout in the form of an area model and a representation of the hierarchical interrelations between layout units, which forms the layout structure. Each of the three domains is cross-referenced with each other to allow the analysis to accurately describe the features and location of each layout unit.

The *rhetorical layer* is used to describe the rhetorical relationships between elements present on a page, based on the *Rhetorical Structure Theory* (hereafter RST), a theory originally developed for describing the rhetorical structure of a linear text (Mann and Thompson 1988). The GeM model extends RST to cover also visual elements and defines an additional set of five «subnuclear» relations, which allow the analysis to extend below the minimal unit of analysis (a sentence), in order to accurately describe text-image relationships (Bateman 2008: 162-163). The rhetorical layer allows the analysis to describe how the content is presented to the reader, allowing a look into the 'internal' structure of an artefact. Finally, the *navigational layer* of the GeM model is used to describe navigational structures present in a multimodal artefact, by defining pointers and entries, which may be realised by both visual and verbal elements.

2. ANALYSIS

2.1 Data

The data set for this study consists of a series of five brochures published between 1967 and 1988 by the city of Helsinki, Finland. The series, named *Helsinki's Four Tourist Islands*, promotes four seaside destinations within the city. The brochures were analysed and compiled into a multimodal corpus. The following sections present the results of the corpus analysis, focusing on the base, layout and rhetorical layers.

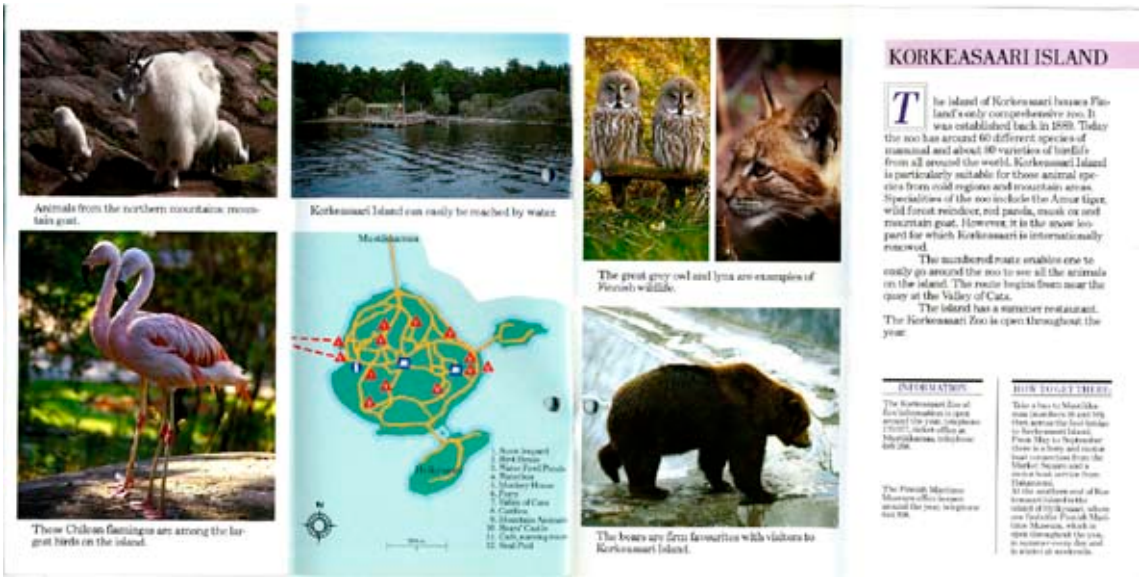

FIGURE 2: A SECTION ADVERTISING THE HELSINKI ZOO IN *HELSINKI'S FOUR TOURIST ISLANDS* (1988)

2.2 The base layer

The analysis of the base layer shows the number of base units growing over time, as shown in Figure 3. However, it should be noted that the brochures published in 1967 and 1972 contain only two double-page spreads, whereas the later brochures have a total of six double-pages.

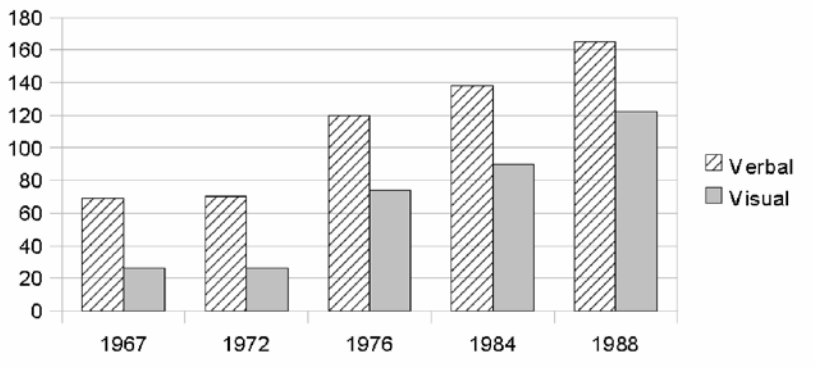

FIGURE 3: THE NUMBER OF BASE UNITS

In the later brochures, the amount of base units increases, although the available semiotic space remains the same. This suggests increased segmentation in multimodal text, which in turn is made possible by complex layouts that break down the layout into sub-areas within sub-areas in the area model, which make the increase in the number of base units possible. Elements with both verbal and visual elements that overlap multiple layout areas are also observed.

2.3 The layout layer

The analysis of the layout layer shows significant variation in the realisation of the various semiotic resources across all five brochures. The variation is reflected in the realisation information of the layout layer, which is used to annotate the typographical features of a text and the general properties of an image. Typical distinctions made in typographical choices of the brochures include using font size, weight and case to distinguish headings and body text, and in some instances, alternating between font families. It appears that the variation is motivated by the function of the particular layout unit: in some cases, such as in headings, alternating features are expected. Within body text segments, typographical variation may be used to highlight information.

In the semiotic resource of image, variation in realisation ranges between illustrations and photography. Both illustrations and photographs can be found in the same brochures, but they typically occur in their own layout areas. In four out of five brochures, photography is used to provide a visual description of the location, whereas illustrations may have a decorative function. In some cases, illustrations contribute to the implicit navigational structure of the brochure by repeating the same illustrations and colours in both maps and header elements. The placement of verbal and verbal elements in the layout shows that grid-based layouts are preferred in the series. In the area model, the verbal and visual elements responsible for carrying the content, which may be identified by consistent patterns in layout, typically occupy their own layout areas. However, layout areas with headers and captions typically allow both verbal and visual elements; in the latter case, this is obvious as the functioning of the caption necessitates the co-occurrence of the caption and image in the same or adjacent layout area.

A reoccurring element combining both text and image is the map, whose major function is the representation of geographical information to the reader. Considering the limited semiotic space available in a small-sized multimodal artefact, the map is the most efficient, if not the only way of representing such information. The representational capability of a map relies heavily on image-text interaction. The realisation of a map may exhibit variation, depending on the function of the map in the brochure, ranging fro m general to detailed representations of space. Points of interest in a map may be designated using lines or numbering to connect symbols and descriptive texts, and a navigational structure may be embedded using colouring and consistent visual/verbal elements.

The GeM analysis in the layout layer can identify typical layout units and compare their realisation in particular occurrences. The descriptive capability of the model is lower than in the interpretative approaches combining the analysis of multiple semiotic resources, resulting from the limited range of properties that can be assigned to each element type. Especially the capability to describe image currently suffers from a lack of a set of features generally agreed upon (Bateman 2008: 121). However, the limitations in the descriptive framework provide a

significant advantage in handing large amounts of data, as the consistent framework enables comparative analysis, which can be used to compare artefacts or to track change over time. With cross-referencing to the rhetorical layer, the GeM model provides a method of looking beyond the surface structure, while enabling cross-layer analysis with the layout layer.

2.4 The rhetorical layer

The rhetorical layer is used to describe the relations between verbal and visual segments, which serve a rhetorical function in the artefact, that is, contribute to the way the content is argued for. Therefore, the purpose of the RST analysis is to describe the 'deep' structure of an artefact. Figure 4 shows the percentages of RST relations in the five brochures. For RST relation definitions and their extension in the GeM model, see Mann and Thompson (1988) and Bateman (2008: 162).

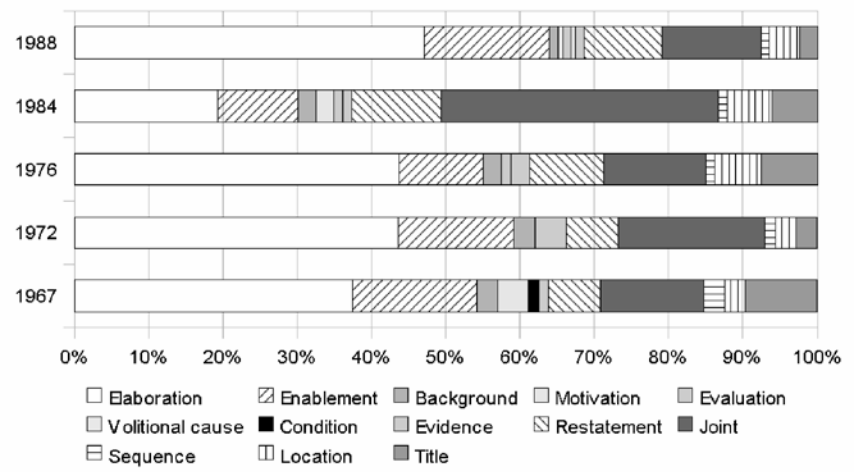

FIGURE 4: RST RELATIONS IN *HELSINKI'S FOUR TOURIST ISLANDS*

Despite the differences in the appearance of the brochure, as described in the analysis of the layout layer, the proportions of RST relations are largely maintained across the five brochures. The 1984 edition is an exception, as it has more JOINT relations than the other brochures; this results from an abundance of similar visual segments of equal value, which are grouped together in the layout layer based on the principle of visual similarity, and thus connected in the RST layer using the JOINT relation.

Common RST relations include ELABORATION and ENABLEMENT, the former being used to describe a location in greater detail, and the latter for enabling the reader to visit the location or to gain more information on it. ELABORATION patterns may include similar patterns within themselves, extending down to a terminal level. Considering the overall function of the artefact, this is not surprising, as the artefact is of *descriptive*, and not of the *argumentative* type. Although relations such as EVALUATION and EVIDENCE can be found, most of the content is presented in a way that the reader is expected to take it as granted. Multinuclear RESTATEMENT and JOINT relations are typically found between visual segments and certain

verbal segments acting as the connecting element between the verbal and visual description of the location. The visual segments are not incorporated in the rhetorical structure as closely as they would be, for example, in an *instructional* artefact, where the visual segments could appear within linear structures consisting of verbal segments.

The main contribution of the RST analysis to the analysis of multimodal artefacts is the capability to look beyond the surface features of an artefact. Interpretative approaches focusing on the particular realisations of the used semiotic resources are unable to describe the deep structure of an artefact, whereas RST may bring new insights into the study of multimodality by looking at the segments, their realisation and the RST relations holding between them. The implications of this to the study of genre in print media are discussed in the following section of the paper.

2.5 Implications

The results of the analysis showed that although the tourist brochures were visually dissimilar, many similarities could be found in their rhetorical structure. In addition, consistent patterns in the layout structures were identified. This implies the existence of higher-level structures that shape the discourse, but at the same time, allow different realisations of the semiotic resources. The question is, how this observation can be used in modelling multimodal genre?

Lemke's (1999) topological view of genre can be used to illustrate the theoretical concept of multimodal genre. In Figure 5, two genres with specific features are shown with overlapping features in artefact B1.

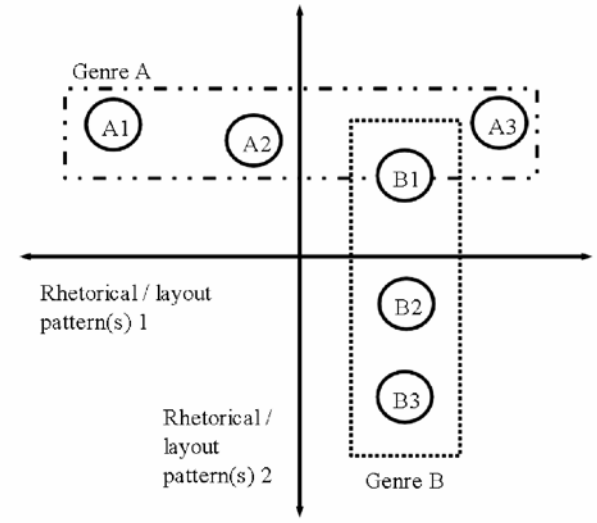

Figure 5: Genre topology, adapted from Lemke (1999)

Let us consider that Genre A is a genre with a descriptive function. In order to define a genre, features typical to the particular genre have to be outlined. In the case of a descriptive genre, these features could include a high probability of RST patterns such as ELABORATION and BACKGROUND. Features in the layout layer may include a high probability of verbal

and visual elements occurring in their own layout areas. If we then consider Genre B to be an instructional genre, we may define, for example, SEQUENCE and ENABLEMENT to be common RST patterns in the genre. Another set of features might be the common occurrence of verbal and visual elements that share the same layout area. An artefact exemplifying Genre A could be a tourist brochure, whereas the artefact designated B1 in the diagram, thus representing Genre B could be, for example, an installation manual for computer software. Although the artefacts serve a different purpose, it is reasonable to expect that certain overlapping features could be found. For example, the artefact B1 from Genre B could have a section describing the product, which would exhibit rhetorical structures similar to those found in Genre A. Similarly in Genre A, sections that guide the reader around a certain location could have similar sequential rhetorical structures as Genre B.

3. CONCLUSION

The GeM model has shown potential as a framework for describing the higher-level organisation in print media, but in order to achieve sufficient empirical backing for a reworked definition of genre, increased comparative research of a variety of artefacts needs to be undertaken. The need for a reworked concept of genre certainly exists within the field of multimodal research, as the current definitions are broad and without empirical foundation. No distinction is often made between a genre and an artefact, and thus print media artefacts are often referred to as genres of their own for the purpose of classification. The traditional functions of genre, classification and categorisation, should be based on empirical research and modelling of artefacts, and complemented by comparative studies.

In another scenario, the definition of genre may be based on the medium, which has lead to a perceived division between 'traditional' and cybergenres (for a brief critique, see Bateman 2008: 209-214). Although the digital environment enables the use of dynamic image and the auditory channel, similarities in linguistic and visual expression can be found across both print and digital media (see e.g. Bateman, Delin and Henschel 2007). Regardless of the medium, increased research on multimodal genre can bring new insights into the whole field of research, complementing the current research focusing on lower level phenomena in multimodal artefacts.

REFERENCES

Baldry, Anthony and Paul Thibault (2005): *Multimodal Transcription and Text Analysis.* Oakville, CT: Equinox.

Bateman, John A. (2008): *Multimodality and Genre: A Foundation for the Systematic Analysis of Multimodal Documents.* London: Palgrave.

Bateman, John A., Judy Delin and Renate Henschel (2007): «Mapping the multimodal genres of traditional and electronic newspapers», in Terry Royce and Wendy Bowcher (eds) *New Directions in the Analysis of Multimodal Discourse*. New Jersey: Lawrence Erlbaum, 147-172.

Kress, Günther and Theo van Leeuwen (2006): *Reading Images: The Grammar of Visual Design*. 2nd edition. London: Taylor & Francis.

Lemke, Jay L. (1999): «Typology, topology, topography: genre semantics». http://www-personal.umich.edu/~jaylemke/papers/Genre-topology-revised.htm

— (2005): «Multimedia genre and traversals», *Folia Linguistica* XXXIX (1-2): 45-56.

Maiorani, Arianna (2009): «Movies 'reloaded' into commercial reality: representational structures in The Matrix trilogy promotional posters», in Eija Ventola and Carys Jones (eds) (2009) *From Language to Multimodality: New Developments in the Study of Ideational Meaning.* London: Equinox, 275-298.

Mann, William C. and Sandra A. Thompson (1988): «Rhetorical Structure Theory: Toward a functional theory of text organisation», *Text,* 8 (3): 243-281.

Martin, James R. (1992): *English Text: System and Structure.* Philadelphia, PA: John Benjamins.

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

Swales, John (1990): *Genre Analysis: English in Academic and Research Setting.* Cambridge: Cambridge University Press.

Teo, Peter (2004): «Ideological dissonances in Singapore's national campaign posters: a semiotic deconstruction», *Visual Communication,* 3 (2): 189-212.