

Genre and Layout in Multimodal Documents: Towards an Empirical Account

Patrick Allen, John Bateman, and Judy Delin

University of East London, University of Bremen, University of Stirling
allen2@uel.ac.uk; bateman@uni-bremen.de; j.l.delin@stir.ac.uk

Abstract

We argue for a model of document genre that encompasses both linguistic and graphical resources as a means of expression. Two bird books are analysed (1972 and 1996) and a simple framework presented for the description of graphical differentiation of communicative function. The analysis predicts that graphical differentiation in this genre is less in earlier publications. This suggestion is tested on a third bird book (1924) and found to be correct. However, the comparison shows that the genre itself has shifted over time, arguing for a 3-D model in which genre is represented as a shifting and interrelated set of parameters. Given this, analysis across genres and within genres across time is a very similar enterprise. A new project, GeM, has begun research into the appropriate model of genre for representing the generic differentiation of layout.

Introduction

As the use of electronic documents increases, and the design identity of both these and more traditional paper documents is negotiated, we need a theory of how far general functional design principles apply to multimodal documents and how far genre-specific principles are needed to account for layout. While it is widely acknowledged that layout is a crucial component of the printed message, very little solid knowledge is available, outside the practical expertise of layout professionals, as to what motivates layout as purposive communication in context. This paper reports the initial stages and theoretical orientation of the GeM project, an investigation into the relationships between genre and all aspects of layout.

We briefly summarise the salient aspects of previous results on the computational modelling of layout constraints and introduce our application of the notion of 'genre' and its motivation. We then give a brief example of the kind of empirical work we are doing, using bird books as an example of a genre, and show the relationships between genre and layout decisions that it illuminates. We conclude with a brief summary of our research orientation.

Previous Results in Motivating Layout

Several significant projects have studied multimodal information presentation in a computational context, examining how information can be distributed coherently across media. Most systems also address the question of 'media allocation' (cf. Arens and Hovy, 1990). Relevant work in multimodal document design includes WIP (Wahlster *et al.*, 1991, 1993), COMET (Feiner and McKeown, 1990), Mittal *et al.* (1995), and Kerpedjiev *et al.* (1997). All these approaches have in common the recognition that graphical design is as purposive as verbal activity: thus, approaches to language generation based on communicative goals could be extended to include graphical presentation. Information presentation of this kind is typically based on the construction of a 'rhetorical structure' for the information to be expressed: individual leaves of this structure can then be expressed in running text or as diagrams as appropriate. This research direction therefore usefully complements the approaches of diagram design, in that communicative goals necessarily augment more data-oriented data classification.

Early considerations of layout were confined to a single medium – i.e. a text. Thus, Sefton (1990) and Hovy and Arens (1991) use rhetorical text organization to motivate a variety of text formatting decisions: for example, a rhetorical sequence can be expressed by a bullet list, and emphasis can require bold face, etc.¹ When considering mixed modes, considerations of layout going beyond textual formatting devices were required. Feiner (1988) proposed a rule-based approach to organizing multimedia content, and subsequently Graf (1992, 1995) adopted a constraint-based approach. Both base layout decisions on the textual-rhetorical organization of the information to be presented: rhetorical relations are mapped directly on to graphical properties of the presented layout. Reichenberger *et al.* (1995) investigated the interrelationships between layout, written content, graphical content, and rhetorical organization in one of the most sophisticated attempts to date: it developed heuristics for proposing pos-

¹ Similar work is reported by Marks (1991) and Roth and Mattis (1991).

sible layouts based on an RST analysis of rhetorical and textual content (cf. Mann and Thompson 1987). This established that there are substantive links between notions of 'communicative purpose' and 'communicative goals' and aspects of text layout which motivate even quite complex layouts.

Although the Reichenberger *et al* model demonstrated both the value of rhetorical analysis and the utility of generating alternative layouts for evaluation, it was found that too many possible layouts still resulted from even the most sophisticated characterisation of the rhetorical relationships holding between the information elements to be conveyed. In addition, the project did not attempt to model independently-known constraints on layout *per se* (i.e., constraints from document type or mode of production), the constraints of generating alternative genres (only one genre was modelled), or results from readers' and experts' reactions to the generated pages. In short, existing models are underconstrained in the kinds of layout that can be produced or motivated by rhetorical structure.

GeM: A Proposed Solution

We begin from the premise that notions of communicative function as we know them are not sufficiently well articulated to motivate layout decisions. It seems clear to us that communicatively-based layout description needs to look beyond RST-style characterisations of textual content, bearing in mind that we need not only to capture 'conventional' layouts, but the layout of publications such as the new magazines whose layout seems as much based on shock value and apparent randomness as on systematic motivated communication.

In addition, it is clear that stereotypes and ordinary expectations of what documents should look like play a crucial role in layout decisions. So far, however, the connection between layout and these considerations of genre has only occasionally been made: for example, Ames (1989) makes reference to a 'special design strategy' necessary for news pages, noting that this must be based on news judgement, while Kress and van Leeuwen (1998) make the link between genre and layout implicitly in their description of newspaper front pages. Design professionals' decisions also need to be based on practical concerns, such as the existence of time constraints (the availability of various component items such as text, advertisements and photographs at deadline, for example, as in newspaper production: cf. Ames (1989:39), Negrine (1994:123)). There has not, however, been a systematic exploration of how, or to what extent, genre is capable of highlighting contrasting layout principles between document types and of including practical design constraints.

The GeM approach incorporates these practical issues in an extension of the linguistic notion of 'genre', applying this to pages involving multimodal elements. Just as genre

constrains linguistic realizations, we argue that it may also, suitably extended, cover layout, and hence provide additional layout constraints. We take two properties of genre as criterial: (i) a genre represents a socially-established set of communicative purposes and accompanying communicative events and (ii) these events are expressed in standardized, prototypical ways that are strongly supportive of their recognition and are formative for their production. A genre thus gives rise to 'generic' structures that reliably recur in instances of communication belonging to that genre.

In the few examples of the notion being employed in layout generation, genres are modelled as separate entities. We take the view that a successful model needs to expect that genres will continually transform and mutate towards and away from one another. In GeM, therefore, we plan to treat genres not as discrete sets of constraints but as representative of a point in a multidimensional genre space. This will allow convergences and divergences between genres to be represented and generated, and will support the exploration of any potential document genre within the space. Description in terms of a genre space allows us to position not-yet-existing layout genres.

Our method is essentially empirical, based on a corpus of professionally produced layouts spanning several distinct types of documents and ranges of physical modes of production (including manuals and instructions, Web pages, illustrated books, and newspapers). An annotation method will be developed as part of the project, allowing us to distinguish texts with respect to content, layout, communicative organization, and linguistic details of any textual elements. The regularities found between any of these descriptions and the document type and mode of production will be embodied in a functioning computational prototype for generating diverse layouts given a specification of intended genre.

As an illustration of the general approach, we take an example in which the textual genres are identical, but a comparison can be made across time to examine development and change. It is upon this type of analysis that the corpus annotation system will be based.

A Diachronic Example: Bird Books

As a discussion example, we take two pages from books for identifying birds (Figure 1)². Bird books represent a well-known example of illustrated documents, are susceptible to variation across time, which sharpens perceptions of the state of the genre in the present day, and

² Production constraints have meant that the text on reproduced pages may not be legible. The layout, however, should still be sufficiently clear. We re-quote text where necessary.

A.1

Gannet

Family SULIDAE. Gannets

Sula bassana

No. 27



This great bird has the magnificent wings and flight of a giant Gull, and a wing-span of nearly 6 ft. It can sight a fish from a great height while on the wing and will drop like an arrow into the water after its prey. The plumage is white with a tinge of buff on the head and neck, and dark brown, almost black, wing-tips. Immatures are first dusky all over, later piebald or white sprinkled with dark spots.

Haunt The coast and sea, and at breeding time rocky isles and stacks, chiefly on the north and west of the British Isles.

Nest Of seaweed and tufts of grass or thift; on the rocky ledge of a stack or island in a great colony with others.

Eggs 1, nearly white, chalky. April or May.

Food Fish.

Notes Short and harsh.

Length 36 or 37 in (91 or 94 cm)

Status Resident

22

A.2

24

GANNET AND IMMATURE

17-200 cm

Gannet

Sula bassana

ID FACT FILE

Size: Larger than

Gull

Wings: White

Black wingtips

Yellow legs

Immature: Grey

Slightly yellow

Legs: Yellow

Eyes: Yellow

Bill: Yellow

Feet: Yellow

Voice: Harsh

Song: Harsh

Call: Harsh

Flight: Harsh

Habitat: Coastal

Diet: Fish

Breeds: April

Lays: 1

Incubation: 40 days

Lifespan: 10-15 years

Status: Resident

birds of the gannet family. Gannets are white with black tips off the wings of Europe. They travel very fast but after nesting, they travel slowly. The young migrate south in the winter. Gannets feed on fish by plunging from 2000 and 3000 ft. into the water. The nest is a pile of seaweed. A single egg is incubated for 40 days. The young bird is fed by fish brought in by the parent birds.



Adults in breeding colors.



How the bird flies.

Figure 1: Two pages illustrating varying layout styles across time

have a high degree of page-to-page similarity within the same publication. This gives another dimension of genre variation: presumption of a history over which the reader can learn the design decisions made. Example A1 is taken from the *Observer Book of Birds* (Warne, 1972), and A2 from the book *Birds of Britain and Ireland* (Harper Collins, 1996). The comparison we wish to make is a diachronic one within a similar genre, but it is clear that the parameters and methods described are equally appropriate for cross-genre comparisons of documents. Rhetorically, the information on each page is a series of 'joint' relationships (to use the terms of RST). This relationship does not presuppose any particular ordering of information presentation, and therefore does not do much work in constraining the layout of the information on the page. The books therefore serve our purpose of investigating further sources of constraint very well.

Text A1 is homogeneous in presentation, and supports a reading of the text in a single block. A1 provides most of the information textually, dividing information into head, running text, and a data-list describing Haunt, Nest, Eggs,

etc. The data-list is not visually highly contrastive from the main text (i.e. it is not distinguished by spacing, location, font family; it is only in fact recognisable by bold font labels and short lines). The text, picture, and title of A1 are clearly identified and united with one another by justification: the text block is both right- and left-justified, as is the title, and no line spaces appear between them and the picture. The picture is therefore completely framed by text, from which it is only slightly indented. What we see from A1, therefore, is a low degree of diversification and contrast, with uniformity provided by font family, physical proximity, and the framing of the whole single block with white space.

A2, in contrast, is made of many more visual 'blocks' than A1, and supports *multiple entry points* for the reader into the page. A2 differentiates its information much more distinctly through appearance and placement, and uses graphics in a much more sophisticated way to both link and dissociate elements. The page is organised into two columns, the title ('Gannet') indicating the primary entry-point for reading to be the wider second column. The

main text block is both linked with and dissociated from the photograph, because, although the body text and photograph are aligned by justification on the left, the text is not right-justified, and is not aligned with the right-hand margin of the photograph. The picture itself runs off into the margin, 'puncturing' the white border of the page as a whole. In addition to the title, main text block, and photograph, there is also a data-list (ID FACT FILE), a table, map, labels and captions, and outline and detailed graphics of the birds. Column 2 is devoted to 'background' information, and contains five elements, each clearly differentiated: a running head and page number, title, body text, photo of 'adults at breeding colony', and colour graphic of 'three immature stages' and 'adult' bird. Other graphical devices are clearly framed by white space, or, in the case of the map and table, by a border. While A1 featured a single colour element (the picture), in A2 there are five: a bird outline with spread wings at the top of Col 1 is blue, below it, a table with the months of the year the bird is present in the UK in green, a map highlighting areas where the birds are found in magenta, and a photograph

and bird pictures below it in full colour. Colour therefore provides an attractant to the eye, but not a cohesive link between these elements, which serve to link page-to-page to facilitate comparison of entries within the book.

Also evident from the comparison of A1 and A2 is the *redistribution* of the information across modalities and realisations. For example, the plumage description of A1's main text block occurs distributed across the lower graphic and one of the description items in the ID fact file' in A2; conversely, A1's data-list item 'Nest' occurs only in the main text block of A2 as a description of the nest: 'The nest is a pile of seaweed'. Such differences make a significant contribution to the page designs but, crucially, are by no means sufficient to explain the full diversity exhibited. That is, whereas part of the difference has been deliberately designed by content redistribution, there has also been a differential response to the *functional requirements* that the texts are expected to fulfil. This link between function and resulting form, including layout, is the essential indication of the potential role of genre.

| Feature Type | Feature | A1 rendering | A2 rendering |
|-------------------------------|---------------------------|--------------|--------------|
| Direct Observation | size:length | DL | I |
| | appearance:adult | T,P | P,DL,G |
| | appearance:sound | DL | DL |
| | size:comparison | T | DL |
| | appearance:juvenile | T | DL, G |
| | appearance:bill | (P) | DL |
| | appearance:when flying | (T) | DL |
| | lookalikes | T | DL |
| Location, Habitat | distribution:type | DL | T, P |
| | distribution:geographical | DL | M,T |
| | seasonal distribution | DL | IT |
| Background Information | nesting | DL | T |
| | how it feeds | T | T |
| | eggs | DL | - |
| | what it eats | DL, (T) | T |
| | eggs:when, incubation | - | T |
| | young:feeding, duration | - | T |

Figure 2: Two Pages compared by Rendering of Information Features

Key: H header IT icon using table P picture G graphic
DL data list I icon using quantity M map T text block

A comparative analysis of the two texts clearly shows the level of functional differentiation that has been achieved through the use of layout devices. Figures 2 and 3 represent the three main kinds of information expressed by the bird books: information available from directly observing the bird, about habitat and where the bird is found, and other background information. In figure 2, we can see first of all the differences in spread of information between the two texts: A1 does not represent appearance of bill, or appearance while flying, other 'lookalikes', and some details of eggs and young. A2, on the other hand, omits information about the appearance of eggs, but overall is the more informative. Second, we get our first impression of the degree of information redundancy in A2 as compared to A1: adult appearance of the bird, for example, is communicated in A2 through three graphically-distinct methods (a picture, data-list, and graphic) while A1 has two (text, data-list). Similarly, where A2 has two methods of communicating geographical distribution (map, text), A1 has one (data-list). In all but one case (what the bird eats), where more than one graphical resource is used for the same information type, A2 uses the greater diversity of means of expression. Readers can therefore find the same information from a range of sources on the page, which increases the usability of the document. Given that, as we noted earlier, A2 supports multiple entry points, the lack of guarantee that a reader will traverse the page in any particular order would make a strategy of providing distributed and redundant information a sensible one to ensure usability.

Figure 3 presents the information in a different way, highlighting how choice of graphical realisation is influenced by the kind of information being expressed. In A1, the older text, there is a high degree of crossover between graphical resources and information type: for example, the data-list is used for information of all three types. A2, on the other hand, uses a data-list for direct-observation information only. The layout resources available in A2 provide for considerably more flexibility, motivated by type of information, thus providing useful additional meta-organization for the reader. We can see, therefore, that the development of a more graphically varied approach to page design has gone hand-in-hand here with the development of genre whose function is to present multiply-faceted information which might require access in a variety of ways.

We can also detect in the progression over time from A1 to A2 a shift of information realisation across modalities. While we have described the kind of information that the older text, A1, does *not* distinguish through layout, we do find that some of this functional discrimination is performed *grammatically* by A1. For example, there are regular but diverse circumstantial prepositional phrases under the DL-item 'nest': 'Of seaweed and tufts of grass or thrift; on the rocky ledge of a stack or island in a great colony with others.'

| Pg. | Background | Location, Habitat | Direct Observation |
|-----|------------|-------------------|--------------------|
| A1 | DL, T | DL | DL, T, P |
| A2 | T, H, P | M, IT, T, P | DL, G, P, I |

Figure 3: Types of Information and Their Rendering

There is punctuation division under the item 'eggs' where the first 'sentence' describes the number and colour of the eggs (but as a modified nominal phrase) and the second 'sentence' describes the months in which the eggs appear (but as the names of months): '1, nearly white, chalky. April or May.' While a large corpus is clearly needed to enable conclusive statements to be made about the development of information presentation in the bird book genre, the framework adopt present is clearly productive in highlighting trends for future analysis.

Towards Cross-Generic Comparison

The kind of analysis we have illustrated so far interrogates further the truly *generic* aspects of the description, setting out those properties that texts share with others. The two (sub-) genres of bird descriptions illustrated by A1 and A2 exist in an overall space of genres within which such particular cases can be placed and compared systemically with genres located differently. We can ask whether the graphic renderings express the rhetorical relations intuitively and consistently; whether distinct information types are rendered by correspondingly distinct graphical or layout resources, and whether there are already existing genres with similar functions which can be applied, thus providing a 'history of interpretation' for the user within which a richer layering of information presentation is possible. In order to test the approach, we will look at one more text: the Gannet entry in the 1924 bird book *British Birds: Description of All but the Rarest Species, their Nests and Eggs* (London & Edinburgh: Jack and Nelson).

This entry is reproduced in figure 4. This text, which we refer to as A0, is strikingly different from both A1 and A2. Most obviously, bird entries in this 1924 book are not given a separate page, so the distinction between entries relies on the title (bold, indented, and the same point size as body text). Neither are the pictures closely aligned with the entries, or in a consistent position in each entry: see, for example, the intrusion of the Shag picture caption at the right-hand top of the page into the space used for describing Gannets, and the resulting necessity to give the page number as well as the figure number in the Shag caption to avoid confusion with the Gannet. Reference between entries is not as polished as in A2, for example, where 'lookalikes' are listed in the same place on every page, and the reference to appearance makes it clear what the purpose of the comparison is. Here, '(See No. 177)'

comes after the description of the Gannet's size, but it can only be presumed that this reference is meant to enable size comparison between the Gannet and entry 177, the cormorant (the latter is three feet in length, so it is hard to see the purpose of the comparison; neither are they easily confused, as the Cormorant is black and the Gannet is white). Apart from the positioning of pictures, there is no informative caption apart from figure number, and the graphics are simple black and white line art. Since proximity of bird illustration to its entry is a problem, the only connection between pictures and text that can be relied upon is the link between the figure number in the caption and the reference to it in the text.

A.0

CORMORANTS—GANNETS. 115

Nest. Usually on ledges in sea-caves, sometimes on a cliff ledge or in recesses among boulders by the sea; generally, therefore, unlike the cormorant's, under cover. *Material:* sea-weed and coarse herbage. The species breeds in colonies.

Eggs. Usually 3-5. Like cormorant's, but smaller. Av. size, 2.51 x 1.61 in. Laying begins March-April. Two broods.



(2) *Family: Sulida—Ganneta.*

179. Gannet or Solan-goose [*Sula bassana* (Linnaeus)]. Resident. Marine species. Fig. 104 (1178). Breeding places: Graaholm (Wales); Bull Rock and Little Skellig (Ireland); Bass Rock, Ailsa Craig, St. Kilda, E. Sullageir and N. Barray (Scotland), Faeroes.

Bird. Length 33 in. (See No. 177.) Bill strong, straight, with hook, and of a pale lead-blue with deeper slate-coloured longitudinal lines. Tail graduated to a point, the central feathers the longest. Wings long and rather narrow. Plumage white, save the wing quills, which are dark brown, not black as usually stated. In breeding dress there is a buff tinge on the head and neck. The fledgling blackish-slate, spotted white. Immature till the third or fourth year, and recognised by the dark brown, chiefly on the back, wings, and tail, which colour diminishes season by season till at maturity reduced to the brown of the wing quills.

Nest. Generally on the ledges of the precipitous sides of sea-washed isles, such as the Bass Rock, Ailsa Craig (Scotland), Bull Rock and Little Skellig (Ireland). *Material:* sea-weed, grasses from the isle top, and any material from the surface of the sea.



Fig. 105.

Egg. One. Blue, covered or nearly so with a white chalky deposit which soon becomes dirty and yellow stained. Av. size, 3.06 x 1.96 in. Laying begins March-May. One brood.

(2.153) 8

Figure 4: 1924 Gannet Entry

However, this is not to say that the entry is not strictly organised: there is a strong generic structure for each. First, there is the bird's name, specialised family name (English and Latin), Resident Status, and where it is found (given as a list). Minimal structuring information is given in the form of the graphically undifferentiated subheading 'Breeding Places'. The list that then follows is much more detailed than that in A1 and A2: in A2, the information about breeding is covered less specifically by a map. The

rest of the text, while it *looks* like continuous prose on the page, it is in fact composed entirely of a description list. After the list of the locations in which the bird is found, the entry goes on to describe the bird itself:

Bird: Length 33in. (see no. 177.) Bill strong, straight, with hook, and of a pale lead-blue with deeper slate-coloured longitudinal lines...

Nest. Generally on the ledges of precipitous sides of sea-washed isles, such as the Bass Rock (Scotland)...

Egg. One. Blue, covered or nearly so with a white chalky deposit which soon becomes dirty and yellow stained. Av. size, 3.06 x 1.96 in.

In the whole of the text, there are only two syntactically complete sentences: 'In breeding dress there is a buff tinge on the head and neck' and 'Laying begins March-May'. The textual appearance, then, barely reinforces the fact that this is not continuous text. Apart from the italicised and indented list labels *Bird*, *Nest*, and *Egg*, and the horizontal space between these labels and the beginning of their associated text, it is only the content of the list (finite ellipsed simple present tense property ascriptions) that indicates the list status of the information. The dominant organization of the page, then, is a series of slots for different sorts of information, one or two of which are given head or subhead status, with the remainder differentiated positionally and by content. This represents a very low level of graphical differentiation per function. Note, too, that there is no differentiation in fonts or sizes: only italic, bold, and plain text within the same font family are used. Punctuation does not mark rhetorical relations consistently, as the labels *Bird*, *Nest* and *Egg* (arguably introducing informative elaborations) are marked with a full stop, rather than the colon that might now be expected (cf. A2's list entries in ID FACT FILE), but the same rhetorical relation, the elaboration of nest material within the Nest entry, is introduced here with a colon (hyphenation is deliberately reproduced):

...Bull Rock and Little Skellig (Ireland). *Material:* sea-weed, grasses from the isle top, and any material from the surface of the sea.

The purpose, here, is not to use contemporary expectations about layout to criticise texts like A0, but simply to point to the far lower degree of graphical differentiation of functions than is present, even in A1. What is striking about A0, however, is that it provides a far greater detail of information than the newer texts. Some of this is culturally explicable: we have seen that 'eggs' disappears as a relevant piece of information in the 1996 text, as current thinking is that enthusiasts should be discouraged from going anywhere near the eggs of wild birds. Egg information is present in some detail in A1, and in even more

detail in A0. A0 is also much more specific about bird locations, plumage description, and nesting material than either A1 or A2.

The prediction that could be made from the analysis of A1 and A2 seems to be correct, therefore, in that, the further one goes back in time within the genre, the less graphical differentiation of informative function is used. But further questions arise about the nature of genre itself. The amount of detail in A0 raises the issue of whether A0 is simply targeted at a more knowledgeable audience, or whether modern bird books are 'dumbing down', and simply include less information. It is tempting to argue that the usability of the information has increased, but this may be merely a contemporary perspective based perhaps on the expectation of the 'multiple entry point' navigation style rather than a style based on expectations of serial reading. The discussion simply highlights the fact that historical, cultural, and cross-generic comparisons are all important planks of a theory of document production, document consumption, and genre description.

In fact, A0 has more in common with a contemporary encyclopedia entry than with texts like A2. What has taken place is that the genre *itself* has shifted its definition, along with the kinds of uses that people expect to make of such texts. While we may not now be happy to accept text A0 as a usable field guide (it looks more like a 'reading' book than a quick field reference, to the contemporary eye), it is clear from the foreword of the 1924 book in which A0 appears that the author's intention was for users to recognise birds, rather than to use the book as a reading resource:

THE object of this book is to make as easy as limits of space permit the recognition of all the British species...The descriptions have been written with an eye to the requirements of the non-specialist observer...they should prove adequate for practical purposes.

This, then, sounds like a field guide, although our more graphically-oriented contemporary expectations would lead us into difficulties differentiating important information sufficiently quickly. What we would want to conclude would be that we cannot simply compare texts on the basis of the assumption that they constitute a single time-extended genre: there are other variables at work. Genres must be described independently of the particular use that a culture makes of them. Genres do not merely reflect convention: each instance of a particular genre *creates* convention and hence generic expectation. Our disbelief that A0 could be anything other than a reference book reflects how far the genre has moved since 1924. A sophisticated model of genre must therefore reflect genres not as static, but as mutable.

Conclusions and Ongoing Work

This brief discussion has focused on the value of examining layout as one of the expressive resources available in written and printed communication. We have illustrated our arguments about genre through the diachronic description of three pages with similar functions. The application to comparisons *across* genres is clear, since one need only go back in time to find that the notion that genres are predictably bounded is itself illusory.

Our research project, GeM (Genre and Multimodality)¹, will over the next two years allow us to focus on the role of genre in layout, based on an layout annotation developed for a corpus of illustrated texts of different types. We propose to explore how far genre can capture the differences between multimodal document types, and what generalisations can be made that go across types; the space of variation within document genres; and the role of practical constraints such as word-limits and time limits in the choices made in everyday document design. Given that the project proposes a wide set of sampling points in the space of possible genres through the different text types that will be examined, the research promises an assessment of the nature and number of modifications and extensions to existing frameworks for describing textual and visual space.

References

- Ames, S. *Elements of Newspaper Design*. 1989. New York: Praeger.
- Arens, Y. and Hovy, E. 1990. How to describe what? towards a theory of modality utilization. In *The Twelfth Annual Conference of the Cognitive Science Society*, pages 487--494. Hillsdale, N.J.:Lawrence Erlbaum Associates.
- Bateman, J. A. and Teich, E. 1995. Selective information presentation in an integrated publication system: an application of genre-driven text generation. *Information Processing and Management: an international journal*, 31(5):753--768, September.
- Bhatia, V. K. 1993. *Analysing genre: language use in professional settings*. Harlow, U.K.: Longman.
- Catford, D. 1965. *A linguistic theory of translation*. Oxford: Oxford University Press.
- Christie, F. and Martin, J. R. 1997. *Genre and institutions: social processes in the workplace and school*. London: Cassell.
- Danes, F. 1974. Functional sentence perspective and the organisation of the text. In Danes, F. ed. *Papers on Functional Sentence Perspective*. 106--128. Prague: Academia.
- Feiner, S. 1988. A grid-based approach to automating display layout. In *Proceedings of the Graphics Interface*,

¹ The authors gratefully acknowledge the support of ESRC for the GeM Project (R000238063).

- pages 192--197, Los Angeles, CA.: Morgan Kaufman.
- Feiner, S. and McKeown, K. R. 1990. Coordinating text and graphics in explanation generation. In *AAAI-90: Proceedings of the 8th National Conference on Artificial Intelligence*, volume I. 442 -- 449. Menlo Park, CA.: AAAI Press and The MIT Press.
- Gläser, R. 1990. *Fachtextsorten im Englischen*. Tübingen: Gunter Narr Verlag.
- Graf, W. 1992. Constraint-based layout of multimodal presentations. Technical Report RR-92-15, DFKI, Germany.
- Graf, W. 1995. The constraint-based layout framework laylab and its applications. In *Proceedings of ACM Workshop on Effective Abstractions in Multimedia, Layout and Interaction*, San Francisco, CA.: ACM.
- Hasan, R. The nursery tale as a genre. In C. Cloran, D. Butt, and G. Williams, editors, *Ways of saying, ways of meaning: selected papers of Ruqaiya Hasan*, pages 51--72. Cassell, London.
- Hatim B. and Mason, I 1990. *Discourse and the translator*. London: Longman.
- Hovy, E. and Arens, Y. 1991. Automatic generation of formatted text. In *Proceedings of the 8th. Conference of the American Association for Artificial Intelligence*. 92—96. Anaheim, CA.: AAAI.
- Kerpedjiev, S., Carenini, G., Roth, S., and Moore, J 1997. Integrating planning and task-based design for multimedia presentation. In *Proceedings of IUI-97*.
- Kittredge, R. and Lehrberger, J. eds. 1982. *Sublanguage: Studies of language in restricted semantic domains*. Berlin and New York: de Gruyter.
- Kress, G. and van Leeuwen, T. 1998. Front pages: the (critical) analysis of newspaper layout. In Bell, A. and Garrett, P. eds. *Approaches to Media Discourse*. 186--219. Oxford: Blackwell.
- Mann, W. C. and Thompson, S. A. 1987. Rhetorical structure theory: a theory of text organization. Technical Report RS-87-190 USC/Information Sciences Institute Reprint series.
- Marks, J.W. 1991. A formal specification scheme for network diagrams. *Journal of Visual Languages and Computing*, 2(4):395--414.
- Maybury, M. T. 1999. Research in multimedia and multimodal parsing and generation. *Artificial Intelligence Review*. Forthcoming.
- Mittal, V., Roth, S., Moore, J., Mattis, J., and Carenini, G. 1995. Generating explanatory captions for information graphics. In *Proceedings of the 14th. International Joint Conference on Artificial Intelligence (IJCAI-95)*, volume 2, pages 1276--1283. Montréal, Canada.
- Neal, J. G., and Shapiro, S. C. 1991. Intelligent multimedia interface technology. In J.W. Sullivan and S.W. Tyler, editors, *Intelligent User Interfaces*. Frontier Series. New York: ACM Press.
- Negrine, R. 1994. *Politics and the Mass Media in Britain*. London: Routledge.
- Reichenberger, K., Rondhuis, K., Kleinz, J., and Bateman, J. 1996. Effective presentation of information through page layout: a linguistically-based approach. In *Effective Abstractions in Multimedia, Layout and Interaction*, San Francisco, CA.: ACM Multimedia '95.
- Roth, S. F. and Mattis, J. 1991. Automating the presentation of information. In *Proceedings of the IEEE Conference on AI Applications*. 90—97. Miami Beach, FL.: IEEE.
- Sefton, P. M. 1990. Making plans for Nigel (or defining interfaces between computational representations of linguistic structure and output systems: Adding intonation, punctuation and typography systems to the PENMAN system). Technical report, Linguistic Department, University of Sydney, Sydney, Australia.
- Snell-Hornby, M. 1988. *Translation Studies - An Integrated Approach*. Amsterdam: John Benjamins.
- Swales, J. M. 1990. *Genre Analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Wahlster, W., André, E., Graf, W., and Rist, T. 1991. Designing illustrated texts: how language production is influenced by the graphics generation. In *Proceedings of the 5th. Conference of the European Chapter of the Association for Computational Linguistics*. Association for Computational Linguistics.
- Wahlster, W., André, E., Finkler, W., Profitlich, H., and Rist, T. 1993. Plan-based integration of natural language and graphics generation. *Artificial Intelligence*, 63(1-2):387--427.