

A Model of Genre in Document Layout

Judy Delin
Department of English Studies
and Centre for Research in Communication and Language
University of Stirling
Stirling FK9 4LA
UK
Tel: +44 1786 467974
Fax: +44 1786 466201 (F)
Email: j.l.delin@stir.ac.uk
<http://www.gem.stir.ac.uk/delin.html>

John Bateman
FB10, Sprach-und Literaturwissenschaften
Universität Bremen
28334 Bremen
Germany
Tel: +49/421-218-9483
Fax: +49/421-218-4283 (or 218-7801)
Email: bateman@uni-bremen.de
<http://www.uni-bremen.de/~bateman>

Patrick Allen
Department of Electronic Imaging and Media Communications
School of Informatics
University of Bradford
West Yorkshire
BD7 1DP
Tel: +44 1274 236139
Email: p.t.allen@bradford.ac.uk
<http://www.eimc.brad.ac.uk/staff.html>

Keywords: genre, layout, illustrated documents, computer generation, linguistics, typography

Abstract: We present a framework for the describing the genres of illustrated documents, based on analysis at five levels: content structure, rhetorical structure, layout structure, navigation structure, and linguistic structure. We also include three sources of constraints under which a document might be produced and interpreted: canvas constraints, production constraints, and consumption constraints. Document genres are conceptualised as complex specifications composed of descriptions at each of the five levels that conform in characteristic ways to the three kinds of constraint. We propose that the eight parameters together form a 'space' of possible identities for documents, electronic or paper. The notion of 'genre space' captures the fact that it is possible for new genres to develop at different positions in the space, as well enabling relationships between

document types, or even transformations between them, to be examined and described.

Biosketches:

Judy Delin is Senior Lecturer in Language and Linguistics at the Department of English Studies of the University of Stirling, and is manager of the GeM project. Her research interests are in syntax, discourse analysis, the structure and usability of information and the layout of illustrated documents.

John Bateman is Professor of Applied Linguistics at the English Department of the University of Bremen, specialising in functional and computational linguistics and translation. His research interests include functional linguistic approaches to multilingual and multimodal document design.

Patrick Allen is Lecturer in Electronic Imaging and Media Communications at the University of Bradford. He has a background in art history and knowledge-based systems. His research interests include illustration and the multimodal text, visual culture and the press.

Introduction

As van Leeuwen and Kress (1995) have pointed out, 'All texts are multimodal' (1995:25). On this view, every communicative act, spoken or written, takes place over more than one 'mode' or channel of communication: spoken language involves gesture, for example, while written language always involves other visual elements, such as even the most basic choices of typeface, margins, and headings.

We take the view that language, layout, image, and typography are all purposive forms of communication. Accordingly, in our research project GeM¹, we aim to describe and analyse all these elements within a common framework, thereby providing a more complete understanding of meaning-making in visual artefacts. By analysing resources across visual and verbal modes, we can see the purpose of each in contributing to the message and structure of a communicative artefact.

The research is designed to formalise and model the role of genre, including local and expert knowledge, in layout and typographical decisions. Through the analysis of four sample types of multimodal documents – newspapers, websites, instruction manuals, and illustrated books – the project aims to develop a theory of visual and textual page layout in electronic and paper documents. The model will be implemented in the form of a computer program that allows exploration of both existing and potential layout genres, generating alternative and novel layouts for evaluation by design professionals.

¹ 'Genre and Multimodality: a computer model of genre in document layout'. funded by ESRC, grant no. R000238063.

In the remainder of this section, we outline some of the fields of research that contribute to our approach to the problem of formalising layout and typographical decision-making in the production of multimodal documents.

Genre and Document Layout

Our use of the term *genre* here is similar to Biber's (1989:5-6), who in his study of linguistic variation states that 'text categorizations readily distinguished by mature speakers of a language; for example...novels, newspaper articles, editorials, academic articles, public speeches, radio broadcasts, and everyday conversations...categories defined primarily on the basis of external format'. We adhere, too, to Biber's view that these categories of text also reflect distinctions in the author's *purpose* (1988:68): the documents look different, and contain different language forms, because they are intended to do different things.

In linguistics, there are many attempts to categorise the kinds of language that occur in different genres of texts (see for example, the detailed analysis of Biber, 1988, or the introductory survey in Delin 2000). There are few attempts to extend genre analysis into other aspects of visual meaning, however: Twyman (1985) provides a preliminary scheme for categorising documents according to the interrelationships between images and text, while Waller (1987) is the only attempt extant, to our knowledge, that attempts to describe the role of language, document content, and visual appearance in the formation of document genre within the same framework. Our work draws upon and extends Waller's in several ways, as we shall make clear below.

A final plank of our approach to genre is the central role given to genres' capacity to shift. It is clear, as writers such as Fairclough (1992) have pointed out, that genres can move, and can hybridise with and colonise one another. We see, for example, the influence of web design on other genres. Modelling genres as single entities, then, will not capture their interrelationships, and will always be slightly out of date: a single genre description will not account for how the genre is shifted by the next example, and the one after that. A model of document genre, therefore, must take seriously the suggestion that texts do not only reflect their contexts, but create them, and genre expectations are similarly both reflected and shaped by texts that are instances of those genres. We are concerned to explore document genres, then, as sets of interrelating parameters, and the resulting framework as a 'space' of possibilities for realisation. In this way, we can not only explore the relationships between existing genres, but hypothesise – and with the help of the computer model, even see – 'genre-bending' examples of document genres that don't currently exist, but which easily might.

Computer Generation of Multimodal Documents

Another source of stimulus to the research has been the attempts that have been made in computer science and computational linguistics to generate documents automatically. Automatic generation is often desirable because, while computers are good at storing vast quantities of complex data, it is often difficult to get it out in human-comprehensible form. Much interest exists, therefore, in extracting

information in ways that suit the human purpose, and presenting it in the form of usable text, graphics, or both. Several significant systems have been designed that have studied multimodal information presentation, including examination of how information can be distributed coherently across media. Relevant work in multimodal document design includes the systems WIP (Wahlster *et al.*, 1991, 1993), COMET (Feiner and McKeown, (1990), Mittal *et al.* (1995), and Kerpedjiev *et al.* (1997)), and ICONOCLAST (Bouayad-Agha et al 2000). Bateman *et al.* (2000) investigated the interrelationships between layout, written content, graphical content, and the rhetorical organization of texts in one of the most sophisticated attempts to date: a system that automatically produces sample pages for an encyclopedia of artists' biographies, using textual and graphical representation styles and laying them out appropriately. The computer modelling aspect of GeM allows us to explore and extend this work, using the additional constraint of genre to further determine the appearance of generated documents.

The Analytical Framework

How can we capture the way in which document content is disposed on the page? Waller (1987: 178ff) represents the constraints on the typographer in producing a graphical document as emerging from three sources:

<i>Topic structure</i>	'typographic effects whose purpose is to display information about the author's argument – the purpose of the discourse'
<i>Artefact structure</i>	'those features of a typographic display that result from the physical nature of the document or display and its production technology'
<i>Access structure</i>	'those features that serve to make the document usable by readers and the status of its components clear'

A given phenomenon on the page may reflect the influence of just one of these constraints, or more than one: for example, a chapter heading may arise from access structure (it shows readers how to orientate within the document), but also serve to delineate a topic boundary.

Although Waller does not produce detailed text analyses based on his model, his view that document appearance results from satisfying goals at different levels is persuasive: we particularly take the force of his point that the physical nature of the document and its method of production play a major role in its appearance. In this way, the 'ideal' layout of information on a page may never occur: it must be 'folded in' to the structures afforded by the artefact, and labelled and arranged according to the structures required for access. Document design is therefore never 'free', in the sense that it is never motivated solely by the dictates of the subject matter.

In a revision of Waller's model, we suggest that there is an advantage to be gained in uncollapsing his 'topic structure' into a separation between content and

rhetorical presentation. We view content to be the ‘raw’ data out of which documents are constructed. What Waller describes as ‘the author’s argument’ is not solely or completely dictated by content: many rhetorical presentations are compatible with the same content. Secondly, we take what Waller terms ‘artefact structure’ to be not a structure in the sense that it is a set of ideas to be incorporated in the document, but a constraint on the combination of all the other elements into a finished form. The levels we propose, therefore, are as follows:

<i>Content structure</i>	the structure of the information to be communicated;
<i>Rhetorical structure</i>	the rhetorical relationships between content elements; how the content is ‘argued’;
<i>Layout structure</i>	the nature, appearance and position of communicative elements on the page;
<i>Navigation structure</i>	the ways in which the intended mode(s) of consumption of the document is/are supported; and
<i>Linguistic structure</i>	the structure of the language used to realise the layout elements.

We suggest that document genre is constituted both in terms of levels of description, and in terms of the constraints that operate on the information at each level in the generation of a document. Document design, then, arises out of the necessity to satisfy communicative goals at the five levels presented above, while also addressing a number of potentially competing and/or overlapping constraints:

<i>Canvas constraints</i>	Constraints arising out of the physical nature of the object being produced: paper or screen size; fold geometry such as for a leaflet; number of pages available for a particular topic, for example;
<i>Production constraints</i>	Constraints arising out of the production technology: limit on page numbers, colours, size of included graphics, availability of photographs; for example, and constraints arising from the micro-and macro-economy of time or materials: e.g. deadlines; expense of using colour; necessity of incorporating advertising;
<i>Consumption constraints</i>	Constraints arising out of the time, place, and manner of acquiring and consuming the document, such as method of selection at purchase point, or web browser sophistication

and the changes it will make on downloading; also constraints arising out of the degree to which the document must be easy to read, understand, or otherwise use; fitness in relation to task (read straight through? Quick reference?); assumptions of expertise of reader, for example.

A model of genre, therefore, must begin by expressing adequately the above five levels of description as well as finding the most appropriate way of satisfying the three sets of constraints.

There must be, in our view, further 'meta-layer' constraints on how pages are realised. A computer model of genre should also be sensitive to the generic qualities that occur across the entire artefact. In many recent bird field guides, for example (a genre of text that we have discussed in Allen *et al.* 1999), there is a high premium on each page looking similar: a computer model should not therefore find a different solution to the layout of information on each page. In a newspaper, however, the opposite is the case: news layout is expected and intended to vary page by page, and appearance changes, sometimes radically, between parts of the paper, such as between a news section and a features section. In the bird book, the model would produce similarity to account for the navigation and indexing required by the bird book form. In the newspaper, the model would generate difference in order to make clear the distinctions being made by the different types of content, in this case between news and feature articles. Both exemplify each of the three constraints described: canvas, production and consumption.

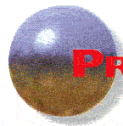
It is precisely these regularly recurrent and stable selections and particular sets of constraint satisfactions that we call *genre*.

In the next section, we illustrate our approach by applying it to a single page, taken from the reference source 'Wildlife Explorer', a set of week-by-week sheets that build into an encyclopedia for young readers. We refer to this as 'the tiger document'².

A Worked Example: The Tiger Document

We selected this example (shown in Figure 1) simply to illustrate our approach: we do not thereby imply any judgement as to its merits as a piece of information design. In working through the five steps of the approach, however, we hope to show as a side-effect how an assessment of the relationships between decisions at different levels can form the basis for critical thinking about design and layout.

² We are grateful to Rob Waller for his comments on the layout of the tiger document, and his suggestions as to how to improve our analysis of it. Any remaining shortcomings are of course our own.



PROFILE BENGAL TIGER

The tiger's powerful muscles and massive build are the keys to its hunting success: in seconds, it can knock down and kill prey that weighs nearly a tonne.

EYES

The tiger's night vision is six times better than our own, aided by a mirror-like layer at the back of the eye that reflects extra light.

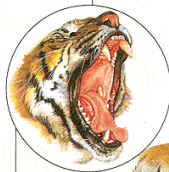
EARS

Hearing is the tiger's sharpest sense. The white spots behind the ears help tigers to identify one another in the dark jungle.



COAT

Every tiger has a unique pattern of black stripes on a deep-orange coat. This breaks up the outline of the body in dense cover.



TEETH

The tiger uses its long canine teeth to stab and kill prey. The molars behind them act like scissors, slicing strips of flesh from a carcass.



CLAWS

The claws are used to grip prey and to scratch trees. They retract when the tiger walks, to remain sharp and to allow it to stalk prey silently.

VITAL STATISTICS	
WEIGHT	180-265kg
LENGTH	1.9-2.2m
Head & Body	80-90cm
Tail	
SHOULDER HEIGHT	90-95cm
SEXUAL MATURITY	Female 3-4 years; male 4-5 years
MATING SEASON	Winter to spring
GESTATION PERIOD	95-112 days
NUMBER OF YOUNG	2 to 4
TYPICAL DIET	Sambar deer, chital deer, buffalo, wild pigs, gaur and monkeys
LIFESPAN	Up to 26 years in the wild

RELATED SPECIES

There were once eight subspecies of tiger. Today, only five remain: the Bengal, the Siberian, (below left), the Sumatran, (below right), the Indochinese and the Caspian. The tiger's closest relatives are the other 'big cats' in the genus *Panthera* — the lion, leopard, jaguar and snow leopard.



CREATURE COMPARISONS

The Siberian tiger is the world's largest cat: adult males may reach more than 3.6m in length. The coat is shaggier and paler than that of the Bengal tiger, equipping the tiger for its icy northeast Asian habitat. The Siberian tiger occupies a vast territory and preys mainly on wild pig. There may be no more than 150 left in the wild.

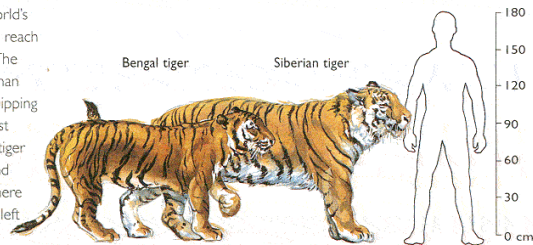


Figure 1: The tiger document

Content Structure

A simple way of getting at the content of the document is to list informally what elements are covered in either medium, textual or graphical. This results in the unordered list in Figure 2.

tigers
length
bengal tiger
mating
other tigers
sexual maturity
hunting success
gestation
how it is so successful in hunting
diet
appearance
young
operation of claws
mating season
function of claws
height
function of canine teeth
size of Siberian tiger
function of molars
appearance of Siberian tiger
appearance of teeth
Siberian tiger
sense of hearing
height of Siberian tiger
white spots behind ears
territory of Siberian tiger
function of white spots
food of Siberian tiger
appearance of white spots
scarcity of Siberian tiger
how eyes work
coat of Bengal tiger
quality of vision
height Bengal tiger
appearance of coat
Relative Siberian tiger size
function of stripes
Relative Bengal tiger size
weight
Relative man size
other cats
other tigers

Figure 2: List of document content elements

This list, however, suggests that there is no logical ordering or hierarchy inherent in the content. There is, of course: the content can be organised into five main segments: an overview (given in bold type just under the title), the appearance and functional description of the Bengal tiger, its size, mating habits, etc. , physical comparisons with other tigers, and related species. Figure 3 represents the hierarchical relationships between the content segments of the tiger document.

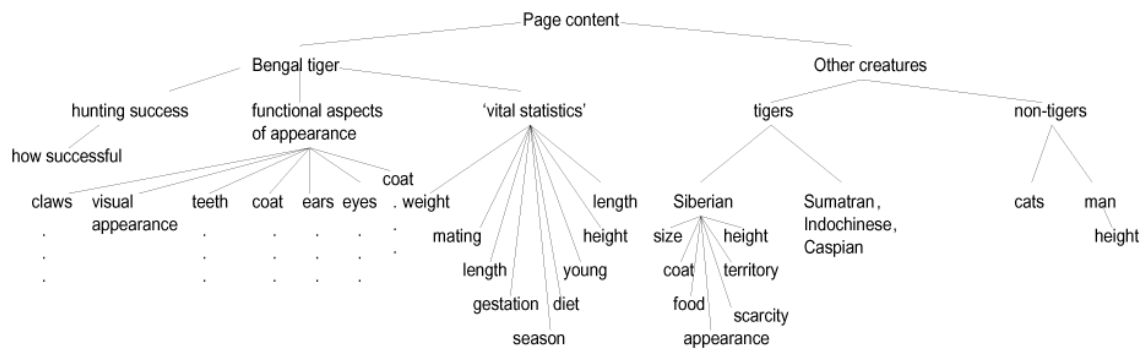


Figure 3: Hierarchical representation of tiger document content

This hierarchical representation makes clear what parts of the content are related, and does not discriminate between what is presented linguistically and what graphically. We can now see that there are two main sets of facts provided in the document: a larger set relates to the Bengal tiger, and the other relates to other tigers and other animals.

The representation of document content does not seek to capture how the content is argued, nor how it is represented on the page. Statements, for example, about the Siberian tiger, other big cats, or about the height of the average human, are treated in the content analysis as factual propositions. Of course, in the document, this content is presented for the purposes of comparison with the Bengal tiger. However, it is not the job of the content representation to make clear *why* this page features such content, or what its role is in the argumentation presented (comparison is an example of such a role). In this way, we diverge from Waller's 'topic structure' which collapses the content and rhetorical levels, making it more difficult to discuss alternative rhetorical presentations, perhaps with other argument or artefactual constraints. This separation is, however, necessary, since we need to explore the space of more or less related genres and to provide a systematic account of the possible variation within that space. It is clearly important to allow the analysis to reflect the fact that content and rhetorical presentation can vary independently of one another.

Rhetorical Structure

The *rhetorical structure* of the text (language and pictorial elements) is the way in which the content is argued and the various segments interrelated textually. In order to analyse rhetorical structure, we use used a framework known as Rhetorical Structure Theory or RST (Mann and Thompson 1988). RST provides a set of concepts and a notation to express the way in which segments of text are hierarchically related to one another in the presentation of a coherent text. Nash

(1980) has proposed a small number of rhetorical structures to which documents can adhere. The closest to the tiger document is what he has described as a 'stack': a topic is established at the outset, and 'becomes the nodal point of divergence and convergence'. This largely captures the idea of a central image around which other content elements are arranged, but the additional flexibility of RST allows us to capture what is in fact the norm – situations in which documents diverge from stereotypical structures.

For the purposes of explication, we will restrict the analysis we present here to the main blocks of the page: the large tiger picture and accompanying text and inset pictures, and the 'vital statistics' panel. What is obvious here is the centrality of the tiger picture, and the role which smaller elements play (descriptions of the functions of its attributes, such as coat, teeth, claws) in *elaborating* upon that central image. Within the smaller elements there are also notable rhetorical relationships: there is a description of the tiger's coat and an explanation of its function (in RST terms, a *purpose* relationship between those two parts of the argument), and there is a description of its eyes and how they work (a *means* relationship), for example. RST was developed simply to deal with texts: here, we assume that any part of a rhetorical relationship can be realised graphically, as well as textually. The text therefore decomposes into segments that are related rhetorically, and which in turn are amenable to further decomposition into sub-segments, again rhetorically related to one another. RST distinguishes between the part of the rhetorical relationship that is core, central or *nuclear*, and the peripheral, additional part that is referred to as a *satellite*. For example, in the segment describing the tiger's coat, we can further analyse the text into a nucleus that states that every tiger has a unique pattern of stripes and a satellite that provides the purpose of these stripes. Similarly, there is an elaboration describing the teeth of the tiger that itself has internal structure with two statements concerning the canines and molars, both of which are subordinate to their location in the tiger's mouth. In both cases, the same content could have been organised rhetorically quite differently, for example, by simply listing the fact that there are teeth, or by providing less detail, for example omitting the purpose.

The diagram in Figure 4 represents this part of the tiger document in RST terms. The curved lines point from satellites to nuclei, so that a concentration of lines converging on one segment shows the most 'central' element of the document. In some cases, however, the sub-segments are all as important as one another, and act more like a list. There is no nucleus-satellite relationship in this case, and the elements are referred to as being in a 'multinuclear' relation, designated by straight lines (as is the case with the tiger's attributes in the second elaboration segment).

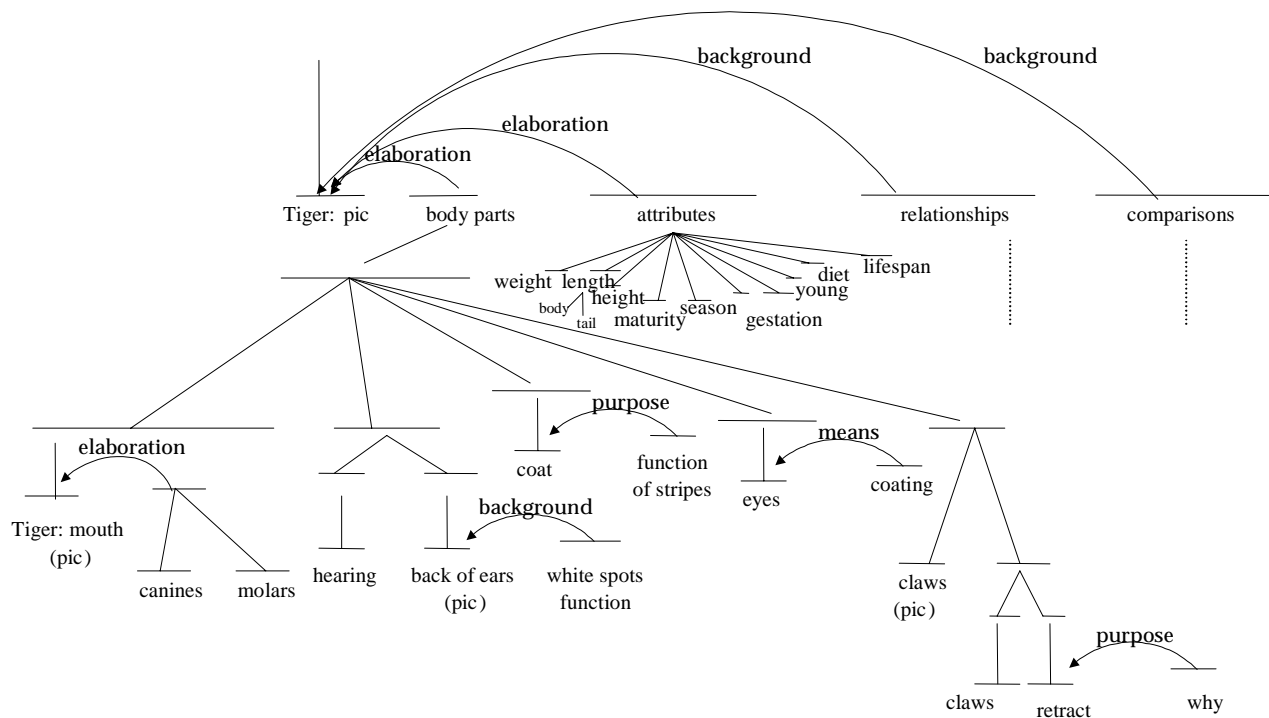


Figure 4: Rhetorical analysis of 'tiger' document

The main structure of the document, then, is based on two main elaborative segments around the nuclear tiger picture: the tiger's body parts, and other attributes. The relationships with other tigers and other animals are presented as background.

At this stage, we can see how the rhetorical structure draws upon the content structure to make an argument. Hierarchy is clear within it: we know, for example, how all the elements belonging to the functional description of body parts relate to one another, and that those elements are themselves complex. For example, since the segment describing the function of the tiger's stripes is a 'purpose' satellite of the nucleus that first mentions the stripes, the nucleus should be presented first and the two should not be split up. Similarly, the three elements describing the teeth have an internal hierarchy, with the two statements about canines and molars subordinate to the picture showing the tiger's mouth.

Layout Structure

The next stage of the analysis provides a detailed characterisation of the concrete layout decisions that have been made in a given document. Layout is described in terms of a hierarchically organised set of layout elements and the properties (graphical, typographical) of those elements. Just as different rhetorical organisations can be selected for communicating a given content structure, diverse layout structures can be selected for any given rhetorical organisation (cf. Bateman et al. (subm.) for extended discussion). Again, this is the primary motivation for maintaining these as distinct descriptive levels. In what follows, we will list briefly some of the layout characteristics to indicate the approach, although a full detailed analysis would take us beyond the scope of this paper.

In the case of the Tiger document, the page falls broadly into three blocks: the top left block (we will call this block A), with labels organised around the central illustration of the tiger, the right-hand column (B), consisting of two panels of information, and a third block (C) consisting of text and illustrations below block A. Each block has a different organising principle. In block A, labelled text entries and detailed illustrative call-outs are ranged around the central illustration, as closely as possible to the relevant feature described, although positioning is clearly also determined by available space. The text around the tiger is in sans serif 10pt mixed case, ranged left, ragged right, while the labels are in the same typeface, 12pt small caps, large initial cap, ranged left, ragged right (note that the document is reduced in reproduction). It is organised to preserve alignment along the top and bottom of the text and to keep white space around the large illustration. The illustrations are full-colour artwork, the smaller ones with a 1pt border. In block B, the first element is another labelled list except this time more simply organised in a column, entry labels are in a serif font, small caps, matching the title of the whole page (although smaller) rather than the labels of the adjoining block B. Spacing of the entries is tight against the red title 'vital statistics' but then generous, including an extra large space between 'typical diet' and 'lifespan' to allow space for the intruding tail of the tiger. The lower panel of block B has a darker red title and a bright red bullet before the bold, sans serif text, two square photographs at the bottom of the panel. In block C, the title 'Creature Comparisons' is in black with a small graphic, with light sans text ranged left and the full-colour and line drawing graphic to the right. The whole is bordered in dark red.

The breaking-down of the layout structure into distinct elements, and elements within those elements, therefore allows us a clear hierarchical analysis for the page, and an association with typographical characteristics with each of the elements.

Navigation Structure

The navigation structure consists of the 'signposts' that enable the rhetorical structure as it has been disposed upon the page to be understood. Given the rhetorical decomposition above, we would expect major segments to be given a label of some kind: this is the case with 'vital statistics', 'related species' and 'creature comparisons'. The dominant status of the information in block A is indicated by the greater space it occupies on the page, its top-left positioning, and its lack of a title: it seems to share the title of the whole page. The text labels and detailed pictures are linked to the illustration by lines, reinforcing their related status.

As we noted in the layout structure, page elements are differentiated by size, positioning, boxes, and panels. The fact that the two panels we are referring to as block B somehow 'go together' is indicated by the fact that they are the same (sandy) colour, although the different colours of titles (bright red top, dark red bottom) have the effect of dissociating them. The overlap of the tiger's tail with the upper panel in block B has the effect of linking that panel with block A, and integrating the page as a whole.

'Creature comparisons' as a unit diverges to quite an extent from content structure, which simply recorded the existence of statements about the Siberian tiger and about man, but as unrelated to one another. In the representation on the page, however, the relationship between the facts about height and size of the different animals is clearly intended to be one of comparison and contrast, and this is represented graphically by the overlapping illustrations and the proximity to the centimetre rule at the left.

The two uncaptioned pictures at the bottom of block B are averted to by references in the text, 'below left' and 'below right', and lines connect the text about the tiger both to the smaller detail pictures and to the main illustration. This explicit labelling to direct the reader within a document is common to several of our genres, such as the newspapers' need to refer readers of a story to more content 'continued on page 34'. The navigation structure is where such internal referencing or 'document deixis' is naturally captured (for a discussion of the phenomenon of document deixis, see Paraboni, 2000).

Linguistic Structure

Language on this page can be divided into several types. Titles and labels consist only of nouns (*eyes, claws*) and noun phrases (*vital statistics, related species*). These elements are selected for word-play: the double-meaning of 'vital statistics', for example, or the alliteration of 'creature comparison' (perhaps also a passing reference to the phrase 'creature comforts', since 'creature' is otherwise an odd word to use in this context). The vital statistics list contains a variety of grammatical forms but none of them complete clauses, and the complete sense of them is only inferable in the context of the labels. Body text around the main illustration, and in the introductory paragraph at the top of the page, is evaluative: the tiger is *powerful, massive*, and it has *success* in hunting, for example. The structure of each text entry around the tiger is to evaluate positively each of the highlighted characteristics: the tiger's night vision is *six times better than our own*, it has a *unique* pattern of black stripes, and it stalks prey *silently*. The text in box C, 'Creature Comparisons', and in the 'related species' panel, is more factual than evaluative, although signs it has been written for young readers exist in the Siberian tiger's coat being described as *shaggier* and its habitat *icy*.

Constraints

We noted at the outset that a description of genre does not only consist of analysis at the five levels proposed, but of observing constraints on production. Here, an obvious constraint is the necessity of fitting all the information on one side of paper, given the canvas constraint presented by the chosen encyclopedia format: one page per entry. This constraint would have an effect in the navigation structure: the status of the new entry is made clear by a title. Although this is a page with several blocks of information, it appears that the user is expected to traverse the page from the top left, a conventional consumption strategy. This consumption constraint is reflected in the layout structure by the positioning of the main graphical and textual element (block A) in the top left of the page. A

consumption constraint also exists in terms of the expectation of the readership: young readers require simple, engaging text (reflected in the linguistic structure) comparing familiar concepts with new ones, and prefer enticing and often highly-coloured productions (reflections in the layout structure) with simple navigation.

Discussion

We have now described the characteristics of the tiger document on all five of the levels we propose in our framework, and briefly sketched the primary additional constraints that play a role in the construction of a complete document. We are currently collecting a set of documents from our selected genres and describing these in detail at each of the five levels. What we are seeking, however, are systematic relationships between the different levels of description. For example, the layout structure informs us that block A (the tiger) is presented as being more important: the natural question then to ask is whether this reflects the RST analysis: one might expect that the interpretation of a document would be complicated if rhetorically core elements are presented as less important by the selected layout structure. Similarly (and much more frequently), the rhetorical organisation shows the relationships that hold between distinct parts of the content elements being presented: a further question is then to what extent the selected layout structure supports or hinders the ability of the reader to recognise those relationships.

The Tiger document illustrates several aspects of this latter issue very well. The content of the first elaboration in the rhetorical structure (cf. Figure 4), the physical, functional characteristics of the tiger, are kept together in the layout structure: they are arranged around the main tiger picture upon which they elaborate. There is no nucleus-satellite relation between elements in this list: they are instead arranged where they would fit around the tiger, as closely as possible to the points they refer to, but kept separate. Furthermore, the rhetorical distinctions within these text labels have been realised to a particular level of delicacy. The text entries for 'eyes', 'ears', 'coat' etc. do not visually distinguish 'means' relations (how the eyes work) and 'purpose' relations (what the stripes are for): all the content is realised as plain text labels. The second elaboration, again unordered, contains the 'vital statistics' information. This is separated in a panel, although the list elements are labelled. In keeping with a left-to-right reading strategy, the large central segment is placed top left.

The layout structure selected in the Tiger document thus broadly observes the rhetorical and content distinctions, although there is slippage between layout and rhetorical or even content structure. The selections of light and dark type, and between serif and sans fonts, do not appear to be functionally motivated from the rhetorical structure and so we need to ask whether this differentiation is helpful or distracting for the reader of the page. Similarly, the alignment of panels on the right also appears not to carry any functional load. We would also want to account for the relative size of the three panels in relation to their function, their placement on the page, and the rationale for using a colour background in only two of the panels. Similarly, the back view of the tiger's ears is

in fact an equal partner in the 'ears' list element that elaborates on the main picture, but is not clearly connected to either the text that is its sister or the picture that is its parent. Indeed, the proximity of the picture of the tiger's ears is closer to the description of its coat, and is thus a stronger cue than the line from the text to the main image: this does not coincide with the rhetorical intent. Finally, the circular picture border for the tiger's mouth lacks coherence with the rectangular borders around the other inset pictures, given that they perform equivalent rhetorical functions.

In general, then, we are seeking to motivate the design decisions taken within the layout structure in terms of the functional discriminations needed to communicate the rhetorical structure. The relationship between levels is to some extent 'natural'; that is, when there are commonalities in layout decisions across distinct elements (be they of size, colouring, type face selections, alignment, or whatever), then there is a natural tendency for the reader to interpret those elements as being in some way rhetorically related. Conversely, when there are distinct layout decisions made, then there is a tendency for rhetorical discriminations to be perceived. The detailed descriptions given by our levels of description now allow us to state and probe these general tendencies with a far higher degree of precision than possible hitherto.

Genre in the Framework

Whereas many of the detailed decisions taken in the layout structure support discriminations required to interpret the rhetorical structure, there are also broader, more 'global' decisions that contribute to a document's appearance. In order for a document to achieve a consistent presentational style, allocation of layout properties to functional motivations has to be fixed rather than being allowed to vary freely. Thus, although it would have been possible to signal the rhetorical discriminations in the Tiger page in many other ways in the layout structure, these decisions should not, of course, be conceived of locally, but rather as solutions that have been decided for the set of Wildlife Explorer documents as a whole. Thus each page concerning an animal will be using similar layout properties to make similar rhetorical and navigational properties clear. We can label these consistent layout decisions for an entire publication a 'micro-genre': this corresponds directly to a fixing of relationships between our levels and the applicable artefact and production constraints and is responsible for a publication having a coherent look-and-feel.

More significantly, however, is that there appear to be broad classes of cross-level decisions that are not limited to single documents but which reoccur. Thus, encyclopedia pages are clearly distinguishable from newspaper pages which are themselves clearly distinguishable from earlier web pages and instruction manuals. Any given microgenre for a publication is not then a unique, out-of-the-blue event; it instead

combines and re-uses many aspects of the overall genre to which the publication is to be identified as belonging. In the case of the Tiger document, this extends down to the placing of some picture in the middle and providing information in call-outs, with extra information round the edges in boxes, in this respect the document is very similar to many other information-giving genres.

By analysing our growing collection of documents, annotated by the levels of analysis that we have described above, we are now seeking to codify these broader classes of consistent layout decisions and motivations in order to describe particular genres. The flexibility of the relationships between levels supports our claim that genres can change and grow, as well as colonise and subvert each other. Thus the tendency of many web pages now to look increasingly like their traditional print-based counterparts can be seen as a change in the allocation of layout resources to rhetorical functions: this is of course supported both by the changing technical basis (artifact and canvas constraints) and perceptions of what function such a web-based document should be fulfilling, and for whom.

References

- Allen, Patrick T., Bateman, John A., & Delin, Judy L. (1999). Genre and layout in multimodal documents: towards an empirical account. In: Power, Richard. & Scott, Donia (eds.) *Using Layout for the Generation, Understanding, or Retrieval of Documents*. Proceedings of the American Association for Artificial Intelligence Fall Symposium, North Falmouth, Massachusetts, November 5-7, 1999. Technical Report FS-99-04. Menlo Park, CA: AAAI Press. 27–34.
- Bateman, John, Kamps, Thomas, Klein, Jörg, and Reichenberger, Klaus. (subm.). The Dartbio system: constructive text, diagram and layout generation for information presentation. Submitted to Computational Linguistics.
- Biber, Douglas (1988). *Variation Across Speech and Writing*. Cambridge: Cambridge University Press.
- Biber, Douglas (1989). A typology of English texts. *Linguistics*, 27: 3-43.
- Bouayad-Agha, Nadjet, Scott, Donia R. & Power, Richard J. D. (2000). Integrating Content and Style in Documents: a case study of Patient Information Leaflets. *Information Design Journal*, 9 (2-3), 161-176.
- Delin, Judy. (2000). *The Language of Everyday Life*. London: Sage.
- Fairclough, Norman. (1992). *Discourse and Social Change*. Cambridge: Polity Press.
- Feiner, Steven K. and McKeown, Kathleen R. (1990). Co-ordinating text and graphics in explanation generation. In: AAAI-90: Proceedings of the 8th International Conference on Artificial Intelligence, volume I. 442-449. Menlo Park: AAAI Press/The MIT Press.
- Kerpedjiev, Stephen, Carenini, Giuseppe, Roth, Steven F. & Moore, Johanna D. (1997). Integrating planning and task-based design for multimedia presentation. In: Proceedings of IUI-97. International Conference on Intelligent User Interfaces, Orlando, Florida. Association for Computing Machinery.

- van Leeuwen, Theo & Kress, Gunther (1995). *Critical Layout Analysis*.
Internationale Schulbuchforschung 17, 1995. Frankfurt: Diesterweg. 25-43.
- Mann, William, C. & Thompson, Sandra A. (1988). "Rhetorical Structure Theory:
Toward a functional theory of text organization." *Text* 8 (3): 243-281.
- Mittal, Vibhu O., Roth, Steven F., Moore, Johanna D., Mattis, Joe, & Carenini,
Giuseppe (1995). Generating explanatory captions for information graphics.
In: Proceedings of the 14th International Joint Conference on Artificial
Intelligence, Montréal, Canada (IJCAI-95), Volume 2. 1276-1283.
- Nash, Walter (1980). *Designs in Prose*. London: Longman.
- Paraboni, Ivandré (2000). Describing Document Parts. *Proceedings of the 3rd UK
Special Interest Group for Computational Linguistics (CLUK-3)*, University of
Brighton, UK. 34-41.
- Wahlster, Wolfgang, André, Elisabeth, Graf, Winfried, & Rist, Thomas (1991).
Designing illustrated texts: how language production is influenced by
graphics generation. In: Proceedings of the 5th Conference of the European
Chapter of the Association for Computational Linguistics. Association for
Computational Linguistics, 1991.
- Wahlster, Wolfgang, André, Elisabeth, Finkler, Wolfgang, Profitlich, Hans & Rist,
Thomas (1993). Plan-based integration of natural language and graphics
generation. *Artificial Intelligence*, 63, 1-2. 387-427.
- Waller, Robert. (1987). *The Typographic Contribution to Language*. Unpublished
PhD Dissertation, Department of Typography and Graphic Communication,
University of Reading.