The DArt$_{bio}$ system: constructive text, diagram and layout generation for information presentation

John Bateman
Bremen University

Thomas Kamps
Darmstadt University of Technology

Jörg Kleinz
IPSI

Klaus Reichenberger
Intelligent Views, Darmstadt

In this paper we present the prototype artist biography information system DArt$_{bio}$, a system for automatically producing multimodal pages of information that appropriately combine textual and graphical representation styles. We concentrate particularly on the selection and function of layout in these pages and demonstrate that layout offers a rich resource for achieving presentational coherence. Layout mechanisms have been developed on the basis of an empirical investigation of professionally produced layouts. These decompose communicative goal structures into layout structures that constrain the placement of automatically generated graphical, diagrammatic and textual elements. Our system builds on an architectural synthesis of data-driven aggregation in information visualization and communicative-goal oriented natural language generation. The paper introduces the component technologies employed, describes the methodology adopted for the empirical investigation of multimodal page layout, and shows the results as implemented within DArt$_{bio}$.

1 Introduction

The desirability of combining text, layout, graphics, diagrams, 'punctuation' and typesetting in order to present information most effectively is widely recognized—indeed, in traditional graphic design, publishing, and information design, they could scarcely be conceived of as separate. It is therefore natural that *computational* attempts to synthesize texts, diagrams, and layout automatically should also have converged (cf. (Feiner and McKeown, 1993; André et al., 1993; Fasciano and Lapalme, 1996; Kerpedjiev et al., 1997)). As such presentations become more complex, it is essential to consider how information presented in a variety of modes (e.g., textual, graphical, diagrammatic, etc.) can be presented coherently *as a whole*—as Maybury (1995) notes,

> "The physical format and layout of a presentation often conveys the structure, intention, and significance of the underlying information and plays an important role in the presentation coherency."

In order to investigate this problem more closely, we have developed DArt$_{bio}$ ('Dictionary of Art: biographies'), a prototype multimedia/multimodal information presentation system which generates multimodal pages from a single source of information shared across the system's presentation modalities; this supports presentation in textual, graphical, and layout form. We argue that pages of this kind are best seen as *single communicative moves* in an ongoing multimodal information dialogue (cf. (Stein and Thiel, 1993; Kamps and Reichenberger, 1995; Dale, Milosavljevic, and Oberlander, 1997)) and are driven in content and form by communicative intentions just as are individual texts. Moreover, in contrast to previous systems for multimodal presentation where layout is generally derived directly from communicative goal structures, we show that a more indirect relationship is necessary for sophisticated layout behavior.

We organize our discussion of the DArt$_{bio}$ system as follows. We first introduce the contributing technologies and components on which DArt$_{bio}$ builds (Section 2). We then set out the particular empirical approach to exploring functionally-motivated layout that we pursued in the development of the layout component of DArt$_{bio}$ (Section 3) and summarize the results of this study as an abstract specification for performing page layout (Section 4). With this apparatus in place, we introduce an application scenario in which information presentation of the kind we envisage is necessary and show how the DArt$_{bio}$ system takes a significant step towards providing it (Section 5); here we work through a detailed example showing the points of controllable flexibility that our approach opens up for automatic page design. Finally, we place our work on DArt$_{bio}$ in the context of the current state of the art in multimedia generation, highlight the points of synergy found across the component technologies, and outline some of the follow-up research and development to which our work is now leading (Section 6).

2 Baseline components and methods

The DArt$_{bio}$ prototype synthesizes a variety of techniques and approaches from both computational linguistics—particularly natural language generation—and visualization. Advanced visualization techniques are used to produce graphical diagrams fully automatically, while state-of-the-art full text generation ensures flexible natural language rendition.

DArt$_{bio}$ was implemented on top and around a generic 'Editor's Workbench' for supporting editorial work for large-scale publications such as encyclopedias. The design of this workbench was itself strongly motivated by earlier user studies, in which it was shown that editorial work is facilitated by the provision of 'content-driven' document access and selective 'overviews' of those documents' content (Rostek, Möhr, and Fischer, 1994; Kamps et al., 1996). The first large-scale application of the Editor's Workbench[1] was the preparation phase of a large-scale encyclopedia on art and art history then being planned by MacMillan publishers. Within the this project, the Editor's Workbench provided visualization and knowledge-based editing for a knowledge base consisting of approximately half a million concepts. This called for the development of a range of flexible au-

1 Carried out within the European Union 'Research and Development in Advanced Communications Technology in Europe' (RACE) project 2042 'EUROPUBLISHING' (Hüser et al., 1995).

tomatic visualisation tools capable of providing a high degree of control over all the visual aspects of their presentations—including the layout of information and diagram design (cf. (Reichenberger, Kamps, and Golovchinsky, 1995; **?**; Kamps, 1997)). These visualization and layout mechanisms were subsequently adapted for the DArt$_{bio}$ prototype and we introduce them briefly in Sections 2.2 and 2.3 below.

The concepts of the Editor's Workbench knowledge base were a combination of facts automatically acquired from around 14,000 encyclopedia articles (mostly biographical information such as birth-dates, dates of working in particular institutions, dates of movements, works of art created, etc.) and information from imported thesauri such as the Getty Arts and Architecture Thesaurus and Princeton WordNet. The knowledge base also contained relevant pictures (e.g., of artists and artworks) and full text representations of several thousand biographies, thus offering an excellent testbed for the multimedia generation functionality that we wished to explore with DArt$_{bio}$.

2.1 Natural Language Generation

While the visualization and layout components of DArt$_{bio}$ were already being developed within the Editor's Workbench, the NLG component employed—KOMET (cf. (Bateman and Teich, 1995))—was part of an ongoing independent research effort aimed at providing generic natural language generation technology. KOMET is based on the PENMAN natural language generation reference architecture (cf. (Mann and Matthiessen, 1985; Matthiessen and Bateman, 1991)) as further developed for practical multilingual generation in the KPML system (Bateman, 1997). In KOMET, the entire generation process is *genre-based*: i.e., properties of text structuring (particularly thematic development), content-selection from the knowledge base, and details of linguistic form within sentences are all determined by a general model of text types, or 'genres'. The ready specification of the linguistic details of particular text types appropriate for particular applications enables *komet* to meet the often conflicting goals of 'genericity' of design' and application-specific customization.

An illustration of this drawn from the Editor's Workbench is the following. The Editor's Workbench knowledge base contains domain knowledge involving any of the activities and events in which particular artists are or were involved during their careers, as well as the circumstances under which these events took place and further information about particular works of art (e.g., the museums or galleries exhibiting them). Crucially, *it is not possible from this information alone to construct any particular text*. Only when a text type has been specified (either implicitly by virtue of a system's design, or explicitly as in KOMET), is it possible to state how the information is to be selected, organized, and presented. The specification of required text type therefore provides both an organizing structure for this information and a filter by which particular application-specific NLG behavior can be achieved.

The resource architecture of KOMET follows closely the proposal for a multistratal set of linguistic resources made in Martin (1992): the strata are, from

Generic Text Structure Stages	
1	Names, nationality and profession
2	Birth and death
3	Education
4	Major activities, major works
5	Influences, analogies to other artists
6	Impact

Figure 1
Example generic structure for a genre of biographies.

most abstract to least, genre, register (or context), semantics and lexicogrammar. These strata are represented as *system networks*, with one network for each stratum, implemented in the manner originally defined for grammar by the PENMAN system (cf. (Matthiessen and Bateman, 1991)). The networks include disjunctive and conjunctive entry conditions and are usually quite large (e.g., of the order of several hundred nodes *before* multiplying out the extensive cross-classifications).

The text generation process begins from a set of pointers into the knowledge base indicating the broad area of focus[2] and a set of 'genre' features that determines the text type of the text to be produced. Generation then proceeds by successively 'traversing' the network of each stratum. Each traversal produces a set of linguistic features, i.e., the set of features lying on all the paths followed through the network during that traversal. The full set of features selected at each stratum specifies constraints that collectively determine a linguistic structure. At the stratum of lexicogrammar, the structure determined is a syntactic structure, at register, a 'contextual structure', and at the stratum of genre a so-called *generic structure* (Hasan, 1978) or text macro-structure.

As shown in the example text macro-structure of Figure 1, text macro-structures are made up of sequences of text stages. Each text stage carries constraints concerning both the form of language required and its content: The former guarantees that the text generated is appropriate to its content and intended reader; the latter guarantees the expression of only those parts of the information available in the domain knowledge that are required by the text type. Such macro structures then guide the selection of material from the Editor's Workbench knowledge base. Figure 2, for example, shows information returned when generation is focused on the Bauhaus-trained artist Anni Albers.[3] The genre features corresponding to the particular text type to be generated then impose constraints that partition this 'pool' of information into individual text stages. A partition compatible with the biography genre given in Figure 1 identifies facts $x2, x3$ (naming) as belonging to the *Name* text stage, $x6$ ('bearing') as belonging to stage

2 In the general case this focus is determined by the previous move in the dialogue between user and information system; in the DArt$_{bio}$ prototype it was generally the results of an information retrieval query placed by the user in the preceding 'turn'—cf. Section 5.
3 Note: The names, facts, institutions, periods, etc. used throughout this paper are selected for illustrative purposes and should not be taken as reliable statements of art history!

```
(x1 Female )
(x2 Name-Relation  :range x1 :domain "Albers")
(x3 FirstNames-Relation  :range x1 :domain "'Anni'")
(x4 Place )
(x5 Name-Relation  :range x4 :domain "Berlin")
(x6 Bearing  :actee x1 :temporal-locating "12. June 1899" :spatial-locating x4)
(x7 Place )
(x8 Name-Relation  :range x7 :domain "USA")
(x9 Settling  :spatial-locating x7 :actor x1 :temporal-locating "1933")
(x10 Person )
(x11 Name-Relation  :range x10 :domain "Brandenburg")
(x12 FirstNames-Relation  :range x10 :domain "'Martin'")
(x13 Studying  :inclusive x10 :actee "art" :temporal-locating "1916 - 1919" :actor x1)
(x14 Institution )
(x15 Place )
(x16 Name-Relation  :range x15 :domain "North Carolina")
(x17 Spatial-Locating  :range x15 :domain x14)
(x18 Name-Relation  :range x14 :domain "Black Mountain College")
(x19 Teaching  :actor x1 :temporal-locating "1933 - 1949" :spatial-locating x14)
⋮
```

Figure 2
Extract of information pool returned from the Editor's Workbench and converted into generation component-compatible terms by means of a simple translation mapping (cf. Bateman, Teich and Alexa, 1998)

Birth&Death, fact $x13$ ('studying') as belonging to *Education*, and facts $x9, x19$ ('settling' and 'teaching') as belonging to the text stage *Major Activities*.[4]

The facts grouped within each text stage are then prepared for linguistic expression by a combination of text structuring (particularly thematic progression) and aggregation. Individual text stages typically specify preferences for particular kinds of thematic progression. For example, the two generic stages *Name* and *Birth&Death* favor realization with thematic development focusing on the artist; this supports single sentences aggregated around the artist and combining all the relevant information provided by the knowledge base for that text stage—e.g., place of birth, date of death, etc. The other generic stages have more scope for variability. Some subgenres select a thematic progression around the artist, where others call for *temporal* thematic development where aggregation is largely limited to grouping events that took place at the same time.[5] The resulting information groupings, already organized into sequences respecting their selected thematic progressions, are then converted by the register network into sequences of semantic specifications and are passed on to the grammar for tactical realization.

Some results of KOMET's genre-based generation for short biographies are shown in Figure 3. Here each successive pair (i.e., (a-b), (c-d), and (e-f)) consists

4 The selection of particular main 'events' in the genre-constraints corresponds to the definition of *register heads* in Bateman and Paris (1991). Essentially, particular semantic domain model types (e.g., 'creating') together with specified roles (e.g., ':actor Annie-Albers') are allocated to particular text stages. All matching concepts from the knowledge base are then returned for processing by the NLG component.
5 Prior to DArt$_{bio}$, aggregation within KOMET was performed in a straightforward way broadly analogous to that described in McKeown, Robin, and Kukich (1995); subsequently, however, more sophisticated techniques imported from visualisation were adopted (cf. Section 2.2 below).

(a) Anni Albers is American, and she is a textile designer, a draughtsman and a printmaker. She was born in Berlin on 12 June 1899. Albers studied art in 1916 - 1919 with Brandenburg. Also, she studied art at the Kunstgewerbeschule in Hamburg in 1919 - 1920 and the Bauhaus at Weimar and Dessau in 1922 - 1925 and 1925 - 1929. In 1933 Albers settled in the USA. In 1933 - 1949 she taught at Black Mountain College in North Carolina.

(b) Anni Albers is American, and she is a textile designer, a draughtsman and a printmaker. She was born in Berlin on 12 June 1899. In 1916 - 1919 Albers studied art with Brandenburg. In 1919 - 1920 she studied art at the Kunstgewerbeschule in Hamburg. In 1922 - 1925 and 1925 - 1929 she studied at the Bauhaus at Weimar and Dessau. In 1933 Albers settled in the USA. In 1933 - 1949 she taught at Black Mountain College in North Carolina.

(c) Josef Albers was American, and he was a painter, a printmaker, a sculptor, a designer, a writer and a teacher. He was born at Bottrop on 19 March 1888. He died in New Haven on 25 March 1976. He trained at the Königliche Kunstschule in Berlin in 1913 - 1915. In 1933 he went to the USA, and in 1950 - 1960 he taught at Yale University. He taught 'art' at Black Mountain College in North Carolina. He created "Untitled (Window Picture)", "Factory", "Bent Dark Grey" and "Bent Back (B)".

(d) Josef Albers was American, and he was a painter, a printmaker, a sculptor, a designer, a writer and a teacher. On 19 March 1888 born at Bottrop. On 25 March 1976 died in New Haven. Trained at the Königliche Kunstschule in Berlin in 1913 - 1915. Went to the USA in 1933 , and taught at Yale University in 1950 - 1960. Taught 'art' at Black Mountain College in North Carolina. Created "Untitled (Window Picture)", "Factory", "Bent Dark Grey" and "Bent Back (B)".

(e) Roger Fenton was an English photographer. He was born at Bury in March 1819. He died in London on 8 August 1869. He studied at University College in 1838, and he entered the studio of Delaroche. In 1841 he traveled to Paris, and in September 1855 he exhibited at the Gallery of the Water-Colour Society in London.

(f) Roger Fenton was an English photographer. In March 1819 he was born at Bury, and on 8 August 1869 he died in London.

Figure 3
Example biographies generated by the KOMET system on the basis of information maintained in the Editor's Workbench. Potential text structure and linguistic forms are given by KOMET's text type descriptions, whereas specific information content is detemined by the information actually present in the Editor's Workbench knowledge base.

of contrasting subgenres generated with respect to the same underlying information from the Editor's Workbench. Biographies (a) and (b) were produced from the information from which Figure 2 was extracted. Within (a) we see contrasting selections of thematic progression in the *Education* and *Career* generic stages: *Education* is organized around the artist, and *Career* shows temporal thematic progression; within (b), both of these stages have selected temporal thematic progression. There are also differences in the strategies adopted for reference; this was controlled by resetting 'text-new' status when beginning a new text stage. Lexical selection was also handled very simply, either by associating particular verbs with domain model concepts (e.g., 'study' for learning events, 'be' for 'nationality' attributes) or by allowing direct entries from the Art and Architecture Thesaurus or strings from the analyzed texts to be passed over for use by the generator (e.g., 'draughtsman', 'Black Mountain College'). Similar differences between closely related subgenres affecting both content and form are illustrated in biographies (c-d) and (e-f).

In summary, the process of text construction supported by KOMET provided suitable generation capabilities for the initial DArt$_{bio}$ prototype. No changes were required in the generic generation architecture and generation performance was sufficiently robust for biographies such as those of Figure 3 to be produced in real-time for any of the approximately 10,000 artists for which information had been extracted by the Editor's Workbench.

2.2 Automatic Visualization

One of the main goals for information visualization in the Editor's Workbench was to be able to present *overviews* of datasets. This functionality was supported by an approach to generative diagram design first reported in Reichenberger, Kamps, and Golovchinsky (1995). This built on the landmark work of Mackinlay (1986) in which relational data is flexibly linked with elements of a 'graphical language' by means of a data classification algorithm. The elements of the graphical language were also allocated particular degrees of 'expressiveness' in order to allow selection of just those graphical resources required to capture the data being described. Reichenberger *et al.* extended this data classification approach by employing a general type hierarchy of data properties that allows algorithmic determination of the *most specific property subtype* (e.g., transitive, acyclic directed graph, inclusion, etc.) that accurately describes the data. This property subtype then allows the selection of those particular forms of diagrammatic representation that are expressively adequate but not over-expressive for the the particular data set to be expressed. For example, inclusion can be expressed visually as a tree diagram, or as nested boxes, but does not need additional expressive resources such as arrows for directionality or size for extension.

The theoretical basis of these methods is given in detail in Kamps (1997) and Kamps (1998). They rest on a new application of Formal Concept Analysis (FCA: Wille (1982)). FCA is an applied mathematical discipline based on a formal notion of concepts and concept hierarchies and allowing the exploitation of mathematical reasoning for conceptual data analysis and processing. In particular, FCA permits the efficient construction of dependency lattices that effectively represent the functional and set-valued dependencies established among the domains of some data relation. Such dependency lattices can then motivate the differential selection of appropriate graphical presentations. This approach has come to influence many components in the DArt$_{bio}$ system, including those responsible for the natural language generation. We therefore show briefly here how such lattices are constructed, demonstrate their use in diagram design, and mention where this also has consequences for our NLG component.

FCA starts from the notion of a formal context (G, M, I) representing a data set in which G is a set of objects, M is a set of attributes and I establishes a binary relation between the two sets. $I(g, m)$ is read "object g has property m" where $g \in G$ and $m \in M$. Such a context is called a *one-valued context*. An example one-valued context corresponding to the attribute 'Profession' for a set of artists is shown in the table to the left of Figure 4. *Concepts* in FCA are defined in accordance with the traditional theory of concepts and consist of an *extension* and

	Architect	Designer	Urban Planner
Gropius	X		X
Breuer	X		X
A. Albers		X	
J. Albers			X
Moholy-Nagy			X
Hilberseimer	X		

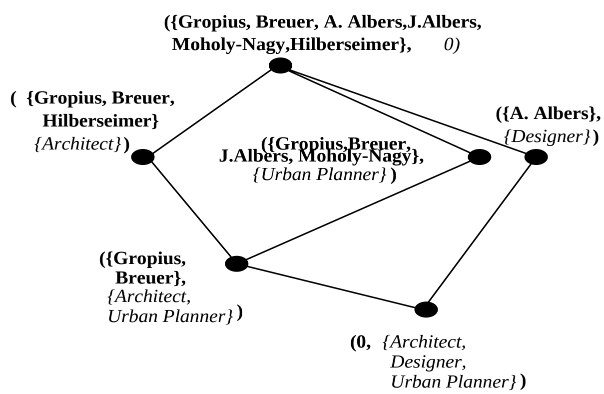


Figure 4
Example for a one-valued context and corresponding lattice

an *intension*. The extension is a subset A of the set of objects G and the intension is a subset B of the set of attributes M. We call the pair (A, B) a *formal concept* if each object of the extension has all the properties of the intension. Thus, for the data shown in Figure 4, the pair ({Gropius, Breuer}, {Urban Planner, Architect}) represents a 'formal concept': each of the members of the extension possesses all the attributes mentioned in the intension. Intuitively, when shown in a table such as that of Figure 4, the concepts of a formal context represent those rectangles of maximum size completely filled with x's after permutation of rows and columns. The set of all concepts for some formal context can be computed effectively using the "Next Closure" algorithm developed by Ganter and Wille (1996).

The main theorem of concept analysis then shows that the set of concepts for a formal context can be organized into a complete lattice structure under the following definition of the "subconcept" relation: a concept (A, B) is a subconcept of (A^*, B^*) if and only if $A \subseteq A^* \Leftrightarrow B^* \subseteq B$ (see Wille (1982)). The concept lattice may be constructed by starting from the top concept (the one that has no superconcepts) and proceeding top-down recursively. In each step we compute the set of direct subconcepts and link them to the respective superconcept until we reach the greatest lower bound of the lattice itself (the existence of the bounds is always guaranteed for finite input data structures). An efficient implementation of this algorithm is given in Kamps (1997). The lattice corresponding to our example one-valued context is shown in the right of Figure 4. This lattice shows the full labelling of formal concepts in order to ease comparison with the originating table. Much of this information is redundant, however, and so we generally use variations on the abbreviated, more concise form shown in Figure 5. Such lattices naturally capture similarities and differences between the values of the specified attributes of objects: each concept of the lattice indicates objects with some set of values in common. Moreover, the generalizations are organized by subsumption, which supports the selection of most specific subtypes.

When considering data sets in general, however, we typically need to express more information than that of single attributes and for this we require *multi-valued contexts*. An example of a multi-valued context is shown in Table 1, which includes our previous one-valued context as one of its columns. The table shows the subject areas, institutions, and time periods in which the indicated artists were active. Although expressed here in tabular form for convenience, this information is precisely analogous to the input to the generation process considered above. Formally, a multivalued context is a generalisation of a one-valued context and may be represented as a quadruple (G, M, W, I) where G, M, and I are as before and W represents the set of values of the attributes—which are, in contrast to the one-valued context case, not trivially either true or false, applicable or not. To identify the value $w \in W$ of attribute $m \in M$ for an object $g \in G$ we adopt the notation m(g) = w and read this as "attribute m of object g has value w".

Kamps (1997) renders multivalued contexts amenable to the techniques for dependency lattice construction by deriving a one-valued context that captures

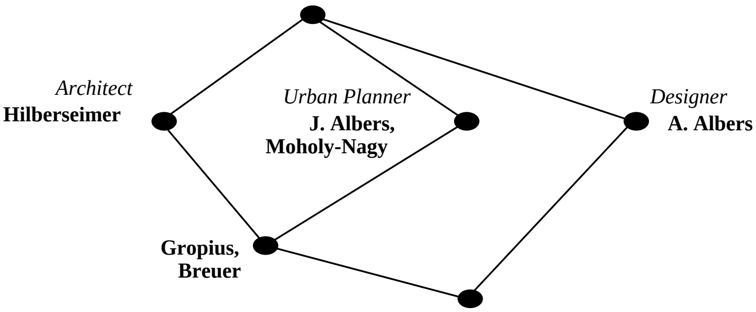


Figure 5
Concept lattice example: more succinctly labelled. Here: the extension label for each node consists of just those elements which are *added* at that node moving *up* the lattice; conversely the members of the intensions are shown moving *down* the lattice, again adding just those elements that are new for that node. For example, the node simply labelled **Gropius, Breuer** corresponds to the full form ({Gropius, Breuer}, {Architect, Urban Planner}) since both Gropius and Breuer are added new to the extension at that node, while no new elements are added to the intension—'Architect' and 'Urban Planner' are both inherited from above.

Table 1
A collection of facts concerning artists and their professions drawn from the Editor's Workbench frame-based domain model and re-expressed as a table of facts and attributes.

	Person	**Profession**	**School**	**Workperiod**
g1	Gropius	Architect	Harvard	1937–1951
g2	Breuer	Architect	Harvard	1937–1946
g3	A. Albers	Designer	Black Mountain College	1933–1949
g4	J. Albers	Urban Planner	Black Mountain College	1933–1949
g5	Moholy-Nagy	Urban Planner	New Bauhaus	1937–1938
g6	Hilberseimer	Architect	Illinois Institute of Technology	1938–1967

	Person	Profession	School	Workperiod
g1g2		X	X	
g1g6		X		
g2g6		X		
g3g4			X	X
g4g5		X		

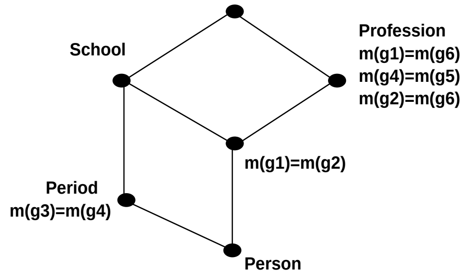


Figure 6
Example dependency context and corresponding lattice

the *functional dependencies* of the original multivalued context. To see how this works, we note first that a functional dependency in a relation table is established when the following implication is always true: for two arbitrary objects $g, h \in G$ and two domain sets $D, D^* \in M$, then $D(g) = D(h) \Rightarrow D^*(g) = D^*(h)$. This implication suggests the following construction for an appropriate one-valued dependency context: for the set of objects take the set of subsets of two elements of the given multi-valued context $P_2(G)$, for the set of attributes take the set of domains M, and for the connecting incidence relation take $I_N(\{g, h\}, m) :\Leftrightarrow m(g) = m(h)$. The required dependency context is then represented by the triple $(P_2(G), M, I_N)$. This is illustrated in the table to the left of Figure 6, which shows the one-valued context corresponding to the multi-valued context of Table 1. An entry here indicates that the identified attribute has the same value for both the facts identified in the object labels of the left-most column: for example, ‘g1’ and ‘g2’ share the values of their Profession and School attributes. This provides a wholistic view on the dependency structure of the original data and is, moreover, computationally simple to achieve.

It is then straightforward to construct a dependency lattice as described above; this is shown on the right of Figure 6. The arcs in this lattice now represent the functional dependencies between the involved domains and the equalities (e.g., m(g1)=m(g2)) represent the redundencies present in the data. For example, the lower left node labelled Period indicates not only that the third and fourth row entries under Period (g3 and g4) are identical but also, following the upward arc, that these entries are equal with respect to School; similarly, following upward arcs, the middle node (m(g1)=m(g2)) indicates that the first and second row table entries (e.g., g1 and g2) are shared with respect to both School and Profession. The lattice as a whole indicates that there are functional relationships from the set of persons into the set of professions, the set of periods, and the set of schools. A further functional relationship exists from the set of periods into the set of schools.

Once such a lattice has been constructed, we also have as a consequence a set of classifications of the original relational input, or data set. This can directly drive visualization as follows. For graphics generation it is important that all domains of the relation become graphically encoded: this means the encoding is complete. Kamps (1997) proposes a corresponding graphical encoding algo-

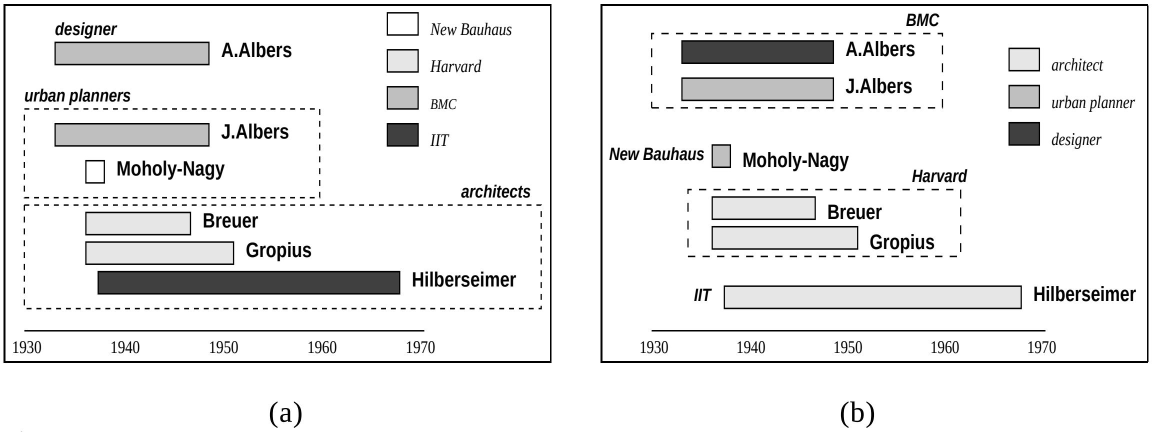


Figure 7
Example generated diagrams for the example data

rithm that starts encoding the bottom domain and walks up the lattice employing a bottom-up / left-to-right strategy for encoding the upper domains. The idea of this model, much abbreviated, is that the cardinality of the bottom domain is the largest, whereas the domains further up in the lattice contain fewer elements. Thus, the bottom domain is graphically encoded using so-called *graphical elements* (rectangle, circle, line, etc.), whereas the upper domains are encoded using *graphical attributes* (colour, width, radius) and *set-valued attributes* that must be attached to graphical elements. This strategy also keeps graphical complexity moderate.

Figure 7 shows two example diagrams that are produced from the data set of Table 1 via the dependency lattice shown to the right of Figure 6. Informally, from the lattice we can see directly that artists ('Person') can be classified on one hand according to work period (following the lefthand arc upwards) and, on the other hand, *jointly* according to school and profession (following the vertical arc). The algorithm first allocates the 'attribute' person, indicated in the lowest node of the lattice, to the basic graphical element 'rectangle'; the individual identities of the set members are given by a graphical attachment: a string giving the artist's name. The functional relationship between the set of persons and the set of time periods is then represented by the further graphical attribute of the *length* of the rectangle. This is motivated by the equivalence of the properties of temporal intervals in the data and the properties of the graphical relationship of spatial 'intervals' on the page. Two paths are then open: first following the functional relationship to a set of schools or to a set of professions. Diagram (a) in Figure 7 adopts the first path and encodes the school relationship by means of the further graphical attribute of the *color* of the rectangle, followed by a nesting rectangle for the relationship to professions; diagram (b) illustrates the second path, in which the selection of graphical encodings is reversed. Both the selection of color and of nesting rectangles are again motivated by the correspondence between the formal properties of the graphical relations and those of the dependencies observed in the data.

Finally here, we note that the work that a dependency lattice performs for organizing a data set for graphical presentation also corresponds to the work needed when *aggregating* information for expression in natural language. For example, the generalization captured in Figure 7(a) that both J.Albers and Moholy-Nagy are Urban planners corresponds precisely to an aggregation of facts $g4$ and $g5$ in Table 1 and would underlie aggregated nominal expressions such as ‘The urban planners J.Albers and Mohology-Nagy’ or aggregated clauses such as ‘J.Albers and Moholy-Nagy were urban planners’. We have discussed this equivalence of classification for graphical presentation and aggregation for NLG in more detail in Bateman et al. (1998). Below we will generally consider dependency lattices as introduced here also as *aggregation lattices* that indicate groupings that may be exploited both for concise graphical presentation and for linguistic aggregation.

2.3 Automatic Page Layout

Page layout, more properly termed ‘typographic design’, is usually divided into three levels: microtypography, macrotypography (layout proper), and style. These three levels roughly relate to the traditional division of labor in design tasks: While microtypographic problems were the domain of a typesetter, the layout was done by a designer, often along stylistic guidelines from an art director. Most of the functions of typography, that go beyond supporting linear reading, i.e. enabling selective reading or looking up information (for a more detailed description see Willberg and Forssman (1997)), are implemented through macrotypographic means of expression: For example, text may be formatted as a footnote or margin-note in order to make clear that it can be skipped without loss of coherence. Another example is a strong separation of sections combined with an emphasis on their headings in order to establish each of them as an alternative entry point for reading an extended text. Macrotypography is also to a large extent responsible for the interpretation of the text: information such as two texts coming from different sources, or one text introducing, commenting or elaborating the other, is commonly expressed in their layout. Dynamically composed documents require autonomous design decisions and to motivate such decisions appropriately it is necessary to model a certain amount of layout-intelligence: we need to know both what layout *is* and what it *does*.

In general we treat layouts as just another kind of diagram. The automatic page layout designed for the Editor’s Workbench is then implemented as a specialization of the general visualization task just described. For the graphical properties we took as starting point a restricted set of general relation types developed by Southall (1992). These include containment, reading order, similarity and reference, and capture the expressive power of layout / typographical design without connecting them to the document content or with specific typographical means. We call this level of description the *Layout Structure*; it is described in more detail in Reichenberger et al. (1996). Layout structures specify: (a) the constituency structure required, (b) the strength of ‘connectedness’ between elements at any depth in the structure (in terms of informational dependence and desired reading order), and (c) the relative 'importance' of individual elements. Given this as input, the automatic visualization component renders it as a restricted form of diagram: Constituency is mapped to nested boxes (i.e., inclusion diagrams), and strength of connection and sequence is mapped to spatial displacement—the boxes included within an enclosing box will also in general be arrayed two-dimensionally in order to influence reading order—and typographic attributes—such as type size, specific type face within the family (bold, italic etc.), arrangement of the type (ragged right, flush matter etc.) and (ever more frequently nowadays) coloring and orientation.

A simple example of a layout structure is shown to the left of Figure 8. Here we can see that in addition to hierarchical structure there are three kinds of links possible between layout units: *type equivalence*, which has the consequence that two or more units are treated identically as far as layout decisions are concerned, *reference*, which indicates the material in one unit makes reference to, or uses, the material in another, and *follows*, which expresses desired relative reading order. The structural links themselves are annotated to show both the amount of material that they contain (e.g., unit 2.3.2's '403 words and 3 pictures') and their relative importance at that level of structure (expressed as a percentage). Combining this information allows further typographical contraints to be derived—e.g., the number of words and importance constrains the size of font. The visualization process then attempts to resolve these combined constraints in order to place the units spatially within a page segment. One such rendering of the layout units is indicated to the right of the figure. Relative importance has been rendered by relative size, the reference relation between unit 2.3.1 and the pictures has been rendered by insetting, and these three units as a block are placed higher on the page to affect reading order.

In order to determine layout structures, it is necessary to refine our view of the *function* of layout. This establishes layout as a class of diagrams with their own particular properties and concerns. Most crucially: they serve to express the *construction of a document*. Thus layout as such does not communicate domain knowledge but rather information about how to read the document and how its utterances relate to each other. Layout accordingly has a number of special features when considered as a form of diagram, i.e., a diagram of the structure of a text, that distinguish it from other kinds of diagrams: In particular, text blocks do not stand as signs for facts or data, but are themselves the significant objects: in contrast to other diagrams, therefore, the question of what is to be realized as graphical element or as graphical attribute does not arise. This iconicity between text structure and layout has as a consequence that, in actuality, the particular allocations of *layout* resources are scarcely interpreted in their own right: visual grouping is interpreted directly as informational content grouping. The fundamental maxim here is that the same layout properties indicate the same content properties, and that different layout properties indicate different content properties. Layout therefore only avails itself of a restricted range of presentation possibilities when compared with the full set available for diagrams in general but requires particular attention to be paid to the intended document structure

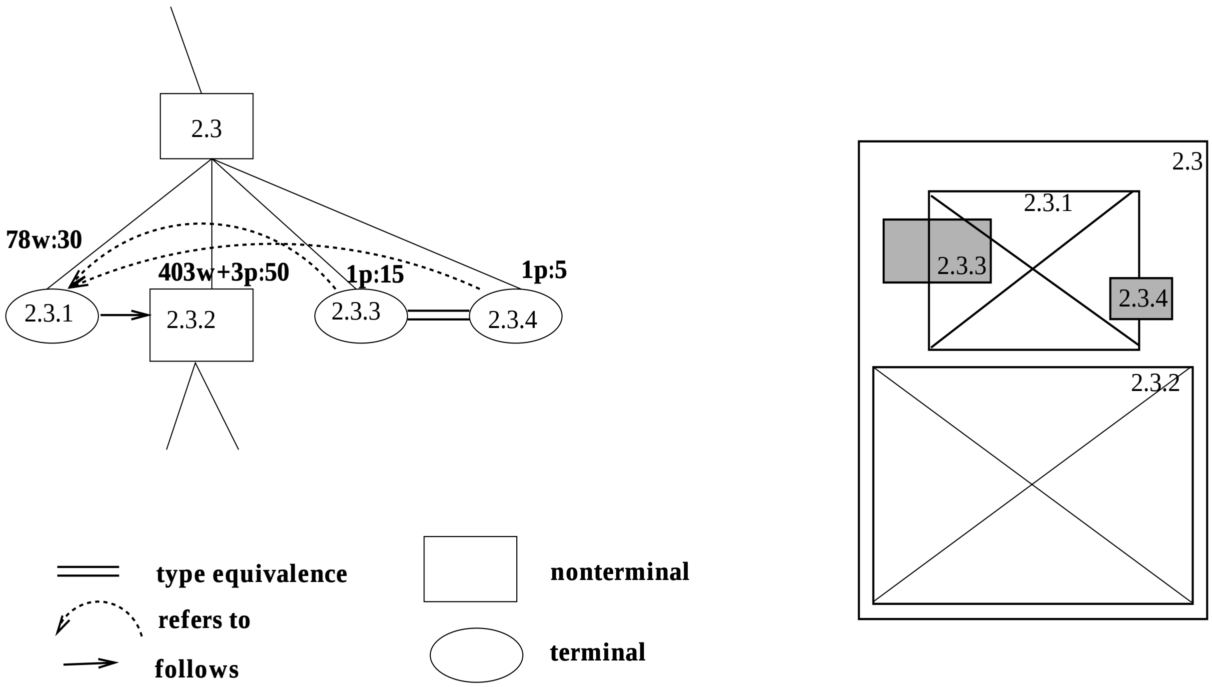


Figure 8
Example layout structure and its correspondence to a segment of page layout (adapted from Reichenberger et al. 1996).

being presented. Characterizing this relationship as the motivation for layout structure thus formed one fundamental starting point for the design of DArt$_{bio}$.

3 Communicative-functional page layout: a methodology for empirical investigation

The acknowledged iconicity between layout and 'document construction' indicated above entails that layout carries important information concerning the 'meaning' of a page: information presented together, or in similar styles, is perceived as related regardless of whether this was intended by the page designer or not. Failing to respect such entailments makes a page or diagram difficult to interpret and possibly misleading. But not all of the possible details of document structure are normally presented in layout: the relationship is substantially more flexible. Moreover, previous systems for multimodal presentation have assumed a rather straightforward view of this relationship which, as we shall see in this section, does not generalize well to cover more sophisticated layout. As an initial preparatory step prior to building the DArt$_{bio}$ prototype, therefore, we set out explicitly to investigate the possible interrelationships between page layout and communicatively motivated text structure more closely. For this, we developed a methodology for exploring the functional basis of page layout. Two caveats here: first, our experimental method was exploratory: as one of the first studies of its kind, we needed to respond flexibly to the results of analysis; and second, since our aim was to move quickly from first analysis results to prototype system in order to evaluate the feasibility and value of the entire scenario at work, the study was deliberately restricted in scope.

3.1 Analysis method

The provision of appropriate multimedia corpora for supporting the kind of analysis we required is still, several years later, very much in its infancy.[6] The main criteria for the selection of pages for our investigation were (a) that the entire page be concerned with a single 'topic', while nevertheless presenting various aspects of that topic by means of varied text structuring and typographical and layout decisions, and (b) that the page demonstrated 'interesting' layout. This led us to consider principally pages drawn from popular magazines since these exhibit very varied typographical and layout decisions in the hope of being eye-catching and interest-maintaining. The detailed structure of our study was as follows. We took a set of pages selected according to criteria (a) and (b), and asked for each page why it was set out as it was. We answered this question by:

1. providing a layout structure representation for each page (cf. Section 2.3),

6 Two new projects currently involved with gathering and annotating such data are ICONOCLAST and GEM: see Section 6. Corio and Lapalme (1998) also present some results of a corpus study.

2. constructing a single 'text' out of the entire 'content' of the page (including headings, picture captions, etc.) that captured as far as possible the perceived purpose of the page,

3. performing a 'functional text structure analysis' of the constructed text respecting the perceived purpose of the page (and therefore of the constructed text),

4. considering whether the page layout structure represented *natural divisions* with respect to the text structure.

We then successively refined our notion of 'natural divisions' by generating (initially by hand, later augmented by the results of the running DArt_{bio} prototype) alternative page layouts of the single constructed text and informally evaluating the result for its perceived effectiveness. To the extent that we judged our alternative layouts not to be effective, this indicated that our understanding of a natural division was still insufficient and needed further investigation. After several rounds of refinement, we had sufficient hypotheses to motivate the first implementation of communicatively-effective layout that formed the basis of the page presentation algorithms of DArt_{bio}.

3.2 An illustrative round of analysis: preliminaries

We illustrate our successive approximation approach by setting out in detail one round of analysis concerning a particular illustrated page. This shows concretely both the approach and the motivation for the initial algorithm adopted for effective layout. Since considerable information is often expressed in a single page of information, space precludes more than a few examples of alternatives even with this single example. The page used is shown in Figure 9 and is drawn from the German illustrated sport and health magazine *Fit for Fun* (1995, Number 5, p92). The article describes various aspects of the game 'Unihoc', presenting background information concerning how the game is played, where it is popular, why it is popular, and some pointers to further information and the equipment needed to play it. The page was typical of a particular series of feature articles being run in the magazine at that time. Particular elements expressed using layout in the page are indicated in the figure: significant here is soley the physical layout of the page, not the content of the article.

Our starting point for analysis was motivated by our initial page selection criteria, i.e., that there is some body of material that the author(s) of the page wish to present. It is then relevant (and necessary) to ask how this information is to be broken up for effective presentation. To do this, we set out the content of each page as a single constructed text: The constructed text for the Unihoc page is given in Figure 10. Here we number independent clauses (or major information units) for ease of reference and also summarize the graphical content of the page. Note that this text already reflects our understanding of the function of the page in its context of use: we assume that the page had the main functions of informing readers about a game that they might not have been familiar with

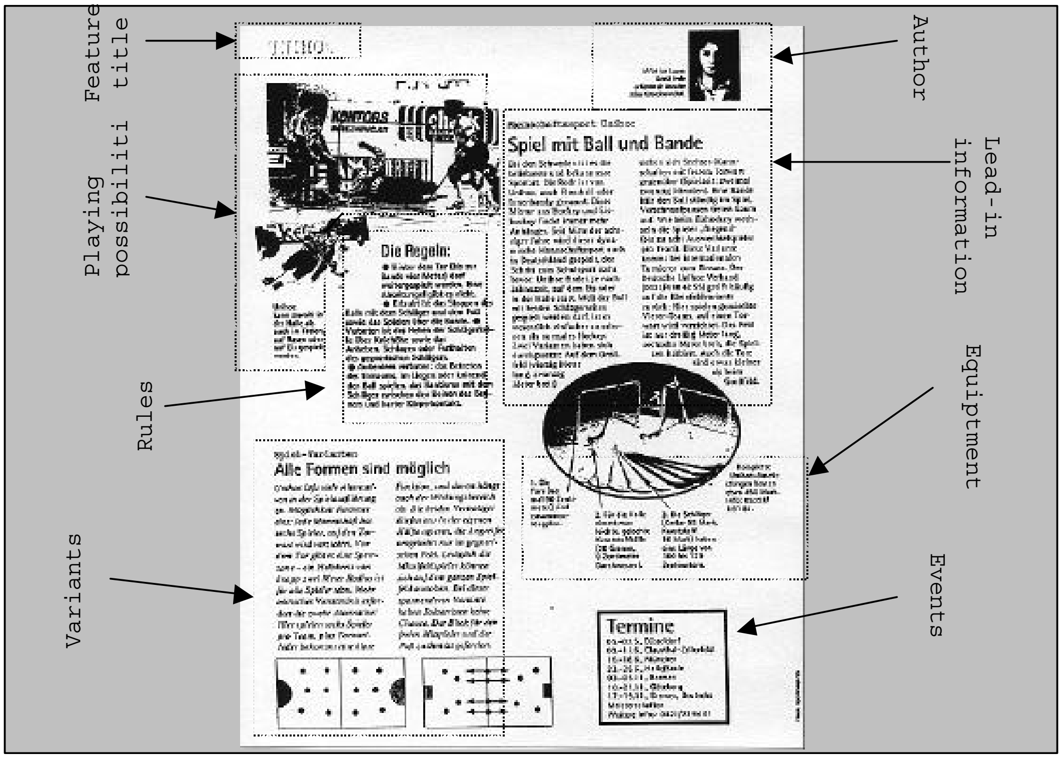


Figure 9
An original page for analysis. Source: 'Fit for Fun' **5**: p92 (1995). The page is annotated here to show the major layout units.

and telling them that the game is in fact finding increasing support both internationally and nationally. The page also shows something of why this might be the case (e.g., it can be played everywhere, allows many different variations, and is much easier than the superficially similar hockey), what you need to play the game, and where to get further information.

This 'pre-analytic' understanding of the communicative function of the page is then further refined by a detailed functional text structure analysis. For this, we adopted 'standard' rhetorical structure theory (RST) as set out in Mann and Thompson (1986). RST seeks to describe the structure of any text in terms of rhetorical relations which hold between the segments of the text. It is a functional theory in that the segments related are functional rather than textual—i.e. a rhetorical relation does not need to have any specific grammatical or lexical realisation. Rhetorical relations are either symmetric (*multinuclear*), in which case the text 'spans' related are considered of equal importance for the text, or assymetric, in which case one text segment among those related by a relation is singled out as being more essential to the writer's purposes than the others. The segment considered most central is called the nucleus, the other segments are called satellites; a nucleus cannot be ommitted from the text without endangering the text's coherence, whereas a satellite can. Segments may themselves have further RST structure, thus building up recursive structures. RST was selected as one of the most elaborated and widely used forms of text analysis that would be capable of providing a detailed communicative-functional decomposition of our selected texts. In principle, we would expect any similarly detailed approach to yield broadly similar results; RST had the additional advantage, however, of already being well established in NLG (cf. (Moore and Paris, 1993; Hovy, 1993)) and has been applied to multimodal information presentation (cf. (André and Rist, 1993)).

Our analysis of the constructed Unihoc text is shown in Figure 11: text segments are represented by horizontal lines, nucleii by vertical lines, and satellites by segments linked to a nuclear segment by a labelled arc. Although space precludes repeating the definitions of the RST relations found in the analysis, the definitions employed are exactly those as given by Mann and Thompson. The analysis in the figure also includes the information presented in the original page as photographs or diagrams. These have been labelled alphabetically (A, B, etc.) as identified in Figure 10 and anchored into the RST-tree at appropriate places with plausible relations.[7]

The RST analysis makes our interpretation of the function of the text very explicit. The central nucleus for the page as a whole is unit [3], i.e., *This mixture of hockey and ice hockey is attracting ever more supporters.* We are therefore considering the primary purpose of the page to be a statement that Unihoc is becoming

7 The linking of graphical material into the RST-structure is not without theoretical implications: the question of *nuclearity* is often particularly unclear: is the diagram supporting the text or vice versa? It is perhaps better to admit a co-dependency relationship where each supports the other, but in different ways. However, since this would have required a change to the 'standard' RST position, we did not adopt this for the stage of analysis reported in the present paper.

[0] Astrid Frula, captain of the German National Unihoc team, writes: [1] Among the Swedes it is the most popular and best-known branch of sport. [2] We are talking about Unihoc, also called Floorball or Indoor Bandy. [3] This mixture of hockey and ice hockey is attracting ever more supporters. [4a] Since the middle of the eighties, this dynamic team sport has also been played in Germany and [4b] the step to becoming a school sport is imminent. [5] Unihoc can be played in the gym as well as outside, on grass or ice. [6a] Because the ball can be played with both sides of the stick, [6b] it is much easier to master than normal hockey. [7] One can continue playing behind the goal (four metres up to the board) and [8] there is no offside rule. [9] Stopping the ball with the stick and the foot is allowed, as well as playing via the board. [10] Not allowed is raising the blade of the stick above knee height, or lifting, hitting and holding the opponent's stick. [11] Nor is it allowed to enter the goal area, to play the ball while lying or kneeling, to move the stick between the legs of the opponent, and to engage in hard body contact. [12] Unihoc allows many alternatives in how it is played. [13] One possibility: [14] each team has six players, and no goalie. [15a] In front of the goal there is a no-go area - [15b] no players are allowed within a semicircle of almost 2 meters radius. [16] The second alternative requires more tactical insight: [17] here there are 6 players per team, plus a goalie. [18] Each receives a clear function, which determines their effective playing area. [19a] The two defenders may only act within their own half; [19b] in contrast, the two attackers may only play within their opponents' half. [20] Only the midfield players can run as they wish over the entire playing field. [21] In this exciting variant solo artists have no chance. [22] Spotting the free team member and passing the ball on are essential. [23] Two variants have become dominant. [24] On the large field (forty metres long, twenty metres wide) two six-person teams with fixed goalie oppose each other (Playing time: two times twenty minutes). [25a] A board keeps the ball continuously in play; [25b] rest periods hardly ever occur. [26] As in ice hockey a player substituting another does not lead to an interruption of play (up to eight substitute players per team). [27] This is the variant that is used in international matches. [28] The German Unihoc Union (0421 / 4984255) frequently goes back to the small field variant: [29] where mixed 4-person teams play without a goalie. [30] The playing field is only thirty metres long and sixteen metres wide, while the playing time is halved. [31] The goals are also smaller than on the large field. [32] The goals (60 $\times$ 90 centimeters) are collapsible. [33] In the gym light holed plastic balls (20 grams, 8 centimeter diameter) are used. [34] The sticks (Kevlar 95 Mark, plastic 10 Mark) are 100 to 120 centimeters in length. [35] Complete sets of Unihoc equipment cost around 450 Marks. [36] Info: 05357 / 18181. [37] 05.-07.5., Düsseldorf, [38] 09.-11.6., Clausthal-Zellerfeld, [39] 16.-18.6., München, [40] 23.-25.6., Halle / Saale, [41] 03.-05.11., Bremen, [42] 10.-21.11., Göteborg, [43] 17.-19.11., Bremen, Deutsche Meisterschaften. [44] Further Info: 0421 / 23 94 01.

Additional graphical material:

A: Astrid Frula (photograph): 'authorial voice'

B: Player positions (diagram): two variations, B1 and B2

C: Unihoc equipment (photograph)

D: Unihoc being played on ice (photograph)

E: Unihoc being played in the gym (photograph)

Figure 10
Constructed Unihoc text and graphical material used

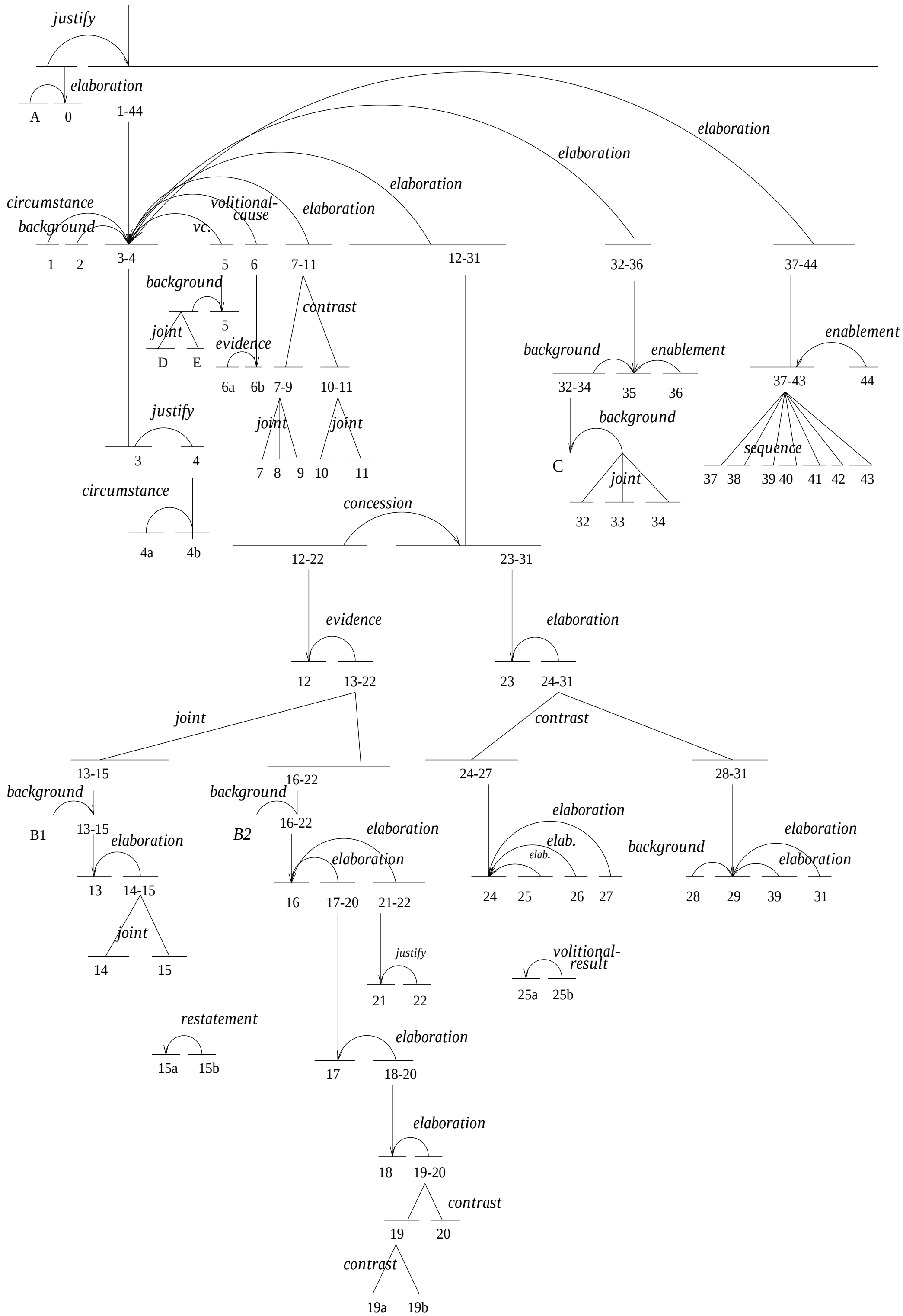


Figure 11
Rhetorical structure theory analysis of the constructed Unihoc text

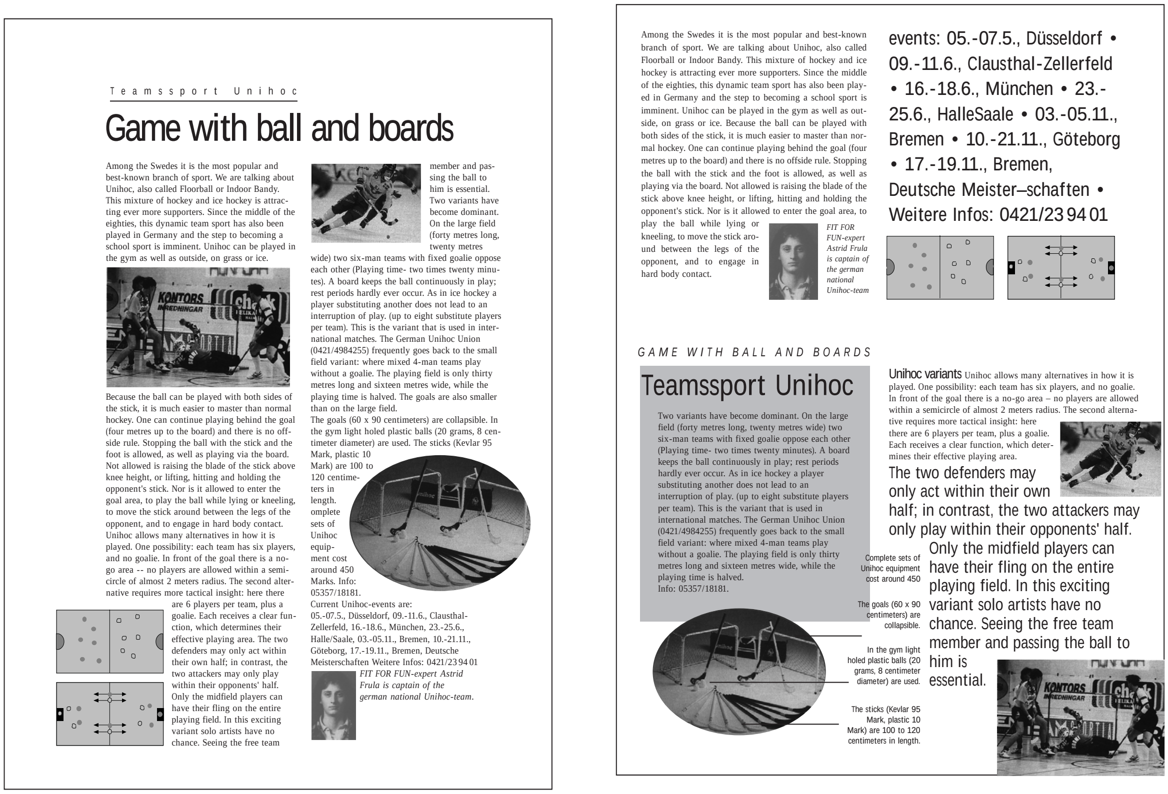
very popular (and so the reader should be well up on it). The segments immediately following the nuclear span, [5] and [6], give some of the reasons why the sport is becoming so popular (with rhetorical relation 'volitional cause'), and then segment [7–11] gives an overview of the do's and don'ts of the game. The main bulk of the constructed text consists of a concession that, although *Unihoc allows many alternatives in how it is played* [12], *two variants have become dominant* [23]. The existence of alternatives is supported by the two possibilities presented in segments [13–15] and [16–22], both related by the relation *evidence* to the nuclear [12]; the two main alternatives are *elaborate*d in the explicit *contrast* drawn in segments [24–27] and [28–31]. Finally, the segments [32–36] and [37–44] provide additional elaborating material concerning where and when the game can be seen and what equipment is necessary to play it.

Although it is possible that alternative analyses of the page be found, with different interpretations of the intended function of the text, it is less likely that any such alternatives will present radically differing structure. In particular, even though the selection of particular relations may vary, our focus here is on the *general constituency and dependency structures* set up by the analysis rather than on precise details of the relations selected. This has now been shown to be considerably less subject to individual variation (cf., e.g., (den Ouden et al., 1998; **?**)) and certainly provided an adequate basis for empirical investigation.

3.3 Alternative renderings of the constructed text and their evaluation

We then considered how the communicative-functional intent represented in such a rhetorical structure analysis can be rendered in a page layout. Our initial hypothesis was that, since the RST analysis represents a statement of the varying degrees of centrality attributed to the text segments present on the page, then these nuclearity assignments should also be reflected in the organization of any page layouts selected. Moreover, we assumed that the units found in the layout structure should correspond only to proper subtrees within the RST structure and that elements grouped in the layout should also be grouped within the RST. Our first approximation to 'natural divisions' was, therefore, that these should be drawn as consistent sub-trees from the RST structure. Our investigation then evaluated this approximation by considering successively more complex 'possible' layouts.

First, for example, we can note that if the aim of the author / editor / multimodal presentation system were *not* to present an interesting page design, then a layout such at that shown in Figure 12(a) might suffice. This presentation contains no layout decisions that subdivide the text into segments or establish relations of similarity and difference among those segments. The only subdivisions are the heading and textbody division; the pictures of the original have been inset into the main text block approximately where their content is touched upon in the text. Nevertheless, this is, we would argue, a perfectly possible (if, by current tastes, dull) rendering of the material to be presented. It could be appropriate, for example, in an extremely densely presented lexicon or encyclopedia where space constraints and tradition suppress layout variation. This layout therefore



(a) (b)

Figure 12
Contrasting page layouts decomposing the source text: (a) Page rendition with minimal layout decisions, (b) Page rendition with random layout decisions.

serves to represent one endpoint in a continuum of possible layouts that need to be accomodated in any general account. It is also indicative of the fact that there is generally a trade-off between information that is expressed linguistically through explicit textual realization and information that is carried by the layout: our constructed Unihoc text needs to do more explicit linguistic signalling of discourse relations and communicative function than does the version employing layout. A reader should be able to recover this information from reading the text but it is not supported by an explicit layout encoding; we will return to this issue at several points below.

We consider this case to emphasize that layout is concerned with *choice*. This is very similar to the state of affairs in NLG: the main principle is that a speaker / writer has to choose *how to present information* and, whenever there is choice, there is meaning: that is, the choice is not free and some choices will be more appropriate than others in particular contexts. Moreover, the layout decomposition that is selected should in some way be ‘coherent’ with respect to the communicative functions of the page. The next example layout, shown in

Table 2
Allocation of information to the layout elements in the 'random' layout of Figure 12(b)

page layout element	*text segments*	*typographical discriminations*
intro+rules	[1]–[11]	neutral
variants 1	[12]–[18]	neutral
variants 2	[19a]–[22]	larger typeface
main body	[23]–[31]	shaded background
equipment info	[32]–[36]	smaller typeface
events	[37]–[44]	very large typeface

Figure 12(b), illustrates this by presenting a layout in which choices, we would argue, have *not* been made coherently. The unmotivated decisions that make this layout incomprehensible include:[8] the "Unihoc variants" section breaks down into two parts of differing visual appearance, additional information (events) becomes excessively important due to its prominent position and font size, and the information about the author is related only to the first textblock and not to the entire article. Such problems are identical to those arising in the NLG task: when more flexibility of expression / presentation is made available—for example, by considering generation with respect to a grammar with broader coverage—it is essential to control this flexibility appropriately in order to avoid wrong decisions.

The problematic nature of the layout of Figure 12(b) can be succinctly stated: it violates our initial hypothesis concerning the desired correspondence between RST and layout structure. The 'constituency structure' of the layout and that of the RST are in direct opposition at several points and, in general, we found that such layout structures were indeed difficult to interpret coherently. The (lack of) correspondence is summarized in Table 2, in which we label layout units according to their content or function, identify the RST span they express, and summarize the typographical decisions taken to render the layout unit on the page. Thus, for example, the first element, 'intro+rules', is made up of material drawn from text units [1]–[11] (cf. Figure 10) and uses a 'neutral' (for this page model) type face; similarly, the first variants element, 'variants 1', is drawn from text units [12]–[18] and adopts the same type face. Both of these elements are presented less prominently than the larger typeface of unit 'variants 2', which is itself less prominent than the very large typeface of the 'events' element. Compared to the RST-structure, it is then curious that the 'intro' element is presented

8 It is perhaps worth clarifying here what is meant by an 'incomprehensible' layout: we mean particularly that it is not possible to reconstruct the communicative functions that we attributed to the original page. It is of course possible to provide some other interpretation of the page, but this does not then constitute an 'alternative' layout: it is instead a layout of some body of information for other communicative goals which may or may not be coherent with respect to the information actually presented.

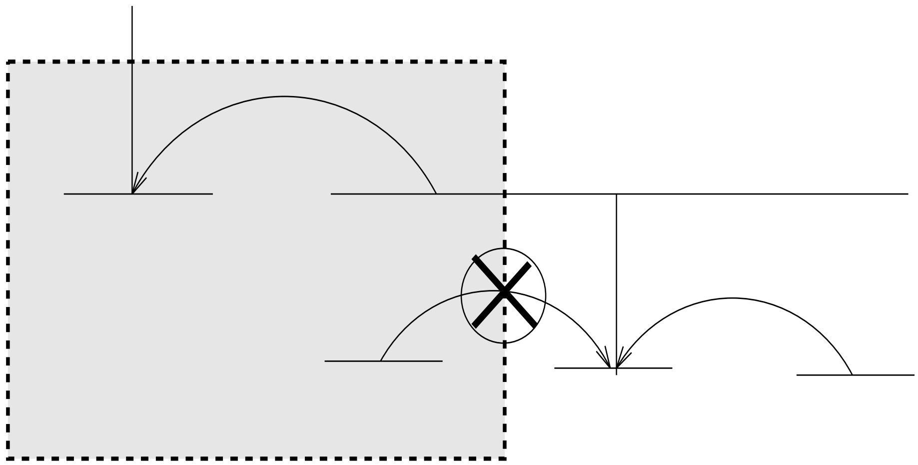


Figure 13
Example of a ‘non-natural’ division of an RST structure

with no particular prominence despite the fact that it contains the nucleus for the text as a whole [3], whereas layout element ‘events’, which expresses material that is non-nuclear in the RST structure, is highly prominent in the layout structure. More seriously, we can also note that some layout elements in their entirety find little motivation in the RST structure: for example, the division into ‘variants 1’ and ‘variants 2’ involves segments which do not correspond to RST-subtrees—indeed, segment [19a]–[22] goes further and breaks the RST structure at two points: segment [19a]–[20] is related by elaboration to unit [18] while segment [21]–[22] is related, also by elaboration, to unit [16]. The combined segment, [19a]–[22], is therefore composed of two completely disjoint and unrelated parts of the RST tree.[9] We show this graphically in Figure 13.

The simplest strategy for producing appropriate layout is therefore to restrict layout decomposition and constituency to the ‘natural divisions’ established in the RST analysis. Any subtree is then, at least in principle, a candidate for selection as a layout unit. Accordingly, we suggest that the constructed layout shown in Figure 14 does fulfil the tasks of rendering the communicative intentions of the original page quite well. A relatively large number of layout decisions have been taken—for instance, the most important statements form a block of their own at the top of the page, additional peripheral information is placed in a vertical grey margin bar, and the main text is divided in two sections: “The rules” and “Unihoc variants”. Despite this diversity, the page remains coherent by virtue of the congruence of its layout structure with the RST structure; this is summarized in Table 3. Here the ‘lead-in’ segment is broken out from the rest of the structure at the point in the RST tree where the long string of volitional means / elaboration satellites of the main nucleus [3–4] shift from being single propositions to more complex substructures; it also receives a larger

9 Interestingly, this configuration also appears to have particular *linguistic* consequences also: cf. **?**) and **?**).

UNIHOC

FIT FOR FUN-expert Astrid Frula is captain of the german national Unihoc-team

Among the Swedes it is the most popular and best-known branch of sport. We are talking about Unihoc, also called Floorball or Indoor Bandy. This mixture of hockey and ice hockey is attracting ever more supporters. Since the middle of the eighties, this dynamic team sport has also been played in Germany and the step to becoming a school sport is imminent. Unihoc can be played in the gym as well as outside, on grass or ice.

Teamssport Unihoc

Game with ball and boards

EVENTS

05.-07.5., Düsseldorf

09.-11.6., Clausthal-Zellerf.

16.-18.6., München

23.-25.6., Halle/Saale

03.-05.11., Bremen

10.-21.11., Göteborg

17.-19.11., Bremen,

Deutsche Meisterschaften

Infos: 0421/23 94 01

The rules Because the ball can be played with both sides of the stick, it is much easier to master than normal hockey. One can continue playing behind the goal (four metres up to the board) and there is no offside rule. Stopping the ball with the stick and the foot is allowed, as well as playing via the board. Not allowed is raising the blade of the stick above knee height, or lifting, hitting and holding the opponent's stick. Nor is it allowed to enter the goal area, to play the ball while lying or kneeling, to move the stick around between the legs of the opponent, and to engage in hard body contact.

Unihoc variants Unihoc allows many alternatives in how it is played. One possibility: each team has six players, and no goalie. In front of the goal there is a no-go area -- no players are allowed within a semicircle of almost 2 meters radius. The second alternative requires more tactical insight: here there are 6 players per team, plus a goalie. Each receives a clear function, which determines their effective playing area. The two defenders may only act within their own half; in contrast, the two attackers may only play within their opponents' half. Only the midfield players can have their fling on the entire playing field. In this exciting variant solo artists have no chance. Seeing the free team member and passing the ball to him is essential. Two variants have become dominant. On the large field (forty metres long, twenty metres wide) two six-man teams with fixed goalie oppose each other (Playing time- two times twenty minutes). A board keeps the ball continuously in play; rest periods hardly ever occur. As in ice hockey a player substituting another does not lead to an interruption of play. (up to eight substitute players per team). This is the variant that is used in international matches. The German Unihoc Union (0421/4984255) frequently goes back to the small field variant: where mixed 4-man teams play without a goalie. The playing field is only thirty metres long and sixteen metres wide, while the playing time is halved. The goals are also smaller than on the large field.

Complete sets of Unihoc equipment cost around 450

The goals (60 x 90 centimeters) are collapsible.

In the gym light holed plastic balls (20 grams, 8 centimeter diameter) are used.

The sticks (Kevlar 95 Mark, plastic 10 Mark) are 100 to 120 centimeters in length.

Figure 14
An example constructed page layout respecting natural divisions in the rhetorical structure

Table 3
Distribution of information for a simple coherent layout

page layout unit	*text segments*	*discriminations*
lead-in	[1]–[5]	larger type face
rules	[6a]–[11]	neutral
variants	[12]–[31]	neutral
equipment info	[32]–[36]	small, sans serif, on margin bar
events	[37]–[44]	small, sans serif, on margin bar
authorship	[A+0]	small, italic, on margin bar

type face and a prominent page position appropriate for its relative nuclearity. Both ‘equipment’ and ‘events’ are minor satellites in the RST structure and so are placed in a non-prominent position with a smaller typeface. Both the large central textblocks ‘rules’ and ‘variants’ correspond to large subtrees (the former combining [6] and [7–11]; the latter the large elaboration under [12–31]). These textblocks are not distinguished from each other, but are distinguished from all the other textblocks.

While this rendering of the RST structure is perhaps acceptable as a simple layout, we must observe that it still does not approach the complexity and diversity of natural layouts (i.e., those produced by people). The layout decisions taken in the original Unihoc page as shown in Figure 9 above, for example, include both significantly more variety in the layout and typeface resources allocated and a greater degree of decomposition. Moreover, although most of the design decisions appear coherent with respect to the notion of natural division developed so far, it is not the case that all of the layout decompositions are covered—in fact, substantially *more* decompositions and discriminations are being made. Something of this complexity is indicated in the correpondence table given in Table 4; here we provide more detail concerning the typeface discriminations and include style of formatting, too.[10] The layout decomposition is also contrasted with the RST structure graphically in Figure 15.

Here it is clear that the notion of natural division corresponding simply to RST-subtrees is by no means sufficient. In fact, there were many cases where more than a simple RST-subtree has been grouped together as a single layout unit; for example, the ‘Intro’-segment of the Unihoc page consists of two distinct RST-subtrees that are not simply linked in the RST structure (i.e., [1]–[6] and [23]–[31]). It is also clear that particular information in the RST-structure

10 For the purposes of the present paper, we will continue to use informal descriptions of the typefaces and formatting options taken up; however, with the increased level of complexity begun here, standard typographical terminology is more appropriate.

Table 4
Distribution of information and styles for the original layout. Here: 'flowing' with respect to a bullet or other list item indicates that items are run on within single lines and do not form separate paragraphs; relative font sizes are indicated by '+'. All text units are left-justified and right-ragged.

page layout unit	*text segments*	*discriminations*		
		typeface	*formatting*	*size*
Intro	[1]–[6] + [23]–[31]	neutral	2-column	neutral
Rules	[7]–[11]	bold	bullet-list, flowing, wrapping picture	neutral
Variants	[12]–[22]	italics	2-column	neutral
Equipment	[32]–[36]	bold	enumeration + summary, wrapping picture, arrow links	small
Events	[37]–[44]	sans serif	enumeration-by-date + trailer, separate items, boxed	neutral
Author	[0]	italic	neutral	smaller++
Caption for Intro	[1]–[6] + [23]–[31]	bold	typewriter	larger
Caption for Rules	[7]–[11]	narrow	typewriter	larger
Caption for Events	[37]–[44]	bold	further distinct face	larger+
Page Title	[0–44]	hollow		larger++
Caption for D and E	[5]	bold	caption	small

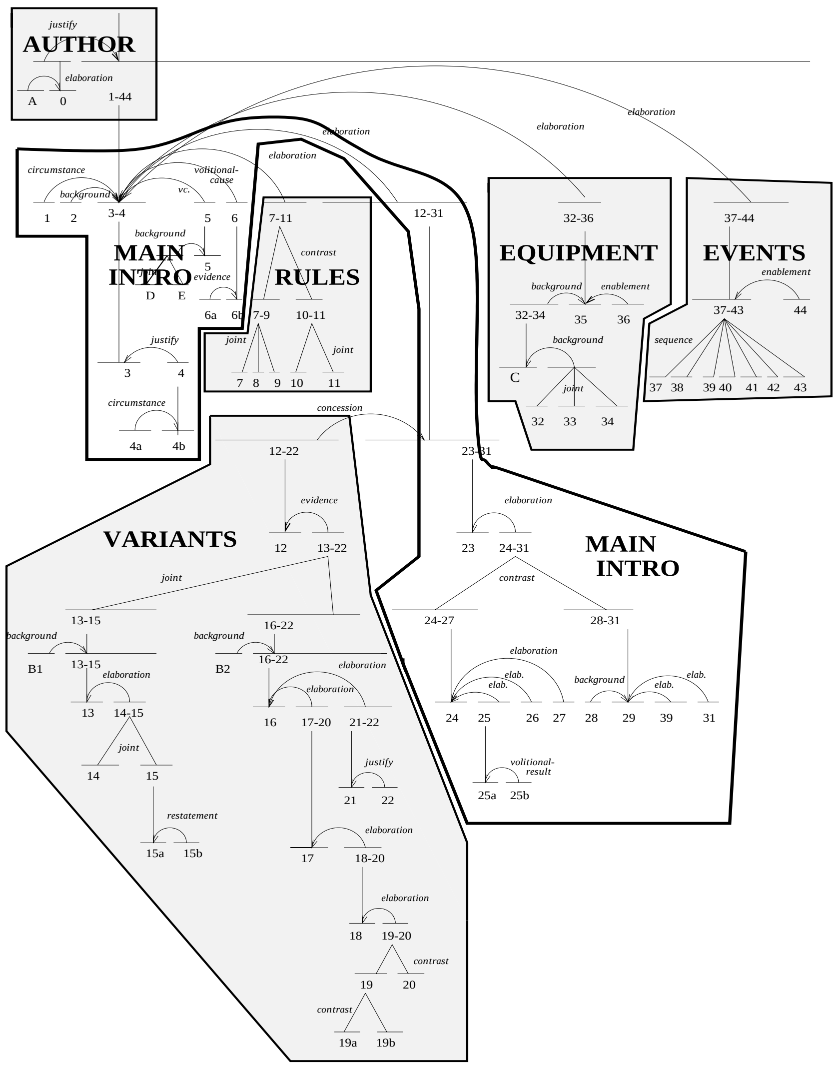


Figure 15
Correspondence between layout units and rhetorical structure in the original page
.

can appear in various guises in the final page. For example, the diagrammatic representation of the two Unihoc game varieties is also presented as text, while segment [5] appears not only as part of the 'Intro' segment but also as the caption to photographs D and E. Taken together, these considerations showed us that the rendering of the information in the RST-structure is not a simple decomposition, but is itself a result of planning activity. Information is taken from the RST-structure for various purposes and, as a consequence, both differing degrees of detail and varying decompositions are appropriate.

By examining distinct page layouts and considering the relationships between RST analyses and alternative layouts as set out here, we continued to refine our notions of what constitutes an appropriate decomposition of the RST structure. Then, for the DArt$_{bio}$ prototype, we summarized our observations as a constructive process for transforming RST structures into layout structures with allocated typographical resources. This transformation takes place in three stages, which we describe in the following section.

4 Towards automatic page layout: a specification for communicatively motivated layout

We have shown that an RST-analysis of the desired content for a page can be used to argue that a layout structure is more or less appropriate, and can indicate possible points of decomposition into layout units. Our decomposition process currently consists of three stages. First a layout structure consisting of hierarchically arranged layout *units* is derived from the RST-structure. Second, the individual layout units are allocated layout *styles*, which can be either graphical or textual and include allocation of those graphical attributes appropriate for layout (cf. Section 2.3). Finally, the information content of a layout unit is distributed across that unit in a manner usually dictated by the style selection. We will present the general approach for each stage, and then illustrate it in detail with respect to the Unihoc page. This brings us to the point where we can motivate the page as it appeared in the final published page layout.

4.1 Determination of Layout Units

The first stage carries out an inspection of the RST structure. This is modeled by interpreting the RST structure as a structure to be traversed. Certain paths through the structure are considered 'mainline', others are 'side-paths'.[11] Different layout options are available for these distinct presentational states. Side-paths can also vary in importance. If their information has only a limited outflow (no or little other information is dependent on it) or if it has a distinct and / or regular structure, then it is a likely candidate for a separate layout item. The content for each such layout item is composed of the summary of most important statements on the central 'theme' of the material—just as is the case within

11 This terminology is selected deliberately to bring out similarities with the treatment of textual development discussed in Paris (1993).

texts, thematic progression plays an important role in structuring the segments selected as individual layout units. These statements are found by gathering branches in the RST structure until a volume consistent with the nuclearity of the subtree has been reached. This process appears analogous to that employed in RST-based summarization (cf. (O'Donnell, 1997)) where nucleii are included in the summary along with some parameterizable set of satellites up to a pre-specified depth. The 'intro' section of the Unihoc page is a good example of this: here some satellites have been ommitted entirely (i.e., segments [5], [6], [7–22], etc.), while another has had only its nucleus included—i.e., within satellite [12–31] only the nuclear component [23–31] is included.

Any multinuclear relation that is encountered while traversing the RST-tree is first considered as a potential source of aggregation. The information content of the nucleii of the related segments are examined in order to see if they share significant regularities or redundancies that would support an additional layout unit consisting of a summary of what is being expressed. As we illustrate in more detail for DArt_{bio}, aggregation is performed as described in Section 2.2 above: that is, the facts of the nucleii being considered are treated as a data set and are subjected to FCA-analysis in order to derive a corresponding dependency lattice. Any aggregation selected may result in a restructuring of the RST-tree for the page by grouping segments that were previously distinct. The search for main- and side-lines then continues with respect to the modified tree.[12] Additional layout units—such as headings and captions—do not need to be explicitly present in the original RST-tree. Headers such as *The Rules* and *Unihoc variants* are drawn instead as summaries or single phrase restatements of the subject matter of their respective layout units and are introduced into the layout structure as dependent elements (cf. Reichenberger et al. (1996)).

Applying these general notions to the Unihoc page, therefore, we must first seek a 'main-line'. This necessarily includes the main nucleus of the text, segment [3], and some portion of the material surrounding it. Three principal considerations delimit that material:

- first, all of the information in satellites [1], [2], [5+D+E], and [6] is relationally diverse (i.e., there are a range of RST-relations employed, including circumstance, background, justify, evidence),
- second, their relationships to [3] are all 'contentful' in that they are not minimal relations such as joint or elaboration,
- and third, none of these satellite groups has a regular internal structure (contrast span [3–4] with [7–11], for example).

This makes a text span [1–6] with two pictures (D and E) a minimal candidate for the main-line. If we assume that no constraint has yet been reached on the

12 One place where this occurs in the example page is in the presentation of the 'rules' layout unit based on RST segment [7–11]. Here a more complex set of positive and negative rules has been aggregated and allocated to textual expression rather than to expression in layout: this leaves for rendering as layout only a modified RST-tree, communicatively equivalent to a simple RST joint relationship.

quantity of material included in the main-line, then we can continue the process to see if further information can be added. Here the remaining elaborations come into question. Of these, only [12–31] has substantial internal varied structure: segments [7–11] and [37–44] are essentially listings of information and are therefore considered good candidates for rendering as separate layout units in their own right.

Following up [12–31] suggests several alternatives. For example, the nucleus of the satellite [12] could be taken alone as representative of the satellite as a whole; this would provide the text fragment:

> . . . Although Unihoc allows many alternatives in how it is played,
> two variants have become dominant. . .

However this leaves the satellite information [13–22] unattached. The alternative taken up in the actual page is to skip over the satellite [12–22] entirely and go straight on with the nucleus [23–31]. This gives the main-line [1–6]+[23–31] as a single consistent 'story' containing the main nucleii of the underlying RST.

Returning to collect side-lines, we can see that the entire satellite [12–22] is taken as a single textblock with the summary 'All forms are possible' as header: this is just a rephrasing of the nucleus of that textblock. The remaining toplevel elaborations also form textblocks with identifying captions—although in their cases the captions identify the kind of elaboration that is at issue: an elaboration of related events [37–44], of necessary equipment [32–36], and of the rules of the game [7–11]. The captions are considered necessary to mark the functions of the individual textblocks because these have now been separated from the possibility of any linguistic indication of their relation to the content of the main-line such as might be indicated, for example, by explicitly marked thematic development or discourse connectives (cf. Section 2.1).

These divisions among mainline and sidelines give the basic decomposition of the page into the layout units summarized in Figure 15 above, as well as attributions of relative centrality (corresponding to nuclearity in the RST structure) and inter-connectedness. This therefore gives the basic range of information necessary for a layout structure as described in Section 2.3.

4.2 Selection of layout styles

The second stage associates the layout units constructed in the first stage with layout styles: examples of textual layout styles are enumerated lists, itemized lists, footnotes, margin notes, headings, captions, etc.[13] These styles can also introduce further decomposition into layout *elements*: for example, an itemized list introduces individual items. Such styles are considered for individual layout units because they generally involve layout decisions that hold for all of the elements that they include. Thus an itemized list may introduce its items by bullets and separate them into paragraph-like blocks, or may choose to run them

13 Examples of *graphical* layout styles, i.e., the resources that can be used for expressing relational information graphically, were given in our discussion of visualization in Section 2.2.

together as a single textblock. The selection of a layout style may also enforce linguistic uniformity on the content that is expressed: for example, rendering items of a list all as nominal phrases or all as sentences, as considered appropriate. A further line of constraint here is that, if two segments are discriminated in the layout structure and its realization, then those segments should also be functionally distinct in the RST-structure. In the Unihoc page, apart from the prominence that should be accorded to the main-line textblock, there is no clear relationship of similarity between any of the remaining decomposed units. Therefore we would expect either the same style to be used for all of them (for a 'quieter' page layout) or differing styles to be used for each (for a 'louder' page layout).[14] Probably unsurprisingly given its intended audience and function, the page as published adopts the latter option yielding the diversity that we saw in Table 4 above.

At present, we consider the motivations for distinct textual styles only very straightforwardly: if there are diverse RST relations present in the content corresponding to a layout unit, then we favour generating running text for that content; if there is a strong multinuclear RST organization (as, in the present example, the sequence seen in [37–44] or the joints in [7–11]), then we favour an itemized list of some kind. Exhaustive presentation combined with strict sequencing lead us to favour an enumeration, less exhaustivity or lack of strict sequencing move us through bullet lists to simple sequences of offset paragraphs. Also relevant here are questions of 'media allocation', and we discuss these more concretely in our presentation of the $DArt_{bio}$ prototype below.

4.3 Distribution of material to layout elements

Material has still to be distributed when layout units themselves have substructure introduced by a layout style: as in, for example, an itemized list. This occurs in several places within the Unihoc page. In the case of the 'Equipment' segment, for example, we have the correspondences indicated in Table 5. The 'Events' segment is similar, with enumeration by date followed by a textblock as the enumeration trailer. In both cases, styles are selected for the overall layout unit and this is enforced for all of the constituent layout elements. This is also used to account for the consistent selection of particular grammatical constructions across a list: thus, the Unihoc rules are presented as a flowing sequence of bulleted items where each is a simple

We separate this out as a separate component of the layout process because it is *not* the case that the distribution is in all cases straightforwardly given by the divisions of the RST tree. In fact, there is a rather complex relationship due to the fact that the rhetorical information to be expressed can still be allocated *either* to the layout resources *or* to the linguistic resources (or both). Space precludes further discussion of the issue here however.

14 Thanks to Judy Delin for suggesting the 'volume control' metaphor.

Table 5
Distribution of information within a layout unit to layout elements depending on the selection layout style.

RST configuration	*rendered as*	*layout style*
joint([32],[33],[34])	↘	enumerate: by number
		[32] ↘ numbered item
		[33] ↘ numbered item
		[34] ↘ numbered item
enablement([35],[36])	↘	enumeration-trailer
		[36] ↘ trailer

4.4 Conclusion

Modularizing the layout construction process as just described provides a general structure for the algorithm investigated further in the DArt$_{bio}$ prototype. Moreover, even though with the extent of our initial study, the process of rendering an RST structure in terms of a page layout remains highly non-deterministic, we can investigate differing layouts experimentally by systematically ‘turning off’ alternatives that are not currently under consideration in order to evaluate the results. It is also worth noting here that many fine-grained decisions for the microtypography of real publications (e.g., what typeface to select, whether text in a layout unit is typeset ragged or justified, etc.) are often not motivated from the communicative-functional intentions as captured by the RST-analysis of single texts, but are instead fixed by ‘higher level’ decisions concerning a magazine’s intended style and feel (cf. Reichenberger et al. (1996) for further discussion). We mention in the paper’s conclusion a new project where we are exploring empirically well-founded ways to reduce the nondeterminacy of our transformation process and to include more fine-grained microtypographical decisions within the purview of the model. First, however, we turn to the particular, application-restricted instantiation of the the approach to layout given in this section as it was applied within DArt$_{bio}$.

5 Page generation within the prototype information system

5.1 Scenario

As we mentioned above, prior to our DArt$_{bio}$ experiment there had already been studies of interface design appropriate for supporting users’ access to the art and art history domain knowledge acquired by the Editor’s Workbench. One usage scenario commonly suggested was that the information system should provide useful ‘starting-off points’, or overviews, of the information maintained in the knowledge base so that a user could shape their subsequent information requests more knowledgeably: it is rarely the case that a user knows their complete information need beforehand—this need, and hence the corresponding information retrieval queries, develops during interaction with the information system as more information is acquired. Our examples in this section build on this and concern possible answers provided by the system when asked a question concerning the spread of the Bauhaus movement. This is based on the particular discussion of interface design and the kinds of interaction that a multimedia information system should support pursued in Kamps et al. (1996).

In this scenario it is presumed that the user is investigating the question ‘How did the Bauhaus spread to the United States?’. The user is seeking and exploring information as provided by the information retrieval component of the Editor’s Workbench. During this preparatory phase of the ongoing multimodal information retrieval dialogue, the user specifies constraints for the selection of lists of artists and source articles concerning those artists. This interaction is very similar to the information retrieval stage of the ‘SPEAK!’ system—another system developed using the Editor’s Workbench but concentrating on spoken dialog (Teich et al., 1997). The retrieval interaction of 'SPEAK!' where users retrieve instances of particular artists has been described in detail and evaluated for distinct user types in Stein (1998) and so we will not discuss this further here. We focus instead solely on the DArt$_{bio}$ system's *presentation* of retrieval results, assuming attention in the dialogue to have been restricted to precisely that particular set of candidate artists just retrieved by the user.

5.2 Abstract page specification

Given as input the selection of artists made by the user during the previous turn, or 'conversational move' of the interaction, page synthesis in DArt$_{bio}$ begins by constructing an RST-like structure for the information to be presented just as in non-multimodal, text planning-based systems. In the DArt$_{bio}$ context, however, the resulting rhetorical structure is interpreted as a preliminary description of the communicative goals and structure of the entire page and not just of an individual text—although, as we suggested in Section 3 above, prior to further information chunking, the structure could well *be* a single text. Although, in general, we would favour using traditional text planning techniques as commonly adopted in NLG (e.g., (Moore and Paris, 1993)) in order to construct the initial RST structure, this was not necessary for the particular application scenario targetted in DArt$_{bio}$—the genre constraints already combine to restrict the required text structures and contents sufficiently for generation. Moreover, once combined with our aggregation techniques, there was little left for a full-blown RST-planner to do.[15] As a consequence, the initial RST-structure constructed in DArt$_{bio}$ could equally well be seen as a structured page template annotated by rhetorical relations.

The initial RST-structure is created by the NLG component in response to genre features prespecified particular for the purposes of the DArt$_{bio}$ prototype, thereby 'customizing' the general text generator (cf. Section 2.1). These generic constraints predispose the system to pursue presentations of evidence for assertions made and, at almost any excuse, to provide short biographies of any artists mentioned as additional background. These design decisions were themselves based on results of previous user studies investigating the browsing strategies preferred by users when engaged in various types of information seeking tasks (see here particularly: (Belkin et al., 1995)). Thus, the 'generic response' for any kind of 'how'-question posed by the user is taken to be of the form: 'Method' followed by 'Instances'. The method gives the means by which, in this case, Bauhaus methods spread, and the instances provide a set of concrete cases: i.e., the particular artists who moved to the States and taught Bauhaus methods there. An initial 'RST-like' structure corresponding to these constraints

15 We emphasize that this is solely a consequence of the restricted nature of our prototype system and is not intended to suggest that RST planning is unnecessary: we see our work as augmenting current work on text planning, not as competing. In this work, we were particularly concerned with using the *output structures* of planning components based on rhetorical relations as possible input forms for page design and, as a consequence, are not at present concerned with how exactly such RST-like structures are being constructed. We expect that successor systems will *combine* both the more 'top-down' orientation of RST planning and the 'bottom-up' orientation of data-driven aggregation that we rely on in DArt$_{bio}$.

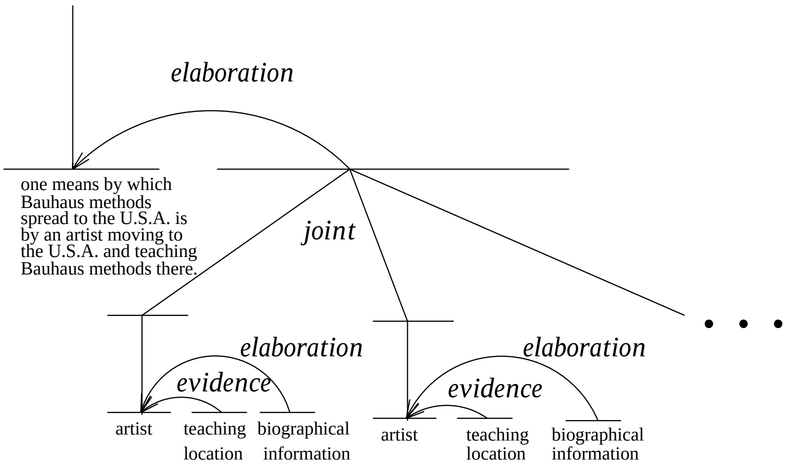


Figure 16
Initial simplistic RST-like structuring of the contents of a potential page

is shown in Figure 16: the initial nucleus is the method, which is then further elaborated by the set of instances found by the information retrieval component; the members of this set are related by the RST 'joint' relation. Each of the artist instances comes ready associated with the dates and locations of teaching in the USA (due to the generic 'present evidence' constraint in the genre specification) and a set of facts constituting a potted biography (corresponding quite directly to a subgenre of biography such as that shown in in Section 2.1). The former information is added by an RST 'evidence' relation, the latter by an RST 'elaboration' relation.

5.3 Two alternative layouts for two contrasting communicative moves

We now show the ways in which DArt$_{bio}$ can render this basic informational content of the page by considering two contrasting layouts that express differing intended emphases in the response. The first alternative puts together the component results used as examples in Section 2 above; the second shows a different path. Moreover, for concreteness in the remainder of the discussion, we focus on page generation for a particular set of artists selected by the information retrieval component in the immediately preceeding communicative move: Gropius, Hilberseimer, Anni Albers, Josef Albers, and Breuer; again we emphasize that the names and information presented here are for *illustrative purposes only*. The initial RST-like structure generated by applying the generic constraints of the DArt$_{bio}$ 'genre' to these five instances is then as already given in Figure 16 but with the actual artists and extracted facts replacing the members of the joint relation that functions as satellite to the main nucleus. It is then the responsability of the DArt$_{bio}$ information presentation components to render this information as a coherent page.

As summarized in Section 4, the first stage of rendering is to seek a mainline for the page. This is substantially simpler in the DArt$_{bio}$ cases than with the extensive page layouts investigated in our empirical study: in fact, there is little option until the RST joint relation is reached and the possibility of aggregation must be considered. Aggregation itself then applies by successively grouping more material from the related text segments: first the nucleii alone are considered, then the nucleii plus one satellite, then plus two satellites, etc. In the present case, just considering the nucleii alone would produce a rather minimal response made up of just the list of artists since in the present case there is no further RST substructure. The dependency lattice constructed on the basis of these nucleii then tells us nothing more than that the entities in question all share the property of being artists. This could be generated by expressing the top-level 'elaboration' RST relation as a "For example, . . . " or "This included the artists:. . . " phrase. Although coherent, we will assume for present purposes that this does not yet exhaust the desired content-quantity for the communicative move and that more information may be included.

Expanding further means that aggregation must be applied to the nucleii together with one or more of the reoccuring satellites: this makes visible either the evidence RST relation or the elaboration relation (or both, although for simplicity here we will consider only stepwise expansion). Although it is obvious in this case from the RST diagram that the two possibilities—nucleii+evidence and nucleii+elaboration are both symmetrical possibilities— this is not the case in general and so the selection of options needs to be motivated. This we can do by again considering an aggregation lattice in order to draw out the most expressive generalizations. In the present case, there is no difference: both 'evidence' and 'elaboration' label the same node in the lattice and so cannot be differentiated according to how they segment the data. We then make the final choice depending on which aggregation best meets the declared communicative goal for the page: that is, whether we want a page providing supporting evidence or a page that is content to elaborate details. Each page conforms to a distinct, but related, subgenre of biography pages. In the remainder of this section, we show the consequences of following each of these possibilities: the first subsection following takes the evidence aggregation, the second takes the elaboration aggregation.

5.3.1 A page presenting evidence When we take the nucleus+evidence combination as the aggregation of the data set to consider, we obtain a set of facts for expression that is exactly equivalent to the data set that we used as an example in Section 2.2: i.e., a range of 'teaching facts' (where, when, who) concerning the selected set of artists. We then need to aggregate just on this data set to see if there is a more concise presentation strategy available. This is given by the subsumption structure of the corresponding aggregation lattice, which was given in Figure 6. This shows that there are indeed regularities available for exploitation in this data set and so the layout process can sensibly factor out the nucleus+evidence combination as a single layout unit. The consequences of this are shown as as a modified RST structure in Figure 17.

This 'aggregated RST' structure provides a suitable starting point for a page of information that expresses the information content required without being unduely tied to the specific structure of the generic constraints pre-specified for the generation component. The structure factors out commonalities so that information from lower leaves of the original tree has been placed at higher branches and offers an alternative, more richly structured presentation plan that is responsive to the instantial regularities found in the retrieved data. This corresponds almost directly to the final layout structure: the corresponding layout structure then consists of the units *Method*, *Examples-summary*, $Bio_{Gropius}$, $Bio_{Hilberseimer}$, and so on. In addition, for the purposes of $DArt_{bio}$, we specify as generic contraints that the biographies should also present (as ever, if the information is available to the Editor's Workbench) pictorial background information for the artist. And, finally, the 'movement' of the aggregated block to form a consolidated layout unit separate from the individual biographies motivates well the inclusion of *Header* segments for each, now isolated, biography block. The completed layout structure then records the constraints that the biographies are all similar to each other in status and are collectively less nuclear than both the *Example-summary* and the *Method*. This means that biographies

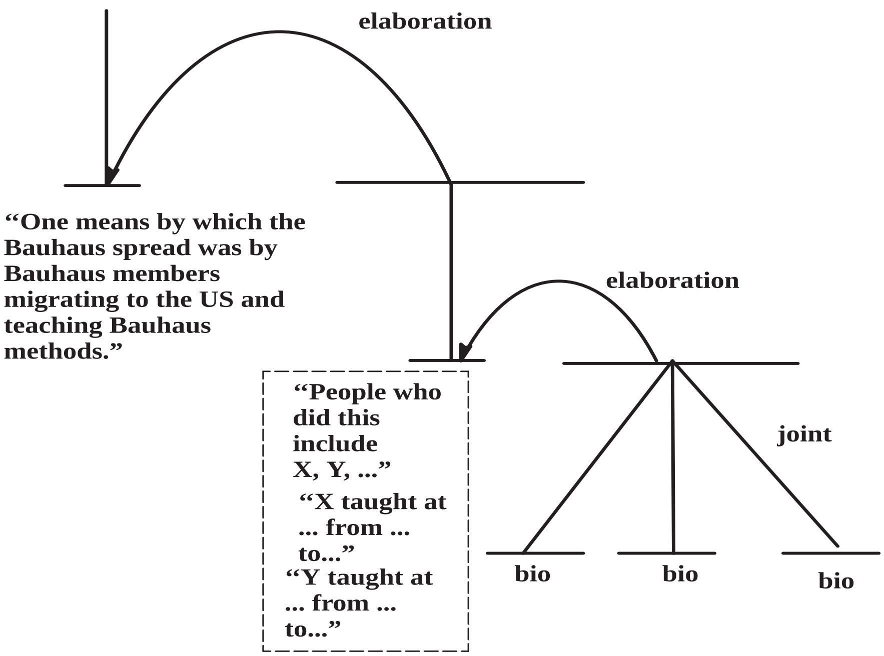


Figure 17
RST-like page structuring after aggregation

should all be laid out with equal prominence and should not be differentiated in type, whereas the overview of times and places should be more prominent and relatively distinct from the set of biographies. The layout style of the individual biographies is then made uniform regardless of possible minor differences that arise due to differences in the particular information to be expressed.

These constraints feed into the next stage, where the layout units indicated by the aggregated RST structure are all allocated layout styles. Since the planned layout units do not themselves specify how their components are to be realized—that is, it is still relatively free whether subtrees are expressed textually or as diagrams or as some combination, this stage includes for us the question of media allocation. In general, this needs to be a motivated selection of presentation mode using criteria such as those discussed by, for example, Arens and Hovy (1990) and (Feiner and McKeown, 1993). Within DArt$_{bio}$ we base our media allocation decision primarily on the results of potential aggregation. Unless a clear preference is pre-determined—for example, by information being available in the Editor's Workbench only in a single modality, e.g., pictures of artists maintained in the knowledge base—the style allocation process examines the constructed dependency lattices in order to see how many co-varying dimensions of potential aggregation are present in the data set corresponding to each layout unit. The more simultaneous dimensions of regularity present, the more likely it is that a diagram will be the more perspicuous representation.[16] In the present case, we can contrast the aggregation possibilities for the Examples-summary layout unit with those for the biography layout units. The former shows several co-varying dimensions while the latter show rather few (in fact, simple progression in time with accompanying changes in location or state revolving around a single individual—the artist forming the subject matter of each biography, and a few repeated activities such as 'studying', etc.). The individual biographies are therefore considered good candidates for textual expression, whereas the Example-summary unit is not.[17] It is also worth noting, however, that quite acceptable and useful text fragments can be generated by dropping some of the less significant dimensions of aggregation: such texts may then be presented in *addition* to a diagram either as a caption (cf. Mittal et al. (1998)) or to draw attention to other aspects of the data (cf. Kerpedjiev et al. (1997)).

Following media allocation, the information to be expressed in the each layout unit is passed to the rendering component considered appropriate—i.e., either to automatic visualization or to the generation component. Two possible results of visualizing the Example-summary unit were shown in the timeline diagrams of Figure 7 in Section 2.2, while the individual biographies correspond to texts shown generated in Figure 3 in Section 2.1.

---

16 This is, of course, only a heuristic and could easily require alteration or parameterization—for example, with different communicative purposes or when different output modalities (e.g., only spoken output) are possible.

17 The problematic nature of realizing several co-varying dimensions of regularity in textual form can readily be seen by attempting to generate a text corresponding to the Example-summary layout unit. We have shown this concretely in Bateman et al. (1998) where we present texts expressing the same aggregations as those represented graphically in Figure 7 above.

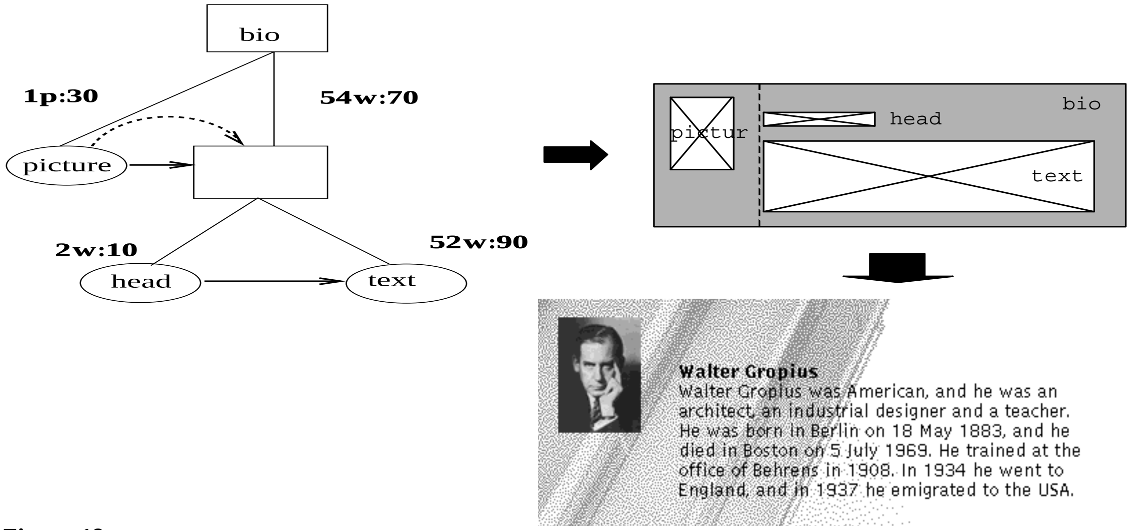


Figure 18
Segment of the layout structure for the layout unit corresponding to a single biography, together with the resulting segment layout and the filled-in page

Finally, the layout structure complete with its leaves filled in with these texts and a diagram is passed to the rendering engine outlined in Section 2.3 in order to situate the layout units on the page in some way that satisfies the collected constraints recorded in the layout structure. We illustrate this here just the layout for a single biography; a similar result for the full page is shown in our next example: Figure 18 shows on the left the segment of the layout structure for a single biography element—including both a header as summary and a picture of the artist as background—and on the right, the corresponding part of the page.

5.3.2 A page presenting elaboration When instead we take the nucleus+elaboration combination, as appropriate for a page that is simply presenting more background information, we obtain a different sets of facts to consider for potential aggregations. This set is equivalent to the union of the individual inputs for the generation process as illustrated for a single artist in Figure 2 above: i.e., the combined information making up the content of the biographies for our selected artists. The aggregation lattice here is more complex but again indicates that there are (limited) possibilities for aggregation: those aggregations that cluster the data most effectively are the presence of places and times of birth and death, and the specification of professions—virtually all of the biographies include this information.

Taking this as a layout unit results in a similar aggregated RST structure to that of the previous example, although the particular details of the aggregated data set are quite different. The media allocation decisions are therefore also as before and a complete page such as that shown in Figure 19 can be generated. Here we can note that the individual biographies are as described above, but the particular automatically generated diagram concerns different properties of the

Figure 19
Example generated page from the art history system

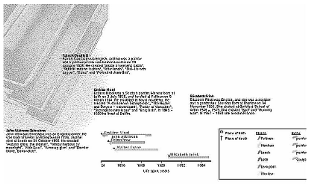


Figure 20
Another example generated page showing a different solution to the layout constraint resolution process

artists selected. We can also see that the generic constraint that a photograph of the artist be presented has not been satisfied in all cases, but this has not affect the general style decisions applied to the set as a whole due to their type identity within the layout structure.

5.3.3 Summary We have now shown how our abstract specification for decomposing a rhetorical structure for a page into a layout structure has informed and guided the design of the concrete design steps automated within the DArt$_{bio}$ prototype. The prototype can produce similar pages to that shown in Figure 19 for any set of artists decided by the user in the course of their interaction with the system. The presentation environment is implemented in Smalltalk, the visualization and layout engines in C, the text generation component in Common Lisp; page generation is in real-time. The final result is considered a generic 'informative' and cooperative response to a user's question, presenting slightly more information than strictly requested. The overhead that the user might not actually want this information is then balanced by the use of a page to set out the information in a way that does not commit the user to reading it all. This feature brings out a striking similarity between page layout and hypertext, which we will mention again in the conclusion below.

The layout process is flexible within the limits that are compatible with the requirement of presenting a fixed amount of information on a page, but we have not explored within DArt$_{bio}$ those additional aspects of design that would be required of a final system: for example, if the user selects 100 artists instead of half a dozen, then this information cannot reasonably be accomodated in the style of page presentation shown here. Nor would it be reasonable just to present a lengthy sequence of individual pages each conforming to our presentation style. But since this kind of overview presentation is already covered well by the Editor's Workbench and the general visualization tools, providing it involves little more than a switch of selected presentation styles when the system notices that it has too much information for expression. Note also that the resemblance of the example page in Figure 19 to a column-based organization is also simply a by-product of the constraint resolution process and its attempt to fit 7 boxes (the 5 biography layout units, the diagram, and the diagram's key) into a limited space. As shown in the thumbnail alternative layout in Figure 20, with more space to 'play with' other potential solutions are also suggested.

We can also reiterate here for the particular pages generated the point we made for the general case in Section 4.4 above: that is, that many fine-grained typographical decisions will not be motivated directly from the layout structure as is but will need to be set according to design goals of the document class being generated. For example, for DArt$_{bio}$ we preset the typeface used and the 'neutral' size, the page background (a corner of the Bauhaus building), and the justification options (right ragged) for the biography layout units. These 'preset constraints' then fed into the layout process just as any other derived constraint and so can be changed for experimentation, but are not motivated from the rhetorical structure. And, as always, the final pages generated can always be criticised on numerous grounds, just as can the text generated. We consider such criticism, however, as a positive motivating force and evaluation that can be used for the next round of experimentation.

6 Conclusions and future work

The last five years has seen rapid growth in the awareness of the importance of consistent and function 'style' selections: notions of style sheets from professional publishing have made their way into the mainstream of web-based document design (with, e.g., Cascaded Style Sheets: CSS) and rendering (e.g., the Document Style Semantics and Specification Language: DSSSL, ISO / IEC 10179:1996). Professional bodies concerned with the closely related theme of 'Information Design'—the application of "processes of design (that is, planning) to the communication of information (its content and language as well as its form)" (Waller, 1996)—have also grown considerably both in number and membership. Our work on DArt$_{bio}$ and its supporting technology interfaces directly with these developments, establishing a bridge between options for consistent microtypography at one end through to high-level communicative goal-based text design at the other.

6.1 Placement of the prototype and consequences drawn

DArt$_{bio}$ establishes the 'page' as a *single, coherent, communicatively-motivated unit of information presentation that is fully designed in terms of both its content and form*. It extends the tradition established in previous work on multimodal presentation systems where the essential leap was the recognition that graphical design is as purposive as verbal activity: approaches to language generation based on communicative goals are therefore naturally applicable to graphical presentation (cf., particularly, the WIP (André et al., 1993; André and Rist, 1993) and COMET (Feiner and McKeown, 1993) systems). Moreover, since we include 'layout' as one kind of graphical design, we also consider decisions concerning text formatting and segmentation to be an integral part of the information presentation process—cf. earlier separate work on the formatting of text according to rhetorical or communicative function (Sefton, 1990; Marks, 1991; Roth and Mattis, 1991; Hovy and Arens, 1991; White, 1995; Pascual, 1996), and automated layout (Feiner, 1988; Graf, 1995). Our treatment of these phenomena as aspects of a single, integrated presentation generation process is also naturally consistent with more recent work on the appropriate combination of media presenting partially overlapping information(Mittal et al., 1995; Fasciano and Lapalme, 1996; Kerpedjiev et al., 1997)).

DArt_{bio} has allowed us to investigate many aspects of the interrelationship between texts, layout, and diagrams and has provided motivation for an extensive architectural synthesis across information presentation modalities. We have found that a variety of generation decisions previously considered in isolation are more usefully treated as closely related phenomena—several of these synergies follow directly from our treatment of layout as a form of diagram and, hence, as meaning-bearing. Among the most significant of these interconnections are the following.

Diagram design and text design Many of the decisions that need to be made for an effective diagram are also found in the construction of an effective text. We have shown how one kind of decision, the determination of aggregation possibilities, could be made from the same formal mechanism for both texts and diagrams. This is then essential for a coherent page where representatives of both modalities might be present side-by-side (as in our biography pages above): diagram and text can mutually reinforce one another by applying common information groupings.

Layout and text Many of the decisions that need to be made for segmenting a text effectively—e.g., into thematic paragraphs, into rhetorically related segments—are also precisely the decisions that need to be made to produce a coherent layout structure. This then entails potential trade-offs between layout and text: segmentation and grouping information may be expressed in language, in layout, or some mixture.[18] By treating such layout and language expressions as arising from the same source, we expect that potentially costly constraint-resolution at a local level (i.e., between arbitrary segments) can be avoided, or reduced, by consistent layout-language decisions for the layout structure of the page as a whole.

Layout and hypertext We can also note that, although hypertext and the web was not a focus of DArt_{bio}, it is clear nevertheless that it has many points of similarity with web-based information presentation projects such as ILEX (Oberlander, Mellish, and O'Donnell, 1997) and PEBA-II (Milosavljevic and Dale, 1996) in that multimedia pages form the basic units of information generated. Interestingly, the decisions as to how much information is presented on a single webpage and how much is to be placed within a single layout unit also appear to resemble each other. We have shown above that one of the possible consequences

18 This is the main emphasis of the recent British EPSRC-funded project ICONOCLAST ('Integrating Constraints on content, layout and style': ITRI, University of Brighton. Cf. http://www.itri.brighton.ac.uk/projects/iconoclast/).

of layout is the avoidance of a commitment to specific rhetorical structures and dependencies. Text with layout would then share the claimed advantage of hypertext over fully linearized text that text structure and text structure problems do not need to be fully resolved by the generation process. We are now investigating this issue more closely, particularly with respect to the question of how the navigational function of layout, hyperlinks between related pages, and explicit discourse markers and textual deixis overlap (Reichenberger et al., 1996).

6.2 Directions and Future work

As natural language generation moves away from producing disembodied texts and interacts more closely with general information presentation, we expect that the kinds of design decisions motivated in DArt$_{bio}$ will become standard. Our study of the relation between the communicative structure of a page of information and its coherent layout demonstrates that this should be treated as an integral and complex part of the generation process and can in no way be treated as a final piece of postprocessing.

Our experiences with DArt$_{bio}$ have led to two main areas where we are now focusing future development and research: one theoretical, one practical.

First, our finding that layout is a rich communicative resource in its own right, with significant points of potential overlap with rhetorical structure planning, hypertext design, and explicit linguistic resources for discourse structuring, highlights that further empirical work concerning a broader range of documents utilising layout is now essential. As we drew attention to above, our proposed decomposition process for mapping RST structures to layout structures is highly nondeterministic with many choices that cannot currently be motivated on any grounds—functional or otherwise. We therefore need to refine the model of layout we have developed, by examining both a greater quantity of professionally produce pages and a greater variety of page types. To this end, a new project, *GeM*—'Genre and Multimodality'—has recently been started (Allen, Bateman, and Delin, 1999): here one aim is to provide additional constraints on the layout process drawn from an explicit representation of possible document types resembling the approach to genres describes in Section 2.1 above but extended to include multimodal aspects as well as features of the situation of production. Here again, analogously to the procedure described in Section 3 above, we will express the final results of our corpus study as a computational specification.

Second, the functioning prototype has helped refine our technology for automatic visualization and layout to the point where it can be applied commercially. These components therefore provide the backbone of the new generation of automatic information presentation services offered by recently formed German company Intelligent Views (http://www.i-views.de). Presentation of data overviews with automatic generated diagrams, as well as detailed document layout, are currently offered.

Acknowledgements

The work in this paper was carried out in the context of the KOMET-PAVE multimedia page generation experiment (GMD-IPSI) KOMET ('Knowledge-oriented production of multimodal documents') and PAVE ('Publication and advanced visualization environments') were two departments of the German National Research Center for Information Technology's (GMD) institute for Integrated Publication and Information Systems (IPSI) in Darmstadt that cooperated closely for this work. The authors of the present paper would therefore like to thank all the members of those departments who contributed, and particularly Lothar Rostek, Melina Alexa, Elke Teich, Wiebke Möhr and Christoph Hüser. A preliminary discussion presenting more background and explanatory material for Section 3 appeared as Reichenberger et al. (1996); we additionally thank Klaas Jan Rondhuis for his contribution to that work.

References

Allen, Patrick, John Bateman, and Judy Delin. 1999. Genre and layout in multimodal documents: towards an empirical account. In *Proceedings of the AAAI Fall Symposium on Using Layout for the Generation, Understanding, or Retrieval of Documents*, Cape Cod, Massachusetts. American Association for Artificial Intelligence.

André, Elisabeth and Thomas Rist. 1993. The design of illustrated documents as a planning task. In Mark T. Maybury, editor, *Intelligent Multimedia Interfaces*. AAAI Press / The MIT Press, Menlo Park (CA), Cambridge (MA), London (England), pages 94 – 116.

André, Elizabeth, Wolfgang Finkler, Winfried Graf, Thomas Rist, Anne Schauder, and Wolfgang Wahlster. 1993. WIP: the automatic synthesis of multimodal presentations. In Mark T. Maybury, editor, *Intelligent Multimedia Interfaces*. AAAI Press / The MIT Press, Menlo Park (CA), Cambridge (MA), London (England), pages 75–93.

Arens, Yigal and Eduard H. Hovy. 1990. How to describe what? towards a theory of modality utilization. In *The Twelfth Annual Conference of the Cognitive Science Society*. Lawrence Erlbaum Associates, Hillsdale, New Jersey, pages 487–494.

Bateman, John, Thomas Kamps, Jörg Kleinz, and Klaus Reichenberger. 1998. Communicative goal-driven NL generation and data-driven graphics generation: an architectural synthesis for multimedia page generation. In *Proceedings of the 1998 International Workshop on Natural Language Generation*, pages 8–17. Niagara-on-the-Lake, Canada.

Bateman, John, Elke Teich, and Melina Alexa. 1998. Generic technologies for selective information presentation: an application of computational linguistic methods. In Peter Fankhauser and Marlies Ockenfeld, editors, *Integrated Publication and Information Systems: 10 years of research and development*, number ISBN: 3088457-968-1. GMD, Forschungszentrum Informationstechnik, Sankt Augustin, Germany, pages 237–258.

Bateman, John A. 1997. Enabling technology for multilingual natural language generation: the KPML development environment. *Journal of Natural Language Engineering*, 3(1):15–55.

Bateman, John A. and Cécile L. Paris. 1991. Constraining the development of lexicogrammatical resources during text generation: towards a computational instantiation of register theory. In Eija Ventola, editor, *Recent Systemic and Other Views on Language*. Mouton, Amsterdam.

Bateman, John A. and Elke Teich. 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.

Belkin, Nicholas J., Colleen Cool, Adelheit Stein, and Ulrich Thiel. 1995. Cases, scripts, and information seeking strategies: On the design of interactive information retrieval systems. *Expert Systems and Applications*, 9(3):379–395. (Also available: "Arbeitspapiere der GMD", No. 875, Sankt Augustin, Germany 1994).

Corio, Marc and Guy Lapalme. 1998. Integrated generation of graphics and text: a corpus study. In M.T. Maybury and

J. Pustejovsky, editors, *Proceedings of the COLING-ACL Workshop on Content Visualization and Intermedia Representations (CVIR'98)*, pages 63–68, Montréal, August.

Dale, Robert, Maria Milosavljevic, and Jon Oberlander. 1997. The web as dialogue: the role of natural language generation in hypertext. In *Proceedings of the AAAI'97 Spring Symposium: Natural Language Processing for the World Wide Web*. American Association for Artificial Intelligence.

den Ouden, J.N., C.H. van Wiik, J.M.B. Terken, and L.G.M.Noordman. 1998. Reliability of discourse structure annotation. In *IPO Center for research on user-system interaction, Annual Progress Report*, number 33. Technische Universiteit Eindhoven, pages 129–138.

Fasciano, Massimo and Guy Lapalme. 1996. PostGraphe: a system for the generation of statistical graphics and text. In *Proceedings of the 8th. International Workshop on Natural Language Generation (INLG '96)*, pages 51–60, Herstmonceux, England, June.

Feiner, S. 1988. A grid-based approach to automating display layout. In *Proceedings of the Graphics Interface*, pages 192–197, Los Angeles, CA. Morgan Kaufman.

Feiner, Steven K. and Kathleen R. McKeown. 1993. Automating the generation of coordinated multimedia explanations. In Mark T. Maybury, editor, *Intelligent Multimedia Interfaces*. AAAI Press / The MIT Press, Menlo Park (CA), Cambridge (MA), London (England), pages 117–138.

Ganter, B. and R. Wille. 1996. *Formale Begriffsanalyse—Mathematische Grundlagen*. Springer-Verlag, Berlin / Heidelberg.

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, California. ACM.

Hasan, Ruqaiya. 1978. Text in the Systemic-Functional Model. In Wolfgang Dressler, editor, *Current Trends in Text Linguistics*. de Gruyter, Berlin, pages 228–246.

Hovy, Eduard H. 1993. Automated discourse generation using discourse relations. *Artificial Intelligence*, 63:341 – 385.

Hovy, Eduard H. and Yigal Arens. 1991. Automatic generation of formatted text. In *Proceedings of the 8th. Conference of the American Association for Artifical Intelligence*, pages 92–96, Anaheim, California.

Hüser, Ch., K. Reichenberger, L. Rostek, and N. Streitz. 1995. Knowledge-based editing and visualization for hypermedia encyclopedias. *Communications of the ACM*, 38(4):49–51, April.

Kamps, Thomas. 1997. *A constructive theory for diagram design and its algorithmic implementation*. Ph.D. thesis, Darmstadt University of Technology, Darmstadt, Germany.

Kamps, Thomas. 1998. A constructive approach to automatic diagram design. In Peter Fankhauser and Marlies Ockenfeld, editors, *Integrated Publication and Information Systems: 10 years of research and development*, number ISBN: 3088457-968-1. GMD, Forschungszentrum Informationstechnik, Sankt Augustin, Germany, pages 223–236.

Kamps, Thomas, Christoph Hüser, Wiebke Möhr, and Ingrid Schmidt. 1996. Knowledge-based information acess for hypermedia reference works: exploring the spread of the bauhaus movement. In Maristella Agosti and Alan F. Smeaton, editors, *Information retrieval and hypertext*. Kluwer Academic Publishers, Boston / London / Dordrecht, pages 225–255.

Kamps, Thomas and Klaus Reichenberger. 1995. A dialogue approach to graphical information access. In W. Schuler, J. Hannemann, and N. Streitz, editors, *Designing user interfaces for Hypermedia*. Springer-Verlag, Berlin, pages 141–155.

Kerpedjiev, Stephan, Guiseppe Carenini, Steven Roth, and Johanna D. Moore. 1997. Integrating planning and task-based design for multimedia presentation. In *Proceedings of IUI-97*.

Mackinlay, Jock D. 1986. *Automatic design of graphical presentations*. Ph.D. thesis, Computer Science Department, University of Stanford, Stanford, CA.

Mann, William C. and Christian M.I.M. Matthiessen. 1985. Demonstration of the Nigel text generation computer program. In James D. Benson and William S. Greaves, editors, *Systemic Perspectives on Discourse, Volume 1*. Ablex, Norwood, New Jersey, pages 50–83.

Mann, William C. and Sandra A. Thompson. 1986. Relational propositions in discourse. *Discourse Processes*, 9(1):57–90, January-March. Also available as ISI / RR-83-115.

Marks, J.W. 1991. A formal specification scheme for network diagrams. *Journal of Visual Languages and Computing*, 2(4):395–414.

Martin, James R. 1992. *English text: systems and structure*. Benjamins, Amsterdam.

Matthiessen, Christian M.I.M. and John A. Bateman. 1991. *Text generation and systemic-functional linguistics: experiences from English and Japanese*. Frances Pinter Publishers and St. Martin's Press, London and New York.

Maybury, Mark T. 1995. Research in multimedia and multimodal parsing and generation. *Artificial Intelligence Review*, 9(2 / 3):103–127.

McKeown, Kathleen, Jacques Robin, and Karen Kukich. 1995. Generating concise natural language summaries. *Information Processing and Management*, 31(5):703–733.

Milosavljevic, M. and R. Dale. 1996. Text generation and user modelling on the web. In *Proceedings of* User Modelling for Information Filtering on the World Wide Web *workshop at the* User Modelling '96 *conference*.

Mittal, V., S. Roth, J. Moore, J. Mattis, and G. Carenini. 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.

Mittal, Vibhu O., Johanna D. Moore, Giuseppe Carenini, and Steven Roth. 1998. Describing complex charts in natural language: a caption generation system. *Computational Linguistics*, 24(3):431–468, September.

Moore, Johanna D. and Cécile L. Paris. 1993. Planning texts for advisory dialogs: capturing intentional and rhetorical information. *Computational Linguistics*, 19(4):651 – 694, December.

Oberlander, Jon, Chris Mellish, and Michael O'Donnell. 1997. Exploring a gallery with intelligent labels. In *Proceedings of the 4th. International Conference on Hypermedia and Interactivity in Museums*, pages 153–161, September.

O'Donnell, Michael. 1997. Variable-length on-line document generation. In *Proceedings of the 6th European Workshop on NaturalLanguage Generation*, Duisburg, Germany. Gerhard-Mercator University.

Paris, Cécile L. 1993. *User modelling in text generation*. Pinter Publishers, London.

Pascual, Elsa. 1996. Integrating text formatting and text generation. In G. Adorni and M. Zock, editors, *Trends in Natural Language Generation: an artificial intelligence perspective,* number 1036 in Lecture Notes in Artificial Intelligence. Springer-Verlag, Berlin, New York, pages 205–221. (Selected Papers from the 4th. European Workshop on Natural Language Generation, Pisa, Italy, 28-30 April 1993).

Reichenberger, K., T. Kamps, and G. Golovchinsky. 1995. Towards a generative theory of diagram design. In *Proceedings of 1995 IEEE Symposium on Information Visualization*, pages 217–223, Los Alamitos, USA. IEEE Computer Society Press.

Reichenberger, Klaus, Klaas Jan Rondhuis, Jörg Kleinz, and John A. Bateman. 1996. Effective presentation of information through page layout: a linguistically-based approach. Technical Report Arbeitspapiere der GMD 970, Institut für Integrierte Publikations- und Informationssysteme (IPSI), GMD, Darmstadt, January. (Paper presented at the workshop: 'Effective Abstractions in Multimedia, Layout and Interaction', held in conjunction with ACM Multimedia '95, November 1995, San Francisco, California.).

Rostek, Lothar, Wiebke Möhr, and Dietrich H. Fischer. 1994. Weaving a web: The structure and creation of an object network representing an electronic reference network. In C. Hüser and W. Möhr and V. Quint, editor, *Proceedings of Electronic Publishing (EP) '94*. Wiley, Chichester, pages 495 – 506. Special issue of the International Journal of *Electronic publishing–origination, dissemination and design*, Volume 6(4).

Roth, S.F. and J. Mattis. 1991. Automating the presentation of information. In *Proceedings of the IEEE Conference on AI Applications*, pages 90–97, Miami Beach, FL. IEEE.

Sefton, Petie 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, November. Batchelor's Honours Thesis.

Southall, Richard. 1992. Presentation rules and rules of composition in the formatting of complex text. In C. Vanoirbeek and G. Coray, editors, *Electronic Publishing '92*, pages 275–290. Cambridge University Press.

Stein, Adelheit. 1998. Usability and assessments of multimodal interaction in the SPEAK! system: an experimental case study. *New Review of Multimedia and Hypermedia*. (Special issue on the evaluation of Information Retrieval and Multimodal systems).

Stein, Adelheit and Ulrich Thiel. 1993. A conversational model of multimodal interaction in information systems. In *Proceedings of the 11th National Conference on Artificial Intelligence (AAAI '93), Washington DC, USA*, pages 283–288. AAAI Press / MIT Press.

Teich, Elke, Eli Hagen, Brigitte Grote, and John Bateman. 1997. From communicative context to speech: integrating dialogue processing, speech production and natural language generation. *Speech Communication*, 21(1–2), Feb.

Waller, Rob. 1996. The origins of the Information Design Association. In *The 1996 Annual Report of the IDA*. Information Design Association.

White, Michael. 1995. Presenting punctuation. In *Proceedings of the Fifth European Workshop on Natural Language Generation*, pages 107–125, Leiden, the Netherlands, 20-22 May 1995. Faculty of Social and Behavioural Sciences, University of Leiden.

Willberg, Hans Peter and Friedrich Forssman. 1997. *Lesetypographie*. Verlag Hermann Schmidt, Mainz.

Wille, R. 1982. Restructuring lattice theory: an approach based on hierarchies of concept. In I. Rival, editor, *Ordered Sets*. Reidel, Dordecht / Boston, pages 445–470.